



TENDÊNCIAS ATUAIS E PERSPETIVAS FUTURAS EM ORGANIZAÇÃO DO CONHECIMENTO

ATAS DO III CONGRESSO ISKO ESPANHA-PORTUGAL
XIII CONGRESSO ISKO ESPANHA

Universidade de Coimbra, 23 e 24 de novembro de 2017

Com a coordenação de

Maria da Graça Simões, Maria Manuel Borges

TENDÊNCIAS ATUAIS E PERSPETIVAS FUTURAS EM ORGANIZAÇÃO DO CONHECIMENTO

ATAS DO III CONGRESSO ISKO ESPANHA-PORTUGAL
XIII CONGRESSO ISKO ESPANHA

Universidade de Coimbra, 23 e 24 de novembro de 2017

Com a coordenação de
Maria da Graça Simões, Maria Manuel Borges

TÍTULO

Tendências Atuais e Perspetivas Futuras em Organização do Conhecimento: atas do III Congresso ISKO Espanha e Portugal - XIII Congresso ISKO Espanha

COORDENADORES

Maria da Graça Simões
Maria Manuel Borges

EDIÇÃO

Universidade de Coimbra. Centro de Estudos Interdisciplinares do Século XX - CEIS20

ISBN

978-972-8627-75-1

ACESSO

<https://purl.org/sci/atas/isko2017>

COPYRIGHT

Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

OBRA PUBLICADA COM O APOIO DE



FLUC FACULDADE DE LETRAS
UNIVERSIDADE DE COIMBRA

2



CEIS 20
CENTRO DE ESTUDOS
INTERDISCIPLINARES
DO SÉCULO XX
UNIVERSIDADE DE COIMBRA

FCT

Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

PROJETO UID/HIS/00460/2013

COORDENADORES GERAIS

Maria da Graça Simões, Universidade de Coimbra, Portugal

Maria Manuel Borges, Universidade de Coimbra, Portugal

COMISSÃO ORGANIZADORA

António Tavares Lopes, Universidade de Coimbra, Portugal

Maria Cristina Vieira de Freitas, Universidade de Coimbra, Portugal

Ana Eva Migueis, Universidade de Coimbra, Portugal

Mário Bernardes, Universidade de Coimbra, Portugal

Ana Catarina Silva, CEIS20, Portugal

Bruno Neves, Universidade de Coimbra, Portugal

Liliana Gomes, Universidade de Coimbra, Portugal

Blanca Rodriguez Bravo, Universidad de León, Espanha

Jose Antonio Frias Montoya, Universidad de Salamanca, Espanha

COMISSÃO CIENTÍFICA

Blanca Rodriguez Bravo, Universidad de León, Espanha

Blanca Gil Urdiciain, Universidad Complutense de Madrid, Espanha

Carlos Guardado da Silva, Universidade de Lisboa, Portugal

Carlos Henrique Marcondes, Universidade Federal Fluminense, Brasil

Carmen Agustín, Universidad de Zaragoza, Espanha

Carmen Caro Castro, Universidad de Salamanca, Espanha

Críspulo Travieso Rodríguez, Universidad de Salamanca, Espanha

Cristina Dotta Ortega, Universidade Federal de Minas Gerais, Brasil

Daniel Martínez Ávila, Universidade Estadual Paulista, Brasil

Dunia Llanes Padrón, Universidad de la Habana, Cuba

Fabio Assis Pinho, Universidade Federal de Pernambuco, Brasil

Fernanda Ribeiro, Universidade do Porto, Portugal

Francisco Javier García Marco, Universidad de Zaragoza, Espanha

Giovana Deliberali, Universidade de São Paulo, Brasil

Isidoro Gil Leiva, Universidad de Murcia, Espanha

Jesús Gascón García, Universidad de Barcelona, Espanha

Jesús Tramullas Saz, Universidad de Zaragoza, Espanha

José Augusto Chaves Guimarães, Universidade Estadual Paulista, Brasil

José Luis Bonal Zazo, Universidad de Extremadura, Espanha

Jose Antonio Frias Montoya, Universidad de Salamanca, Espanha

Juan-Antonio Pastor-Sánchez, Universidad de Murcia, Espanha
Luciana de Souza Gracioso, Universidade Federal de São Carlos, Brasil
Manuela Moro Cabero, Universidad de Salamanca, Espanha
María del Pilar Ortego de Lorenzo Cáceres, Universidad de Extremadura, Espanha
Maria José López Huertas, Universidad de Granada, Espanha
Maria Luisa Alvite Díez, Universidad de León, Espanha
Maria Luíza Almeida Campos, Universidade Federal Fluminense, Brasil
Maria Cristina Vieira de Freitas, Universidade de Coimbra, Portugal
Maria da Graça Simões, Universidade de Coimbra, Portugal
Maria Manuel Borges, Universidade de Coimbra, Portugal
Mariângela Spotti Lopes Fujita, Universidade Estadual Paulista Júlio de Mesquita Filho, Brasil
Maurício Barcellos Almeida, Universidade Federal de Minas Gerais, Brasil
Natália Bolfarini Tognoli, Universidade Estadual Paulista, Brasil
Olívia Pestana, Universidade do Porto, Portugal
Pedro Manuel Díaz Ortuño, Universidad de Murcia, Espanha
Renato Rocha Souza, Fundação Getúlio Vargas, Brasil
Rodrigo Sales, Universidade Federal Fluminense, Brasil
Rosa San Segundo Manuel, Universidad Carlos III, Espanha
Vera Lúcia de Mattos Dodebei, Universidade do Rio de Janeiro, Brasil

SUMÁRIO

INTRODUÇÃO	13
FOUNDATIONS AND MOVING FORWARD: THE HORIZONS AND THE FUTURE OF RESEARCH IN KNOWLEDGE ORGANIZATION <i>Joseph T. Tennis</i>	15
PERSPETIVAS EPISTEMOLÓGICAS EM ORGANIZAÇÃO DO CONHECIMENTO	21
TODOROV, A FILOSOFIA SIMBÓLICA E A ORGANIZAÇÃO DO CONHECIMENTO: SOBRE O TRANSBORDAMENTO DO SIGNIFICANTE DE TESAURO AOS DESAFIOS DA WEB PRAGMÁTICA <i>Gustavo Silva Saldanha</i>	23
CONTRIBUIÇÕES AO ESTUDO DA ORDENAÇÃO DE DOCUMENTOS PARA ARQUIVÍSTICA <i>Clarissa M. S. Schmidt, Renato de Mattos</i>	31
OS DISCURSOS INSTITUCIONAIS SOBRE INTERDISCIPLINARIDADE ANALISADOS POR MEIO DA HERMENÊUTICA E DE CATEGORIAS PMEST <i>Marco Donizete Paulino da Silva, Luciana de Souza Gracioso, Maria da Graça Melo Simões</i>	41
TEORIAS DA LINGUAGEM NOS ESTUDOS DE INDEXAÇÃO: UMA ANÁLISE A PARTIR DOS PERIÓDICOS <i>JOURNAL OF DOCUMENTATION</i> E <i>KNOWLEDGE ORGANIZATION</i> <i>Raquel Luíse Pret, Rosa Inês de Novais Cordeiro</i>	53
A EXPANSÃO DO DOMÍNIO DO ARQUIVO: O CASO DO CSAC – CENTRO STUDI E ARCHIVIO DELLA COMUNICAZIONE. <i>Giulia Crippa</i>	65
UM ESTUDO PRELIMINAR SOBRE A APLICABILIDADE E EXTENSÃO DO MODELO SECI COM O USO DE DISPOSITIVOS DA INTERNET DAS COISAS <i>Antônio José Rodrigues Neto, Renato Souza Rocha</i>	75
ORGANIZAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO: INTERFACES INFORMACIONAIS, COMUNICACIONAIS E LINGUÍSTICAS <i>Nair Yumiko Kobashi</i>	85
ANÁLISE TEÓRICA CONCEITUAL DA TERMINOLOGIA, GARANTIA CULTURAL, MULTICULTURALIDADE E TRANSCULTURALIDADE NA CONSTRUÇÃO DE LINGUAGENS DOCUMENTÁRIAS <i>Ricardo Biscalchin, Érica Fernanda Vitorini, Walter Moreira</i>	97
O DOCUMENTO CIENTÍFICO COMO PRODUTO E COMO OBJETO DA ORGANIZAÇÃO DO CONHECIMENTO <i>Luciana Cortes Mendes</i>	109
ACHADOS TEÓRICO-EPISTEMOLÓGICOS EM TRABALHOS DO GT2-ENANCIB: UMA ANÁLISE BASEADA EM CATEGORIAS <i>Hildenise Ferreira Novo, Bruna Bomfim Lessa dos Santos, Carolina de Souza Santana Magalhães, Daniel Almeida Marins, Ivana Aparecida Borges Lins, Luís Carlos Batista de Jesus, Maíra Salles</i>	119
AS CONTRIBUIÇÕES DA TERMINOLOGIA PARA A ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO DE TEXTOS NARRATIVOS DE FICÇÃO <i>Mariana da Silva Caprioli, Larissa de Mello Lima, Walter Moreira, João Batista Ernesto de Moraes</i>	133
CENÁRIO DA REPRESENTAÇÃO DA INFORMAÇÃO NA ORGANIZAÇÃO DO CONHECIMENTO: VOCABULÁRIOS CONTROLADOS DE ACERVOS ARQUIVÍSTICOS NO RIO DE JANEIRO <i>Rosale de Mattos Souza, Antonio Victor Rodrigues Botão, Julia de Vasconcellos Machado da Silva</i>	143
POSSIBILIDADES E VIABILIDADE DA DESCRIÇÃO ARQUIVÍSTICA DESDE A PRODUÇÃO DA INFORMAÇÃO DIANTE DOS DESAFIOS DE SUA RECUPERAÇÃO <i>Angélica Alves da Cunha Marques, Thaianie Honda Cotts</i>	155

FORMAÇÃO E ATUAÇÃO PROFISSIONAL EM ORGANIZAÇÃO DO CONHECIMENTO	165
ATUAÇÃO DO PROFISSIONAL DA INFORMAÇÃO NA ORGANIZAÇÃO DO CONHECIMENTO NA WEB DE DADOS	167
<i>Francisco Carlos Paletta, Armando Manuel Barreiros Malheiro da Silva</i>	
O ENSINO DA CLASSIFICAÇÃO ARQUIVÍSTICA NA FORMAÇÃO DO PROFISSIONAL DA INFORMAÇÃO: ANÁLISE QUALITATIVA DOS CURRÍCULOS DA UFF E UNESP	177
<i>Juliana de Mesquita Pazos, Clarissa Moreira dos Santos Schmidt</i>	
AÇÕES PARA O PROCESSO DE ENSINO APRENDIZAGEM EM DISCIPLINAS DE ORGANIZAÇÃO E TRATAMENTO DA INFORMAÇÃO: INOVAÇÃO PEDAGÓGICA DO GRUPO DE PESQUISA RECRI/UFMG	189
<i>Elisângela C. Aganette, Benildes Coura M. S. Maculan, Célia da Consolação Dias, Gercina Ângela Lima, Cintia de Azevedo Lourenço</i>	
ATUAÇÃO PROFISSIONAL EM AMBIENTES DIGITAIS: PERSPETIVAS PARA O DESENVOLVIMENTO DA COMPETÊNCIA EM INFORMAÇÃO	203
<i>Gabriela Belmont de Farias, Maria de Fátima Oliveira Costa</i>	
POLÍTICA DE INDEXAÇÃO PARA BIBLIOTECAS: ESTUDO DE ELABORAÇÃO, IMPLANTAÇÃO E AVALIAÇÃO COM PESQUISA-AÇÃO INTEGRAL	213
<i>Mariângela Spotti Lopes Fujita</i>	
A ESTRUTURA/CARGO EM BIBLIOTECA E ARQUIVO: O PODER ORGANIZACIONAL	225
<i>Samir Elias Kalil Lion, Zeny Duarte</i>	
GESTÃO POR COMPETÊNCIA EM BIBLIOTECAS UNIVERSITÁRIAS	237
<i>Andreza Nadja Freitas Serafim, Izabel França de Lima Giulianne Monteiro Pereira, Fernanda Mirelle A. Silva, Bernardina Maria Juvenal Freire Oliveira</i>	
RECONSTRUCTING NEWS SPREAD NETWORKS AND STUDYING ITS DYNAMICS	249
<i>Elisa Mussumeci, Flávio Codeço Coelho</i>	
A ÁREA DE DOMÍNIO NO ÂMBITO DA DISCIPLINA DE LINGUAGENS DOCUMENTÁRIAS	261
<i>Heliomar Cavati Sobrinho, Mariângela Spotti Lopes Fujita, Natanna Santana de Moraes</i>	
PRÁTICA DA CATALOGAÇÃO DE MATERIAIS ESPECIAIS COMO PARTE DA FORMAÇÃO DO PROFISSIONAL DA INFORMAÇÃO	271
<i>Luciana Moreira, Monica Marques Carvalho Gallotti</i>	
COMPETÊNCIAS E COMPORTAMENTO INFORMACIONAL NA FORMAÇÃO CONTINUADA DE BIBLIOTECÁRIOS DO SISTEMA UNIVERSITÁRIO DE BIBLIOTECAS DA UNIVERSIDADE FEDERAL DA BAHIA	283
<i>José Carlos Sales dos Santos, Sinara Abreu Alves</i>	
GESTÃO DO CONHECIMENTO: FERRAMENTAS PARA RETENÇÃO E DISSEMINAÇÃO DO CONHECIMENTO ORGANIZACIONAL NO TRIBUNAL REGIONAL FEDERAL DA PRIMEIRA REGIÃO	293
<i>Marcia Mazo Santos, Marília de Souza de Mello, Márcio da Silva Albuquerque</i>	
CONCEITOS EM ORGANIZAÇÃO DO CONHECIMENTO A PARTIR DO TEMA ÉTICA: DISCIPLINA DE USO DO SISTEMA CDU	307
<i>Benildes Coura Moreira dos Santos Maculan</i>	
GESTÃO DA INFORMAÇÃO OU DO CONHECIMENTO? UM ESTUDO DE CASO NA DIREÇÃO-GERAL DOS ESTABELECIMENTOS ESCOLARES-DIREÇÃO DE SERVIÇOS DO CENTRO	317
<i>Ana Isabel Coelho, Maria Cristina Vieira de Freitas</i>	
PERSPETIVAS DE INVESTIGAÇÃO EM REPRESENTAÇÃO E ORGANIZAÇÃO DO CONHECIMENTO: ATUALIDADE E TENDÊNCIAS	327
UMA ABORDAGEM FENOMENOLÓGICA ÀS ORGANIZAÇÕES INTELIGENTES: A PERSPETIVA DOS ESTUDANTES DE PÓS-GRADUAÇÃO	329
<i>Luís Corujo, Jorge Revez</i>	

O PAPEL DAS DEFINIÇÕES NA FORMAÇÃO DE CORPUS TERMINOLÓGICO PARA ELABORAÇÃO DE ONTOLOGIAS DE DOMÍNIOS: UMA EXPERIÊNCIA A PARTIR DE EXTRAÇÃO AUTOMÁTICA E MANUAL <i>Maria Luiza de Almeida Campos, Renato Rocha Souza</i>	341
REQUISITOS PARA FERRAMENTAS DE VISUALIZAÇÃO DE ONTOLOGIAS <i>Cristiane Mendes Netto, Gercina Angela Borém de Oliveira Lima</i>	351
COERÊNCIA NA REPRESENTAÇÃO TEMÁTICA DE ARTIGOS CIENTÍFICOS NO REPOSITÓRIO DE SAÚDE PÚBLICA DA FUNDAÇÃO OSWALDO CRUZ <i>Fatima Cristina Lopes dos Santos, Maria Cecilia Magalhães Mollica, Vânia Lisboa da Silveira Guedes</i>	363
A PROFUNDIDADE TEMÁTICA DA ANÁLISE DE ASSUNTO DE INDEXADORES NO BRASIL E PORTUGAL: ESTUDO DA COGNIÇÃO COM MAPAS CONCEITUAIS <i>Dulce Amélia de Brito Neves, Mariângela Spotti Lopes Fujita</i>	375
TEXTO VERSUS IMAGENS? FOLKSONOMIAS E INDEXAÇÃO SOCIAL EM ARQUIVOS <i>Ana Margarida Dias da Silva, Leonor Calvão Borges</i>	391
PROPOSTA DE UM MÉTODO PARA O DESENVOLVIMENTO DE OBJETOS DE APRENDIZAGEM BASEADO NA SEMIÓTICA E NO DESIGN DA INFORMAÇÃO <i>Marcia Cristina dos Reis, Edberto Ferneda</i>	401
GESTÃO TERMINOLÓGICA: SUBSÍDIOS PARA A CONSTRUÇÃO DO VOCABULÁRIO CONTROLADO DA CARNAÚBA <i>Jacqueline A. Souza, Monica Marques Carvalho Gallotti, Jailma Santos</i>	413
OS PROCESSOS DE NEGÓCIO, A GESTÃO DE DOCUMENTOS E OS FLUXOS DOCUMENTAIS: ALGUMAS PERSPECTIVAS E RELAÇÕES <i>Livia Marangon Duffles Teixeira, Elisângela Cristina Aganette</i>	421
RELAÇÕES SEMÂNTICAS EM FOLKSONOMIA NO DOMÍNIO DA DIABETES <i>Jacqueline A. Souza</i>	431
ANÁLISE BIBLIOMÉTRICA DA PRODUÇÃO CIENTÍFICA DO GRUPO DE PESQUISA MHTX <i>Gercina Ângela de Lima, Benildes Coura Moreira dos Santos Maculan, Filipi Miranda Soares</i>	441
CIÊNCIA DA INFORMAÇÃO E WEB SEMÂNTICA: UMA RELAÇÃO EFETIVA OU APÓCRIFA? <i>Luís Machado, Maria da Graça Simões, Renato Rocha Souza, Maurício Almeida</i>	453
EL ÁRBOL DE DOMINIO COMO HERRAMIENTA EN ORGANIZACIÓN DEL CONOCIMIENTO: APLICACIÓN AL ÁREA DE EDUCACIÓN SUPERIOR <i>Mario Barité, Mirtha Rauch</i>	467
ANÁLISE DO CONCEITO DE "IDENTIFICAÇÃO ARQUIVÍSTICA" À LUZ DOS PRINCÍPIOS DA TEORIA DO CONCEITO DE DAHLBERG <i>Danilo André Bueno, Ana Célia Rodrigues, Maria Luiza de Almeida Campos</i>	477
REUSO DE ONTOLOGIAS: UMA REVISÃO DE LITERATURA NO CONTEXTO DA WEB SEMÂNTICA <i>Helder Firmino, Gercina Lima</i>	487
RELACIONES CIENTÍFICAS Y ORGANIZACIÓN DEL CONOCIMIENTO <i>Carlos Olmeda-Gómez, Daniel Martínez-Ávila, María Antonia Ovalle-Perandones</i>	499
LA INDIZACIÓN DE ARTÍCULOS CIENTÍFICOS CON EL SISTEMA DE INDIZACIÓN AUTOMÁTICA SISA COMPARADA CON LA INDIZACIÓN EN LAS BASES DE DATOS AGRICOLA, WoS Y SCOPUS <i>Isidoro Gil-Leiva</i>	509
ORGANIZAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO EM CONTEXTO EMPRESARIAL: A APLICAÇÃO DA ANÁLISE POR FACETAS <i>Olívia Pestana</i>	525
PERSPETIVAS DE HARMONIZAÇÃO DOS SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO NAS ORGANIZAÇÕES <i>Ana Rita Ferreira, Maria da Graça Simões</i>	533

ORGANIZACIÓN AUTOMÁTICA DE DOCUMENTOS MEDIANTE TÉCNICAS DE ANÁLISIS DE REDES SOCIALES <i>Carlos G. Figuerola</i>	545
VISIBILIDADE DE PESQUISADORES NA TEMÁTICA POLÍTICA DE INDEXAÇÃO: ANÁLISE DE CITAÇÃO E COCITAÇÃO <i>Roberta Cristina Dal' Evedove Tartarotti, Paula Regina Dal' Evedove, Mariângela Spotti Lopes Fujita</i>	557
INFORMAÇÃO E COMUNICAÇÃO DOS CONTROLES INTERNOS DAS UNIVERSIDADES FEDERAIS DO NORDESTE DO BRASIL <i>Wagner Miranda Gomes, Zeny Duarte, Dulcinéia Assunção, Adriana Rosa</i>	567
A FUNÇÃO DO REGISTRO DE AUTORIDADE DE ASSUNTO NA CONSTRUÇÃO E USO DE LINGUAGENS DE INDEXAÇÃO PARA CATÁLOGOS ONLINE <i>Mariângela Spotti Lopes Fujita, Luciana Beatriz Piovezan, Natany de Souza dos Santos</i>	577
RECORDS IN CONTEXTS Y LA PUBLICACIÓN DE CONJUNTOS DE DATOS ARCHIVÍSTICOS INTEROPERABLES <i>Juan-Antonio Pastor-Sánchez, Dunia Llanes Padrón</i>	587
UTILIZACIÓN DE CATEGORÍAS DE WIKIPEDIA EN PROCESOS DE ORGANIZACIÓN DE INFORMACIÓN: HACIA UNA REVISIÓN CUALITATIVA. <i>Jesús Tramullas, Ana I. Sánchez-Casabón, Piedad Garrido-Picazo</i>	601
CONSTRUÇÃO DE TESAURO NO DOMÍNIO DOS METADADOS PARA DESCRIÇÃO DE RECURSOS BIBLIOGRÁFICOS <i>Jefferson Leite Oliveira Ferreira, Heliomar Cavati Sobrinho, Virgínia Bentes Pinto</i>	611
O USO DE VOCABULÁRIO DE NOMES POPULARES DE DOENÇAS COMO FERRAMENTA DE REPRESENTAÇÃO DA INFORMAÇÃO EM SAÚDE <i>Virgínia Bentes Pinto, Lidia Eugenia Cavalcante, Hamilton Rodrigues Tabosa, Odete Máyra Mesquita Sales</i>	617
FRENTE E ELITE DE PESQUISA DA ISKO-IBÉRICO, A PARTIR DOS TRABALHOS APRESENTADOS NO PERÍODO DE 2005 A 2015 <i>Ely F. Tannuri de Oliveira, Leilah Santiago Bufrem, Bruno Henrique Alves, Natanael V. Sobral</i>	629
ANÁLISE SOBRE AS ABORDAGENS E APLICAÇÕES DO SKOS NO ÂMBITO DA LITERATURA PERIÓDICA INTERNACIONAL DA CIÊNCIA DA INFORMAÇÃO <i>José Carlos Francisco dos Santos, Walter Moreira</i>	641
LOS MODELOS CONCEPTUALES EN LA CIENCIAS DE LA INFORMACIÓN: UN NUEVO RETO PARA LA DESCRIPCIÓN Y EL ACCESO INTEGRADO DE RECURSOS <i>Dunia Llanes-Padron, Manuela Moro-Cabero</i>	651
ORGANIZAÇÃO E REPRESENTAÇÃO DA INFORMAÇÃO CORPORATIVA <i>Maurício B. Almeida, Christiano P. Pessanha</i>	663
DIMENSÕES DAS PRÁTICAS DE AUTORIA EM ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO NO BRASIL <i>Murilo Artur Araújo da Silveira, Sônia Elisa Caregnato</i>	673
ELEMENTOS DE INDEXAÇÃO PARA ARQUIVOS PERMANENTES: APROXIMAÇÕES TEÓRICAS PRELIMINARES <i>Natália Bolfarini Tognoli, Suellen Oliveira Milani, José Augusto Chaves Guimarães</i>	683
MAPEAMENTO SISTEMÁTICO DA PRODUÇÃO CIENTÍFICA NO ÂMBITO DA BASE DE DADOS EM CIÊNCIA DA INFORMAÇÃO: PERIÓDICOS CIENTÍFICOS EM DISCUSSÃO <i>Jorge Santa Anna, Maria Elizabeth de Oliveira Costa</i>	693
O CONTROLO DE AUTORIDADE DE DOCUMENTOS FOTOGRÁFICOS NOS ARQUIVOS <i>Joana Rodrigues</i>	705
PROPOSTA METODOLÓGICA PARA MAPEAMENTO E ANÁLISE DE CONFIGURAÇÕES DE DISCURSOS SOCIAIS: AS REPRESENTAÇÕES DISCURSIVAS DE DILMA ROUSSEFF <i>Ludmila Salomão Venâncio, Renato Rocha Souza</i>	715

A BIBLIOMETRIC STUDY OF ALGERIAN JOURNALS IN AGRONOMY. METHODOLOGICAL AND POLITICAL CONSULTING ASPECTS <i>Radia Bernaoui, Peter Ohly</i>	727
CARTAS COMO GÊNERO DO DISCURSO CIENTÍFICO NA ÁREA DE MEDICINA TROPICAL: ANÁLISE BIBLIOMÉTRICA E MAPEAMENTO DE REDE TERMINOLÓGICA <i>Maria José Veloso da Costa Santos, Vania Lisboa da Silveira Guedes</i>	737
ESTRUTURAS SEMÂNTICAS DA REPRESENTAÇÃO DO CONHECIMENTO ARQUIVÍSTICO: ANALISANDO ELEMENTOS ESTRUTURANTES E DE CONTEÚDO <i>Wanessa Rodrigues Martins, Thiago Henrique Bragato Barros</i>	747
AS PESQUISAS SOBRE PERIÓDICO CIENTÍFICO PUBLICADAS NO BRASIL: UMA REVISÃO SISTEMÁTICA DE LITERATURA <i>Jorge Santa Anna, Maria Elizabeth de Oliveira Costa</i>	761
ONTOLOGÍA EN GESTIÓN DE PROYECTOS. UN ENFOQUE PRÁCTICO <i>Anisleiby Fernández Hernández, María José López Huertas</i>	775
REVISÃO SISTEMÁTICA BASEADA EM PESQUISA BIBLIOGRÁFICA ESTRUTURADA – PPBE: UM MAPEAMENTO SOBRE ANÁLISE FACETADA APLICADA À ARQUITETURA DA INFORMAÇÃO <i>Graciane Silva Bruzanga Borges, Gercina Ângela de Lima</i>	791
O ATO LEITOR NO PROCESSO DE INDEXAÇÃO: UMA ABORDAGEM COGNITIVA <i>Ana Lúcia Terra</i>	803
TRATAMENTO INFORMACIONAL DA IMAGEM BIOMÉDICA: ESTADO DA ARTE EM PERIÓDICOS DE CIÊNCIA DA INFORMAÇÃO <i>Joice Cleide Cardoso Ennes de Souza</i>	811
O RESUMO CIENTÍFICO COMO RECURSO DE ACESSO EQUITATIVO À INFORMAÇÃO <i>Helena Leitão, Maria da Graça Simões</i>	825
PRINCÍPIOS PARA MODELAGEM DE DOMÍNIOS EM SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO (SOC) <i>Simone Torres, Maurício Barcellos Almeida, Maria da Graça Melo Simões</i>	841
COMPUTAÇÃO EM NUVEM E SISTEMAS DE GESTÃO DOCUMENTAL: AVALIAÇÃO DE RISCOS E RECOMENDAÇÕES <i>Ariovaldo Veiga de Almeida, Maria Cristina Vieira de Freitas</i>	853
A CATEGORIZAÇÃO COMO ORGANIZAÇÃO DO CONHECIMENTO NA PERSPECTIVA DE CHAIM ZINS <i>Guilherme Ataíde Dias, Josiana Florêncio Vieira Régis de Almeida</i>	869
INTERACCIÓN TECNOLÓGICA Y FUNDAMENTOS EPISTEMOLÓGICOS EN ORGANIZACIÓN DEL CONOCIMIENTO <i>Rosa San Segundo, María Adelina Codina-Canet</i>	879
MAPAS CONCEPTUALES PARA MODELAJE CONCEPTUAL DE ONTOLOGIAS DE DOMINIO <i>Gracy Kelli Martins, Denysson Axel Ribeiro Mota</i>	889
ONTOLOGÍA DE FENÓMENOS NATURALES: PLANEACIÓN Y DISEÑO <i>Adriana Suárez Sánchez, Ariel Alejandro Rodríguez García</i>	899
TAXONOMIA COMO INSTRUMENTO AUXILIAR PARA ESCOLHA DE PROCEDIMENTOS DE AUDITORIA CONTABIL <i>José Humberto da Cruz Cunha</i>	913
GESTÃO DO CONHECIMENTO E SISTEMAS DE GESTÃO DO CONHECIMENTO NAS MICROEMPRESAS: FATORES CRÍTICOS, REQUISITOS E RECOMENDAÇÕES <i>Paulo Madeira Gonçalves, Maria Cristina Vieira de Freitas</i>	923
ESTUDO DE PROSPECÇÃO E VIGILÂNCIA TECNOLÓGICA PARA BUSCA E RECUPERAÇÃO DA INFORMAÇÃO <i>Maria do Carmo Oliveira Ribeiro, Fernanda Maria Melo Alves</i>	935

REVISITANDO WILLIAM SHAKESPEARE EM SISTEMA DE ORGANIZAÇÃO DO CONHECIMENTO NA ÁREA DE LITERATURA INGLESA	947
<i>Márcia Barcelos Gomes, Vânia Lisboa da Silveira Guedes, Maria José Veloso da Costa Santos</i>	
CATEGORIZAÇÕES DE SOC: UMA INVESTIGAÇÃO LITERÁRIA	957
<i>Márcio Bezerra Da Silva, Zeny Duarte de Miranda</i>	
ANÁLISE DA PAISAGEM RURAL SOB A ÓTICA DA TEORIA DA ESPACIALIDADE: SUA APLICAÇÃO AOS CENSOS AGROPECUÁRIOS DE 1970 E 2006 E O CAMPO TRAÇADO NO ATLAS DO ESPAÇO RURAL BRASILEIRO (2011)	967
<i>Vera Lucia Punzi Barcelos Capone, Rosa Inês de Novais Cordeiro</i>	
DESAFIOS À ORGANIZAÇÃO E ACESSO AO PATRIMÔNIO CULTURAL	977
TRATAMENTO E ANÁLISE DOS OBJETOS PRESENTADOS AOS REITORES DE UNIVERSIDADES: QUESTÕES SIMBÓLICAS E DIPLOMÁTICAS	979
<i>Samanta Prado, Luzia S. F. Costa, Luciana S. Gracioso</i>	
ANÁLISE DO TERMO COMPOSTO NA INDEXAÇÃO DE UMA COLEÇÃO DE POSTAIS ILUSTRADOS (1940-1960) À LUZ DA NF Z 47-200 (1985)	989
<i>Maria da Graça de Melo Simões, María Begoña López-Ávila, Blanca Rodríguez-Bravo, Maria de Fátima Moura Carvalho, Giovana Deliberali Maimone</i>	
O LEGADO DO MÉTODO DIPLOMÁTICO E A IDENTIFICAÇÃO DOCUMENTAL: CONTRIBUIÇÕES PARA A ORGANIZAÇÃO DO CONHECIMENTO	1001
<i>Gabrieli Aparecida da Fonseca, Sonia Troitiño</i>	
ARQUIVOS PESSOAIS E SEUS TIPOS DOCUMENTAIS: A PERSPECTIVA DA REPRESENTAÇÃO	1011
<i>Lucia Maria Velloso de Oliveira, Camilla Campoi de Sobral</i>	
REFLEXÕES SOBRE UMA ORGANIZAÇÃO DO CONHECIMENTO FOTOGRÁFICO SEGUNDO SEU CONTEXTO DE PRODUÇÃO DOCUMENTAL	1019
<i>Bruno Henrique Machado, Rafael Semidão, Telma Campanha de Carvalho Madio</i>	
DOCUMENTOS AUDIOVISUAIS, INFORMAÇÃO E MEMÓRIA: IDENTIFICAÇÃO DE ACERVOS FOTOGRÁFICOS E FÍLMICOS DAS REGIÕES ADMINISTRATIVAS DO DISTRITO FEDERAL DO BRASIL	1029
<i>Eliane Braga de Oliveira, Miriam Paula Manini, Sérgio Peçanha da Silva Coletto</i>	
MEMÓRIA E PATRIMÔNIO CULTURAL: O MUSEU DO CEARÁ COMO LUGAR DE MEMÓRIA	1039
<i>Fátima Portella Cysne, Maria Cleide Rodrigues Bernardino, Carla Façanha de Brito³, Adriana Nóbrega Silva, Denise Marques Rodrigues</i>	
COLECCIONES DE DATOS ABIERTOS ENLAZADOS: DE LA BÚSQUEDA AL DESCUBRIMIENTO DE INFORMACIÓN	1049
<i>María Luisa Alvite Díez</i>	
LA VISIBILIDAD DEL TRADUCTOR EN CATÁLOGOS Y REPERTORIOS BIBLIOGRÁFICOS ESPAÑOLES (1970-2010): POSIBILIDADES Y ESTRATEGIAS DE ANÁLISIS	1061
<i>Belén Santana López, Crispulo Travieso Rodríguez, Manuel de la Cruz Recio</i>	
OS VOCABULÁRIOS CONTROLADOS NA ORGANIZAÇÃO E GESTÃO DO PATRIMÔNIO CULTURAL: ORIENTAÇÕES PRÁTICAS	1071
<i>Natália Jorge, Filipa Medeiros, Juliana Rodrigues Alves, Susana Medina</i>	
A MEDIAÇÃO DA INFORMAÇÃO NA REDE DE ARQUIVOS DIPLOMÁTICOS IBERO-AMERICANOS – RADI	1087
<i>Ana Cláudia Medeiros, Bruna Lessa, Carolina Magalhães, Daniel Marins, Hildenise Novo, Ivana Lins, Luis Carlos Batista, Maira Salles</i>	
METODOLOGIAS DE SALVAGUARDA DOS BENS CULTURAIS: CONTRIBUIÇÕES DA CIÊNCIA DA INFORMAÇÃO NO CAMPO PATRIMONIAL	1097
<i>Vitória Gomes Almeida, Izabel França de Lima, Henry Poncio Cruz de Oliveira</i>	

DIAGNÓSTICO DA SOBREVIVÊNCIA E DISPONIBILIDADE DOS PERIÓDICOS CIENTÍFICOS ON-LINE DE ACESSO ABERTO: ESTUDO DO ESTADO DE MINAS GERAIS (BRASIL) <i>Edna da Silva Angelo, Marlene Oliveira</i>	1107
USO DE UM PERFIL DE APLICAÇÃO PARA A GESTÃO E AGREGAÇÃO DA INFORMAÇÃO AUDIOVISUAL DA UNIVERSIDADE DO PORTO: O CASO TVU <i>Sara Oliveira</i>	1121
PATRIMÔNIO: IDENTIDADES CULTURAIS NO SANTUÁRIO BOM JESUS DO MATOSINHOS EM CONGONHAS-MG <i>Cátia Rodrigues Barbosa, Luana Rodrigues Godinho Silveira</i>	1131
AVALIAÇÃO DOS MODELOS CLÁSSICOS DE RECUPERAÇÃO DA INFORMAÇÃO EMPREGADOS NA ÁREA DA SAÚDE <i>Josiane Mello, Angel Freddy Godoy Viera</i>	1141
REPRESENTAÇÃO DOCUMENTAL DE GRAFFITI COMO PATRIMÔNIO CULTURAL: PROPOSTA METODOLÓGICO- CONCEITUAL <i>Zaira Regina Zafalon, Fábio Rogério Batista Lima, Paula Regina Ventura Amorim Gonçalez, Paula Regina Dal'Evedove</i>	1151
U. PORTO DIGITAL MUSEUM PROJECT: TOWARDS CONVERGENCE IN UNIVERSITY'S INFORMATION MANAGEMENT <i>Maria Manuela Pinto, Susana Medina, Paula Menino Homem, Augusto Ribeiro, Marisa Monteiro, Rodolfo Matos, Alexandre Lourenço, Sara Oliveira</i>	1161
POSTERS	1173
PRÁTICA INFORMACIONAL EM DOMÍNIOS INTERDISCIPLINARES: UM ESTUDO SOBRE A PRODUÇÃO DO CONHECIMENTO NA GOVERNANÇA DA ÁGUA <i>Marianna Zattar, Regina Maria Marteleto, Marta Pedro Varanda</i>	1175
CLASSIFICAÇÃO DA INFORMAÇÃO ARQUIVÍSTICA SEGUNDO MÉTODOS FILOMEMÉTICOS: METADADOS COMO UNIDADES FENOTÍPICAS? <i>L. S. Ascensão de Macedo</i>	1181
O PAPEL DOS SITES NOTICIOSOS NA DIVULGAÇÃO DA INFORMAÇÃO CIENTÍFICA E SEU USO NOS PROCESSOS DE EDUCAÇÃO INFORMAL DE CTI <i>Benjamín Camacho Vega, Zeny Duarte de Miranda</i>	1197
A PRESERVAÇÃO DA MEMÓRIA DOS GEOPARQUES NO BRASIL: O CENTRO DE MEMÓRIA DO QUADRILÁTERO FERRÍFERO EM MINAS GERAIS <i>Cátia Rodrigues Barbosa, Mônica Elisque Carmo, Leonardo Barreto Oliveira</i>	1205
O PROCESSO DE REFERÊNCIA NAS BIBLIOTECAS UNIVERSITÁRIAS BRASILEIRAS COM USO DAS REDES SOCIAIS <i>Ana Roberta Sousa Mota</i>	1213
A ORGANIZAÇÃO DE DOCUMENTOS TÉCNICOS DE AERONAVES À LUZ DA CLASSIFICAÇÃO FACETADA <i>Ana Patrícia Guimarães, Renato Rocha Souza</i>	1219

INTRODUÇÃO

No mundo contemporâneo, marcado pela dualidade do global vs. local e pelos impactos de inovações tecnológicas que afetam os ambientes informacionais e comunicacionais, emergem novos ecossistemas nos quais a apropriação e o uso da informação ocupam um lugar de destaque. É neste contexto plural, porque heterogéneo e transversal, que assume particular relevo a organização do conhecimento.

Revisitar teorias e práticas torna-se obrigatório, de modo a assegurar a sua compatibilidade e robustez face às configurações morfológicas dos sistemas de organização do conhecimento. Do mesmo modo, apontar para caminhos a serem explorados num ambiente propício à fertilização cruzada de distintas áreas do conhecimento é fundamental num ambiente de cunho marcadamente digital.

O espaço ibérico, que partilha características históricas e culturais, assume-se como uma área privilegiada para o desenvolvimento de projetos de investigação capazes de responder a desafios sociais, culturais e tecnológicos. Nos encontros ibéricos da ISKO, a presença da comunidade científica latino-americana tem contribuído para a produção científica nesta área de conhecimento.

O *III Congresso ISKO Espanha e Portugal/XIII Congresso ISKO Espanha*, subordinado ao tema “Tendências atuais e perspetivas futuras em organização do conhecimento”, definiu os seguintes painéis:

- Perspetivas epistemológicas em organização do conhecimento;
- Formação e atuação profissional em organização do conhecimento;
- Perspetivas de investigação em representação e organização do conhecimento: atualidade e tendências e
- Desafios à organização e acesso ao património cultural.

As 109 propostas aprovadas foram sujeitas ao processo de revisão por pares na modalidade ‘double-blind’ que envolveu a colaboração de 39 revisores cuja prestimosa colaboração foi essencial para a seleção dos trabalhos apresentados, garantindo, deste modo, o elevado nível científico deste congresso.

Não podemos terminar sem endereçar um agradecimento muito especial a todos os apoios diretos recebidos - Faculdade de Letras, Departamento de Filosofia, Comunicação e Informação, Centro de Estudos Interdisciplinares do Século XX (CEIS 20) e FCT -, bem como aos patrocinadores, que tornaram possível a realização deste encontro. Agradecemos, ainda, aos moderadores das sessões cujo rigor facilitou a execução de um programa intenso ao longo dos dois dias do evento. E, finalmente, um agradecimento muito particular à equipa que nos ajudou a preparar o *III Congresso ISKO Espanha e Portugal/XIII Congresso ISKO Espanha*, sem a qual nada disto teria sido possível: António Tavares Lopes, Ana Miguéis, Bruno Neves, Cristina Freitas, Mário Bernardes e ainda aos alunos de licenciatura e mestrado que se mobilizaram para secretariar este encontro. Bem hajam pelo vosso generoso e empenhado contributo!

Coimbra, 24 de novembro de 2017.

Maria da Graça Simões e Maria Manuel Borges

FOUNDATIONS AND MOVING FORWARD: THE HORIZONS AND THE FUTURE OF RESEARCH IN KNOWLEDGE ORGANIZATION

Joseph T. Tennis

University of Washington

INTRODUCTION

In the context of the history of knowledge organization, we have attempted to intellectualize the methods and assumptions behind the meaningful, but everyday, work of classification. This has, over the years, been expanded to other traditional forms of representation and organization, like cataloguing. Further, advances in technology and the technological context have precipitated expansions in this field. These are not limited to the contemporary linked open vocabulary environment, but also include questions around data intensive discovery, reproducibility, critical algorithm studies, and the quantitative self. Today, the well-intellectualized concerns of traditional knowledge organization are reborn in this new context at the horizon of new technologies of inquiry, knowledge production, and fragmentations of storage.

What then constitutes the foundations of contemporary knowledge organization? What changes in the context of new technologies, contemporary knowledge production, and the increased fragmentation of information storage? In this paper we explore the way in which constructs and methods, developed in the past 140 years of knowledge organization work, form the foundation of our contemporary thinking in knowledge organization. We then move from that to discuss what we must do, in our discipline, to move forward with research fronts, innovation, and inquiry.

FOUNDATIONS

Elsewhere, I have commented on layers of classification theory (Tennis, 2015; Tennis 2016). In these conversations, I lay out what I see as foundational issues in the design, implementation, and evaluation of classification schemes. In these texts I identify three layers: foundational, first, and second-order classificatory processes. Foundational classification theory deals with philosophical and definitional issues of classification.

“In foundational classification theory, the nature of the process and the products of classification are called into question. Examples of this kind of theory production are primarily conceptual papers that argue for a perspective. We might point to early examples of this work by Richardson (1901/1964) and Broadfield (1946), who were intellectualizing a practical work mode in libraries. This work is alive and well today with many arguments for particular philosophical stances toward classification (e.g., Hjørland and Pederson 2005; Lee 2011), or arguments for particular ontological understanding of concepts in the field (Furner 2009).”

Likewise, the foundations for knowledge organization lie in the philosophical work that precedes the description and representation of documents. In a Dahlberg conception: we must understand what we mean by concepts (or as she often said, knowledge units), and the way we order them (cf., Dahlberg 2014). If we follow Smiraglia (2014), we will also need to comprehend what is meant by document, work, and author or even more broadly the entities we will describe in order to access knowledge.

Examples of foundational work:

Definitional work, especially if it is contested definitional work.

E.g., what is a knowledge unit? How do we know it when we see it? (cf., Hjørland, 2009).

Taxonomic work, especially if it provides theoretical advancement.

E.g., What is the taxonomy of KOS (cf., Sousa, R. R. et. al., 2010).

Mission work, especially if it outlines the rationale for the field.

E.g., What do we mean by identity in knowledge organization? (cf., Furner, 2009).

Exploration based on applied or first order work, like Ranganathan's *Abstract Classification* (1967).

FIRST ORDER KNOWLEDGE ORGANIZATION WORK

The first order of knowledge organization research consists of discussions on how to build knowledge organization systems. This also contains work that I have called warrant studies, that is citation analysis, domain analysis. These are instrumental to building out systems, but in the majority of cases, these are examples of applied knowledge organization.

As I said in the context of classification theory, “Exemplar literature in this area is the work of S. R. Ranganathan (1937; 1957; 1967), the CRG, and contemporary summaries of design patterns and discussions of semantic web work (cf., Frické 2012; Hlava 2014),” (Tennis, 2015 p. 245). I currently have a student engaged in domain analysis in the context of neurodiversity and autism (Zoloymi and Tennis, 2017).

In the context of our work we are not necessarily advancing the research agenda of knowledge organization to better understand how to design and critique knowledge organization systems. We are identifying terms and how they relate to identity, caregiving, and problem solving. It would be possible to feed this work into foundational work in KO, but, in my opinion, on its own, this domain analysis does not advance KO.

SECOND ORDER KNOWLEDGE ORGANIZATION WORK

As I said above, research in second order knowledge organization asks what we do with knowledge organization systems once they are built. Here we want to know how systems change over time, how they interoperate with other systems, how modeling affects KOS, and where foundational issues conflict in the interpretation of change, interoperation and modeling. Interoperability is a well established literature in knowledge organization, at least with regard to thesauri. We have an evolving literature that is looking at KOS interoperability more broadly (Zang and Chan, 2004).

The modeling effect is one that is also emerging as a well-documented space because of semantic web requirements. For example, Panzer and Green sought clarity about the nature of hierarchical relationships when they moved to encode the Dewey Decimal System in the Web Ontology Language (OWL) (Green and Panzer, 2010).

My own work on how indexing languages change over time is a second order activity in my view.

HORIZONS OF KNOWLEDGE ORGANIZATION WORK

And it seems to me that the horizons of knowledge organization research can be pulled from the work that can be labeled second order. Even with well established fields like interoperability, we still do not have all the answers. And what is even more fruitful for those of us, like myself, that are more theoretically inclined, is to contemplate the relationship between the definitional, taxonomic, mission, and exploration work that will influence, philosophically, the work going on in the second order. That is, how does our apprehension of concepts, language, and the entities we care about in knowledge organization change if our philosophy surrounding them changes? This is an exciting horizon for knowledge organization, especially with our long history.

A final note on some of the externalities that affect knowledge organization. We do have foundational practices and our own theoretical conversations that continue within the field, but the horizons of knowledge organization must also reckon with things outside the field.

In my thinking knowledge organization must deal with metadata for data driven decisions, quantified self and classification; and the disaggregation, reconfiguration, and reinvention of description. The order of the disciplines, the order of knowledge is not stagnant, nor does it admit to a single description. Data, as we are learning more and more, is out us, not about knowledge, and yet once reified, it admits to description and ordering, not unlike documents. What will we make of this development?

CLOSING

Great thinkers have come before us, and provided us with a rich literature. We are standing on the shoulders of giants. We must take that responsibility, for their work, their contributions, to our hearts. We must do good work that extends their mission to better understand how best to design, study, and critique the processes of organizing and representing documents that societies see as worthy of preserving.

REFERENCES

Araújo, P. C., Tennis, J. T., and Guimarães, J. A. C. (2017a). "The Concept of Epistemology in Knowledge Organization: O Conceito de Epistemologia na Organização do Conhecimento." In *Memória, cultura e tecnologia na Organização do Conhecimento*, , 4:71–78. Estudos Avançados Em Organização Do Conhecimento. Recife: Ed. UFPE. <http://isko-brasil.org.br/wp-content/uploads/2013/02/livro-ISKO-2017.pdf>

Araújo, P. C., Guimarães, J. A. C., and Tennis, J. T. (2017b). "Metatheory and Knowledge Organization." In Proceedings of the North American Symposium on Knowledge Organization. (Urbana-Champaign, IL). Available: <http://dx.doi.org/10.7152/nasko.v6i1.15154>

Dahlberg, I. (2014). Brief Communication: What is Knowledge Organization? In Knowledge Organization. 41(1): 85-91.

Furner, J. (2009). "Interrogating "Identity": A Philosophical Approach to an Enduring Issue in Knowledge Organization." In Knowledge Organization 36(1): 3-16.

Garfield, E. (1979). Citation Indexing: Its Theory and Application in Science, Technology, and Humanities. New York: Wiley.

Green R. and Panzer, M. (2010). "The ontological character of classes in the Dewey Decimal Classification." In Proceedings of the 11th International Conference for Knowledge Organization. (Rome, Italy). Advances in Knowledge Organization vol. 12. Ergon: Würzburg: 171-179.

Hjørland, B. (2009). "Concept theory." In JASIST 60(8): 1519-1536.

Hjørland, B. (2016). Knowledge organization. Knowledge Organization 43, no. 6: 475-84

Ranganathan, S. R. (1967). *Prolegomena to Library Classification*. Bombay: Asia Publishing House.

Smiraglia, R. P. (2014). *The Elements of Knowledge Organization*. Springer.

Sousa, R. R. et al. (2010). "The KOS spectra: a tentative typology of Knowledge Organization Systems." In Proceedings of the 11th International Conference for Knowledge Organization. (Rome, Italy). Advances in Knowledge Organization vol 12. Ergon: Würzburg.

Tennis, J. T. (2005). *Conceptions of Subject Analysis: A Metatheoretical Investigation*. (2005). (University of Washington: Seattle). Advisor: Allyson Carlyle Committee: Clare Beghtol, Jens-Erik Mai, and Stuart Sutton

Tennis, J. T. (2008). "Epistemology, Theory, and Methodology in Knowledge Organization: Toward a Classification, Metatheory, and Research Framework." (2008). In Knowledge Organization. 35(2/3): 102-112.

Tennis, J. T. (2012). "Forms, Knowledge and Forms of Knowing: Correct and Useful Epistemology and Knowledge Organization." (2012). In Guimarães, J. A. C. and Dodebei, V. (eds.) *Desafios e perspectivas científicas para a organização do conhecimento na atualidade. Estudos avançados em organização do conhecimento; 1*. (Marília: ISKO-Brasil): 22-29.

Tennis, J. T. (2015). "Foundational, First-Order, and Second-Order Classification Theory." In *Knowledge Organization* 42(4): 244-249.

Tennis, J. T. (2016). "Structure of Classification Theory: On Foundational and the Higher Layers of Classification Theory." (2016). In Guimarães, J. A. C., Milani, S. O., and Dodebei, V. eds. *Knowledge Organization for a Sustainable World: Challenges and Perspectives for Cultural, Scientific, and Technological Sharing in a Connected Society*. Proceedings of the Fourteenth International ISKO Conference 27-29 September 2016, Rio de Janeiro, Brazil. Advances in Knowledge Organization, Vol 15: 84-87.

Tennis, J. T. (2017). "Through Formal Analysis and Theories of Meaning: Extending the Trajectory of S. R. Ranganathan's Theoretical Frameworks in Classification Research." 2017 Sarada Ranganathan Endowment Lectures / Ranganathan Memorial Lectures. (Chennai, India), October 2017.

Tognoli, N. B., Guimarães, J. A. C., and Tennis, J. T. (2013). "Diplomatics as a Methodological Perspective for Archival Knowledge Organization." In Proceedings of the North American Symposium on Knowledge Organization. (Milwaukee, WI). Available: http://iskocus.org/NASKO2013proceedings/TOGNOLI_GUIMARAES_TENNIS_DiplomaticsAsAMethodologicalPerspective.pdf

Zeng, Marcia Lei and Lois Mai Chan. 2004. Trends and issues in establishing interoperability among knowledge organization systems. *Journal of American Society for Information Science and Technology (JASIS&T)* 55(5): 377 – 395

Zoloymi, A. and Tennis, J. T. (2017). "The Autism Prism: A Domain Analysis Paper Examining Neurodiversity." In Proceedings of the North American Symposium on Knowledge Organization. (Urbana-Champaign, IL). Available: <http://dx.doi.org/10.7152/nasko.v6i1.15153>

**PERSPETIVAS EPISTEMOLÓGICAS
EM ORGANIZAÇÃO DO CONHECIMENTO**



TODOROV, A FILOSOFIA SIMBÓLICA E A ORGANIZAÇÃO DO CONHECIMENTO: SOBRE O TRANSBORDAMENTO DO SIGNIFICANTE DE TESAURO AOS DESAFIOS DA WEB PRAGMÁTICA

Gustavo Silva Saldanha

IBICT – UNIRIO, saldanhaquim@gmail.com

RESUMO Da obra de Emanuele Tesauro aos desafios da organização do conhecimento na *web*, podemos identificar o que Tzvetan Todorov chama de transbordamento do significante pelo significado. Trata-se da compreensão do papel da filosofia simbólica como domínio teórico de estudo e modalidade de percepção do real. O objetivo geral deste trabalho é discutir as ideias de Todorov a partir da centralidade ocupada pelo simbólico na organização do conhecimento, junto de aportes de outros teóricos do simbolismo, como Umberto Eco e Ernst Cassirer. Trata-se de uma reflexão epistemológica, pela via da filosofia da linguagem, baseada na abordagem pragmática wittgensteiniana, sendo os processos metodológicos efetivados na interpretação do pensamento de Todorov em diálogo com teóricos do campo informacional. As conclusões apontam para a pertinência das abordagens simbólicas como fórum teórico para a reflexão contemporânea em organização do conhecimento, bem como uma tópica fundamental no plano epistemológico-histórico do estudo da representação do conhecimento, principalmente a partir de cinco dimensões em aberto, a saber, a assertiva hipotético-epistemológica capurriana da centralidade retórica na fundamentação da Ciência da Informação, o papel de uma teoria da linguagem em Ranganathan, a atualidade crítico-discursiva das ideias de Tesauro, a potencialidade do discurso como categoria para a crítica social em organização do conhecimento, e, por fim, as margens de compreensão e resolução dos desafios de uma *web* pragmática.

PALAVRAS-CHAVE *Todorov, Simbólico, Organização do Conhecimento, Retórica, Discurso.*

ABSTRACT From the work of Emanuele Tesauro to the challenges of the knowledge organization on the web, we can identify what Tzvetan Todorov calls the “overflowing of the signifier by meaning”. It is the understanding of the role of symbolic philosophy as a theoretical domain of study and a mode of perception of reality. The general objective of the research is to discuss the ideas of Todorov from the centrality occupied by the symbolic in the knowledge organization e, along with contributions from other theorists of symbolism, such as Umberto Eco and Ernst Cassirer. This is an epistemological reflection, based in the philosophy of language and the Wittgensteinian pragmatic approach. The paper discusses the interpretation of Todorov's thinking in dialogue with informational field theorists. The conclusions point to the relevance of symbolic approaches as a theoretical forum for contemporary reflection on knowledge organization, as well as a fundamental topic on the epistemological-historical plane of the study of knowledge representation, mainly from five open dimensions, namely, the hypothetical-epistemological assertion of the rhetorical centrality in the foundation of Information Science; the research highlights also the role of a theory of language in Ranganathan; the critical-discursive actuality of Thesaurus ideas; the potentiality of discourse as a category for social criticism in the organization of knowledge; and the margins of understanding and solving the challenges of a pragmatic web.

KEYWORDS *Todorov, Symbolic, Knowledge Organization, Rhetoric, Discourse.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

1 INTRODUÇÃO: DA POTÊNCIA DA LINGUAGEM À FILOSOFIA SIMBÓLICA

“A interação possibilitada pela linguagem – a conversa - é [...] superior a tudo que a precedeu, tanto em sutilidade quanto em eficácia, de tal maneira que ela se torna a atividade humana por excelência, influenciando, por sua vez, todas as outras formas de interação.” (Todorov, 2014b, p. 102)

As diferentes abordagens contemporâneas dedicadas à compreensão dos desafios pragmáticos da organização do conhecimento podem ser reunidas, em nossa visão, naquilo Tzvetan Todorov (2014a) considerada como um transbordamento do significante pelo significado. Trata-se da compreensão do papel do simbolismo, ou da simbólica, como domínio teórico de estudo e modalidade de percepção do real.

O objetivo geral deste trabalho é discutir as ideias de Tzvetan Todorov a partir da centralidade ocupada pelo simbólico na organização do conhecimento, junto de aportes de outros teóricos do simbólico, como Umberto Eco e Ernst Cassirer. Reconhecemos hipoteticamente que o desdobramento das perspectivas de estudo do simbólico tendem a encontrar grandes horizontes na resolução de problemas aplicados da organização do conhecimento a partir de distintos avanços teóricos e técnicos.

O percurso da reflexão sobre o que Umberto Eco (2001) chama de “modo simbólico”, identificado nas tentativas de Emanuele Tesauro, no século XVII, aponta para um conjunto de questões inerentes ao “humano” como “produto do simbólico”, como colocado na abordagem cassireriana (Cassirer, 1994, 2001, 2004, 2011), demarcando a linguagem com uma centralidade na edificação das ações de produção, representação e organização do conhecimento.

Como afirma Day (1992), a potência da linguagem no âmbito da socialidade se aplicará pontualmente no desenvolvimento das perspectivas teóricas e práticas da organização do conhecimento. No mesmo contexto, a reflexão sobre o simbólico aponta para uma relação objetiva com a Retórica (antes de técnica, tomada aqui como uma filosofia do discurso), colocando em ênfase posicionamentos epistemológicos, como o de Capurro (1992), e teórico-aplicados, como o de Feinberg (2010).

Abordagem metodológica parte de um plano epistemológico-histórico e de um plano estruturado na filosofia da linguagem. O enfoque está na procura por conceitos constituídos no tempo-espço que permitem relações (elos) com os modos de constituição epistêmica da Ciência da Informação a partir da linguagem, ou seja, de sua própria feitura conceitual do campo informacional. Os elos teóricos permitem, ainda, perceber como, de Emanuele Tesauro (1670), aos estudos contemporâneos de uma *web* pragmática (Gracioso, Saldanha, 2010), tais questões foram abordadas e são pauta para a reflexão em organização do conhecimento, demonstrando a potencialidade das ideias todorovianas no campo, incluindo a abertura para horizontes críticos e sócio culturais de reflexão e intervenção.

2 PROCEDIMENTOS METODOLÓGICOS: A METAESCRITA DAS LINGUAGENS PRIMITIVAS

“[...] dar uma razão para algo que se fez ou disse significa mostrar um caminho que conduz a esta ação.” (Wittgenstein 1992, p. 30)

O estudo se trata de uma reflexão epistemológica, sendo os processos metodológicos baseados na interpretação do pensamento de Todorov em diálogo com teóricos do campo informacional. A problemática nasce em dois projetos distintos, especificamente, na passagem das conclusões do estudo “Trilhar o trivium: a filosofia da Ciência da Informação na tradição filosófica da linguagem”, iniciado em 2013, e “Ciência da Informação, Filosofia da Linguagem e Filosofia da Cultura: da epistemologia histórica informacional ao pensamento linguístico-simbólico nos estudos informacionais”, de 2016.

Preocupadas com o papel sócio teórico metodológico da linguagem na constituição e na “feitura” da epistemologia e da práxis informacional, ambas as investigações seguem como estrutura metodológica central a proposta wittgensteiniana (Wittgenstein, 1979, 1992) de “apresentação panorâmica”, orientada para a identificação de conceitos e suas semelhanças de família no escopo de indícios pragmáticos (o uso conceitual e suas apropriações pelos teóricos).

Parte-se aqui da preocupação constante em identificar “linguagens primitivas” em curso (ou em jogo) nas movimentações teóricas do campo. O método, pois, aponta para uma constante recontextualização das esferas de comunicação (ou jogos de linguagem), uma metaescrita que se orienta pela compreensão de como se registra o conhecimento de um campo do conhecimento.

3 RESULTADOS E DISCUSSÕES: O CAMINHO SIMBÓLICO NA ORGANIZAÇÃO DO CONHECIMENTO

“[...] nossos retóricos, o catálogo mais rico de que a tradição dispõe sobre os fatos simbólicos, veem a semelhança como uma relação simples e não analisável. Mais uma vez, isto não ocorre na poética sânscrita, que sabe identificar até 120 variedades da comparação.” (Todorov, 2014a, p. 64)

Todorov (2014a) aborda a relação entre língua e discurso a partir dos enfoques da abstração e da concretude. A língua pode ser reconhecida como uma abstração da linguagem, enquanto o discurso é tomado como manifestação concreta da língua. A categoria do discurso adentra um universo amplo, intersubjetivo, que reúne não apenas elementos linguísticos (como peças de um quebra-cabeças), mas os contextos particulares e as circunstâncias singulares de produção, envolvendo os elementos de interlocução, tempo e lugar. Para além de frases, pode-se falar, segundo Todorov (2014a), em “frases enunciadas”, e, mais especificamente, em “enunciados”.

Todorov (2014a) “encontra” aqui, de um plano, no plano sócio epistemológico, o pensamento de Foucault (2010) e a noção de formação discursiva” como “conjunto de enunciados que se apóia em um mesmo sistema de formação”. (Foucault, 2010, p. 121-122). Trata-se de compreender o complexo de enunciados que compõem o discurso. Por outro lado, a preocupação com o discurso, no plano epistemológico-informacional, faz com que a abordagem todoroviana entre em diálogo com reflexões que vão de Tesouro à Frohmann (1990), passando por Capurro (1992) e outras abordagens críticas,

como García Gutierrez (2011) que entraram em diálogo pontuam com as perspectivas de desenvolvimento da organização do conhecimento.

Especificamente, a abordagem todoroviana “dialoga” com a “filosofia das formas simbólicas” de Cassirer (2001), preocupada com a “atuação da imaginação da língua”, que estabelece um método próprio de criação – atuação esta que permite tanto a vigência do mito quanto da lógica. É compartilhada aqui com Aristóteles a hipótese original da relação entre discurso e linguagem. Seguindo os passos cassirerianos, a abordagem todoroviana busca por estas analogias e correspondências no plano da linguagem e a potência da noção de discurso.

Segundo Todorov (2014), se o sentido próprio do discurso poderia ser chamado de “direto”, ou seja, obtido a partir de uma possível relação de identidade, existe um terreno distinto que povoa tal sentido. Trata-se aqui do sentido discursivo indireto, também chamado por Todorov como “simbolismo linguístico”, enquadrado dentro do campo de estudos da “simbólica da linguagem”. Para o teórico, é importante demonstrar que a noção “negativa” da “não direção” do simbólico não aponta para uma redução ou para uma marginalidade. Ao contrário, faz-se relevante exatamente pela sua presença intensiva.

E que o prefixo ‘indireto’ não faça pensar um fenômeno marginal, apêndice esporádico do sentido direto: a produção indireta de sentido está presente em todos os discursos, talvez dominando inteiramente alguns deles, e não os menos importantes [...]. (Todorov, 2014, p. 14)

É neste sentido, do ponto de vista dos regimes, das políticas e da economia política da organização do conhecimento, como já nos apontaram González de Gómez (1996), Rosali Souza (1996), Frohmann (2011), Hjørland (2002) e García Gutiérrez (2011), que o simbólico alertado por Todorov (2014) se aplicaria tanto à compreensão do discurso central vislumbrado pela organização do conhecimento, saber, o discurso científico, como nas demais formações discursivas. Uma categoria comum em autores resolveria a questão: contexto, categoria essa igualmente central para Jacob & Albrechtsen (1999) em sua compreensão dos desafios os sistemas de informação a partir da pragmática.

Dada a reflexão, as dimensões nucleares identificadas na perspectiva simbólica de Todorov (2014a) para a organização do conhecimento são:

- a) Simbólica e a macro-hipótese capurriana aplicada à organização do conhecimento: a partir de Todorov (2014a), podemos aproximar a hipótese de Capurro (1992) de fundamentação da Ciência da Informação como uma “subdisciplina da Retórica”, orientada que está, estruturalmente, em razão da organização do conhecimento, para um modo específico de hermenêutica. Trata-se de abordar, como Cassin (2005), uma leitura logológica (centrada no discurso) da obra de Aristóteles, reunindo Tópicos, Retórica e Poética. Sob a visão todoroviana, podemos problematizar os papéis da teoria e da aplicação no tocante à interpretação do sujeito histórico que representa os saberes, compreende os artefatos e cria possibilidades para o seu acesso. Tanto Todorov (2014a), quanto Capurro (1992), quanto distintos teóricos da organização do conhecimento, “antigos”, como Tesouro (1670), e “novecentistas”, como Ranganathan (1967), bem como “contemporâneos, como Paling (2004), recorrem à Aristóteles, para, a partir das categorias (Tópicos), tecer o esclarecimento de enunciados e sua multiplicidade, ou seja, cada asserção, desde o Estagirita, é compreendida como uma implicação de várias consequências. Com o Aristóteles da Retórica e da Poética, esta

dimensão se torna ainda mais flagrante: a linguagem é variável. Neste sentido, segundo a visão todoroviana sobre o simbólico, podemos acatar a assertiva capurriana, transpondo a hipótese para o domínio, ou seja, a organização do conhecimento é uma subdisciplina da Retórica.

- b) A estrutura filosófica ranganathaniana à luz da teoria da linguagem do sânscrito: em razão das dificuldades que vão da apropriação da língua à compreensão da cultura, o pensamento de Ranganathan (1949, 1967, 2009) permite-nos ainda inúmeras interpretações. Sabemos que alguns conceitos, como *vastu-tantra* e *digvijaja*, não recebem tradução para o Ocidente, principalmente pela intersubjetividade local que carregam. A partir da compreensão de Todorov (2014a, p. 54), é com o modo de construção da gramática do sânscrito que encontramos elementos desconhecidos no mundo ocidental. Por exemplo, o reconhecimento das funções indiretas na abordagem sânscrita e permitem a identificação de fatos simbólicos no que, a nosso ver, participam da fundamentação de Ranganathan (1949, 1967, 2009), bem como são aplicadas no desenvolvimento de experiências na *web*, como o caso dos processos chamados “ontológicos”. Ranganathan (1949, 1967) abre, assim, não uma perspectiva de redução de relações, mas de multiplicação de potenciais verossimilhanças, antevistas pelo simbolismo todoroviano.
- c) “Discurso” como objeto da organização do conhecimento: a partir de Todorov (2014a) podemos reconhecer a potencialidade do “discurso” como categoria para a crítica social em organização do conhecimento. Trata-se aqui de perceber como as teorias e as experiências no domínio apontam para a compreensão do discurso como objeto central da organização do conhecimento. Esta lente sobre o fundamento objetual do domínio está presente em frentes teóricas como Frohmann (1990), Day (2005) e no próprio Capurro (1992). O fato que mais nos chama atenção na discussão está na “abertura” à crítica social na teoria e nas práticas de organização do conhecimento. Como lembra Todorov (2014a, p. 69), “a evocação simbólica é fundamentalmente múltipla”. Pensar os objetos como discurso (ou, à moda foucaultiana como desdobramentos chamados “formações discursivas”) permite uma amplitude no conjunto de diversidades culturais. Em outros termos, pensar o “discurso”, é pensar a alteridade. Como lembra Eco (2001, p. 213), o signo sobrevive sob a dialética do significante e do significado. O símbolo, por sua vez, existe a “ideia do reenvio que de certa forma encontra o seu próprio termo: uma reconjugação com a origem” – trata-se de conhecer o outro, sua cultura, para se aproximar dos significados de sua produção discursiva.
- d) Tesouro e a vida aberta do significante na metáfora: com Todorov (2014a) podemos retomar os “diálogos teóricos” entre Tesouro e Aristóteles no Século XVI e perceber suas possibilidades de aplicação na atualidade. O foco aqui está em reposicionar a vivência do discurso em sua produção constante de metáforas, fundamentais para compreender mais do que a “língua do outro”, ou seja, indo objetivamente à cultura de formação de suas linguagens, ao discurso propriamente dito. Tesouro e seu índice categórico coloca em discussão o simbólico nos potenciais modos de organização da linguagem, visando não apenas sua compreensão, a articulação de suas combinações, mas a possibilidade, como a organização do conhecimento busca, de reencontra-la (recupera-la).
- e) Os desafios de uma *web* pragmática: por fim, a abordagem todoroviana nos permite o encontro com os desdobramentos de uma rede digital sintática e semântica, ou seja,

estruturadas em relações de “diretas” (no vocabulário todoroviano) para a multiplicidade das funções “indiretas”. Trata-se de enfrentar a dinâmica de um ponto de vista da rede baseado estruturalmente no uso (Gracioso, Saldanha, 2010), como modo de reconfiguração constante do significante, ou seja, seu transbordamento absoluto. Trata-se, em outros termos, da grande mutação dos contextos, identificados por Jacob e Albrechtsen (1999) em sua análise pós-estruturalista dos sistemas de informação, identificando os elementos pragmáticos responsáveis por tais transformações.

Estas são algumas, porém não todas as possibilidades de reconhecimento das contribuições de Todorov para a organização do conhecimento. Seus estudos, acreditamos, à busca pela compreensão da linguagem segundo uma antropologia geral, chegando ao simbólico, ainda guardam muitas margens de investigação no campo.

4 CONCLUSÕES: AS “SUTILEZAS” DO SIMBÓLICO

“A sutileza na descrição do simbólico estará, a partir de então, na medida da sutileza que é manifesta naquilo que concerne às relações discursivas. [...]” (Todorov, 2014a, p. 82).

As conclusões apontam para a pertinência das abordagens simbólicas como fórum teórico para a reflexão contemporânea em organização do conhecimento, bem como uma tópica fundamental no plano epistemológico-histórico do estudo da representação do conhecimento e de seus instrumentos. Como na metodologia filosófica wittgensteiniana, podemos reconhecer o simbólico como uma espécie de “linguagem primitiva” nos estudos e nas práticas de organização do conhecimento.

Em outros termos, a partir da reflexão, trata-se de tomar o que Todorov (2014a) chama de “sutilezas” do universo simbólico como um exercício epistemológico fundamental e emergencial a organização do conhecimento. Tais “sutilezas” abrem, em nossa visão, um amplo horizonte de abordagens críticas e sócio culturais, dada a prerrogativa da alteridade presente nas abordagens simbólicas, com visto em Todorov (2014b).

Retomando, as dimensões centrais para a aproximação das ideias todorovianas estão, principalmente, em cinco dimensões em aberto, a saber, o papel de uma teoria da linguagem em Ranganathan, a assertiva hipotético-epistemológica capurriana da centralidade retórica na fundamentação da Ciência da Informação, a potencialidade do discurso como categoria para a crítica social em organização do conhecimento, a atualidade das ideias de Tesouro e os desafios de uma *web pragmática*.

AGRADECIMENTOS

A pesquisa foi desenvolvida com o fomento do Conselho Nacional de Desenvolvimento Científico e Tecnológico do Brasil (CNPq) e da Fundação Carlos Chagas de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ).

REFERÊNCIAS BIBLIOGRÁFICAS

- Blair, D. C. (1992). Information retrieval and the philosophy of language. *The Computer Journal*, v. 35, n. 3, p. 200-207.
- Cassin, B. (2005). *O Efeito sofisticado*. São Paulo: ed. 34.
- Cassirer, E. (2011) *A filosofia das formas simbólicas*; Terceira parte: Fenomenologia do conhecimento. São Paulo: Martins Fontes.
- Cassirer, E. (2004). *A filosofia das formas simbólicas*; Segunda Parte: O pensamento mítico. São Paulo: Martins Fontes, 2004.
- Cassirer, E. (2001). *A filosofia das formas simbólicas*; Primeira parte: A linguagem. São Paulo: Martins Fontes.
- Cassirer, E. (1994). *Ensaio sobre o homem: introdução a uma filosofia da cultura humana*. São Paulo: Martins Fontes.
- Capurro, R. (1992) What is Information Science for? a philosophical reflection In: Vakkari, P.; Cronin, B. (Ed.). *Conceptions of Library and Information Science; historical, empirical and theoretical perspectives*. In: International conference for the celebration of 20th anniversary of the department of information studies, University of Tampere, Finland. 1991. *Proceedings...* London, Los Angeles: TaylorGraham. p.82-96.
- Day, R. (2005). Poststructuralism and information studies. *Annual review of information science social and technology (ARIST)*, v. 39, p. 575-609.
- Eco, U. (2001). *Semiótica e filosofia da linguagem*. Lisboa: Instituto Piaget.
- Feinberg, M. (2010). Two kinds of evidence: how information systems form rhetorical arguments. *Journal of Documentation*, v. 66, n. 4, p. 491-512.
- Foucault, M. (2010). *A Arqueologia do saber*. 7. ed. Rio de Janeiro: Forense Universitária.
- Frohmann, Bernd (1990). Rules of indexing: a critique of mentalism in information retrieval theory, *Journal of Documentation*, Vol. 46 Iss 2 pp. 81 – 101.
- García gutiérrez, A. (2011). Desclassification in Knowledge Organization: a post-epistemological essay. *Transformação*, Campinas, v.23, n.1, p. 5-14.
- González de Gómez, M. N. (1996). Da organização dos saberes às políticas de informação. *Informare – Cad. Prog. Pós-grad. Ci. Inf.*, Rio de Janeiro, v. 2, n. 2, p. 58-66.
- Gracioso, L. de S.; Saldanha, G. S. (2010). *Ciência da Informação e Filosofia da Linguagem: da pragmática informacional à web pragmática*. Rio de Janeiro: ed. do autor.
- Jacob, E.K.; Albrechtsen, H. (1999). When essence becomes function: Post-structuralist implications for an ecological theory of organisational classification systems. In T.D. Wilson & D.K. Allen. *Exploring the contexts of information behaviour: proceedings of the second international conference on*

research in information needs, seeking and use in different contexts. London: Taylor Graham. p. 519-534.

Hjorland, B. (2002). Domain analysis in information science: eleven approaches traditional as well as innovative. *Journal of Documentation*, Londres, v. 58, n. 4, p. 422-462.

Paling, S. (2004). Classification, rhetoric and the classificatory horizon. *Library Trends*, v. 52, n. 3, p. 588-603.

Ranganathan, S. R. (2009). *As Cinco leis da Biblioteconomia*. Brasília: Briquet de Lemos.

Ranganathan, S. R. (1967). *Prolegomena to Library Classification*. 3d ed. New York: Asia Publishing House.

Ranganathan, S. R. (1949). Self-perpetuating scheme of classification, *Journal of Documentation*, v. 4., n. 4, p. 223 – 244.

Souza, R. F. (1996). Organização do Conhecimento em uma estrutura classificatória no contexto da indexação e recuperação da informação. *INFORMARE*, RIO DE JANEIRO, v. 2, n.2, p. 37-49.

Tesauro, E. (1670). *Il canocchiale Aristotélico*. Berlin: Verlag Gehlen; Zürich: Bad Homburg.

Todorov, T. (2014a). *Simbolismo e interpretação*. São Paulo: Editora Unesp.

Todorov, T. (2014b). *A vida em comum: ensaio de antropologia geral*. São Paulo: Unesp.

Wittgenstein, L. (1979). *Investigações Filosóficas*. 2o ed. São Paulo: Abril Cultural.

Wittgenstein, L. (1992). *O livro azul*. Lisboa: Ed.70.



CONTRIBUIÇÕES AO ESTUDO DA ORDENAÇÃO DE DOCUMENTOS PARA ARQUIVÍSTICA

Clarissa M. S. Schmidt¹, Renato de Mattos²

¹Universidade Federal Fluminense, 0000-0003-1555-4594, clarissaschmidt@id.uff.br

²Universidade Federal Fluminense, 0000-0001-7993-8672, Renato_mattos@id.uff.br

RESUMO As recentes pesquisas acerca da organização do conhecimento arquivístico apresentam a classificação e a descrição como seus procedimentos elementares. Ambas as funções têm por objetivo, respectivamente, demonstrar o contexto de produção dos documentos e representar as informações que identificam o acervo, visando seu uso e disseminação. Entretanto, nenhuma delas preocupa-se com a organização, intelectual e física, dos documentos em suas unidades de arquivamento de maneira a facilitar seu uso e recuperação. Mediante a isso, entendemos que a ordenação de documentos de arquivo é a atividade destinada à disposição dos documentos com vistas a recuperá-los, sendo fundamental para organização da informação na Arquivística. Todavia, é importante destacar que não encontramos, na literatura especializada da área, discussões consistentes sobre o conceito de ordenação, sua aplicação prática e importância. Nesse sentido, os problemas que justificam as análises propostas neste trabalho são a coexistência de diferentes e poucas definições sobre o significado de ordenação de documentos, além de as distintas maneiras de compreensão de sua aplicação prática. Por conseguinte, buscamos com este estudo analisar como a ordenação está sendo significada pela Arquivística, bem como seus desdobramentos no tocante às funções arquivísticas classificação e descrição. Em termos metodológicos, torna-se importante demarcar que nosso campo teórico será o conceito de ordenação de documentos e a compreensão acerca de sua aplicação prática na Arquivística. Do ponto de vista empírico, trata-se de pesquisa bibliográfica e de natureza qualitativa. Deste modo, espera-se contribuir para equacionar lacunas de definição e operacionalização sobre a temática em questão, permitindo a ampliação das discussões no âmbito da organização do conhecimento arquivístico.

PALAVRAS-CHAVE *Organização do Conhecimento Arquivístico, Ordenação de documentos de arquivo, Aplicação prática da ordenação de documentos de arquivo.*

ABSTRACT The recent researches regarding archival knowledge organization demonstrate the classification and description of elementary procedures. Both are developed with the objective to, respectively; demonstrate the context of document production and to represent the information which identifies the collection, aiming its use and dissemination. However, none of them is concerned with the intellectual and physical organization of the documents in its archives facilities in a way to ease their use and recovery. Thus, we understand that ordering archival documents is an activity directed to the availability of the documents aiming to recovering being fundamental to information organization in Archival Sciences. On the other hand, it is important to highlight that consistent discussion or arguments were not found in specific related literature about the concept of ordering and its application in theory and practice. In this way, the issues which justify the proposed analysis in the present investigation are: the coexistence of different definitions about the meaning of ordering documents as well as the different ways of understanding its practical applications. Therefore, we hope with this study to analyze how the ordering is being signified by the Archival Sciences, as well as its unfolding in relation to the archival classification and description functions. In methodological terms, it is important to point out that our theoretical field will be the concept of document ordering and the understanding of its practical application in Archival Sciences. From the empirical point of

view, it is bibliography and qualitative research. Thus, we expect to contribute to equalize the definition and operationalization gaps concerning this issue, allowing a broader scope discussion in archival knowledge organization.

KEYWORDS *archival knowledge organization, ordering archival documents, Practical application of file ordering.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

1 INTRODUÇÃO

As recentes pesquisas acerca da organização do conhecimento arquivístico apresentam a classificação e a descrição como seus procedimentos elementares (Silva, Moreira, Guimarães & Moraes 2015; Tognoli 2013; Barros, 2013; Barros&Tognoli, 2015; Schmidt&Smit 2013, 2015; Orrico&Silva, 2016). Tal afirmação está fundamentada pela possibilidade de tais funções “permitirem que um conhecimento produzido pela sociedade e devidamente registrado em documentos de natureza orgânica possa ser apropriado por essa sociedade para a geração de novos conhecimentos” (Silva, Moreira, Guimarães & Moraes, 2015, p. 461). Assim, entendemos tratar-se de processos que estão a serviço da organização dos documentos e das informações neles contidas, sendo orientados por teorias e metodologias.

Classificação e descrição têm por objetivo, respectivamente, demonstrar o contexto de produção dos documentos e ser a “ponte que comunica o documento com os usuários” (Heredia Herrera, 1991a, p. 300, tradução nossa). Entretanto, nenhuma delas preocupa-se com a organização, intelectual e física, dos documentos em suas unidades de arquivamento de maneira a facilitar seu uso e recuperação. Mediante a isso, entendemos que a ordenação de documentos de arquivo é a atividade arquivística capaz de suprir tal preocupação, sendo fundamental para organização da informação na Arquivística.

Nessa linha de pensamento, compreende-se que, se tais procedimentos e atividade não forem operados de maneira conjunta, pode-se dizer que não existe organização efetiva, porque é possível manter arquivos bem classificados e descritos, mas mal ordenados, e vice-versa. A organização é verificada não só ao examinarmos fisicamente os documentos, ela também pode ser representada nos instrumentos de gestão (plano de classificação, quadro de arranjo) ou instrumentos de pesquisa como guias, inventários, dentre outros (González, 2003).

Todavia, é importante destacar que não encontramos, na literatura especializada da área, discussões consistentes sobre o conceito de ordenação, sua aplicação prática e importância na Arquivística. O arquivista italiano Elio Lodolini (1993, p. 115, tradução nossa), por exemplo, apresenta como um problema fundamental na Arquivística, a ordenação – para conservar de forma permanente e para dar acesso – dos documentos nos arquivos históricos, questionando as dificuldades, dos profissionais dos arquivos, em ordenar documentos produzidos durante vários séculos, ou ainda, várias décadas.

Nesse sentido, os problemas que justificam as análises propostas neste trabalho são a coexistência de diferentes e poucas definições sobre o significado de ordenação de documentos, além de as distintas maneiras de compreensão de sua aplicação prática. Por conseguinte, buscamos com este estudo analisar como a ordenação está sendo significada pela Arquivística, bem como seus desdobramentos no tocante às funções arquivísticas classificação e descrição. Deste modo, espera-se contribuir para equacionar

lacunas de definição e operacionalização sobre a temática em questão, permitindo a ampliação das discussões no âmbito da organização do conhecimento arquivístico.

Por fim, em termos metodológicos, torna-se importante demarcar que nosso campo teórico será o conceito de ordenação de documentos e a compreensão acerca de sua aplicação prática na Arquivística. Do ponto de vista empírico, trata-se de pesquisa bibliográfica e de natureza qualitativa. A análise dos conceitos se dará em dois casos específicos: dicionários de especialidade brasileiros e espanhóis. Já no tocante à compreensão do uso de ordenação de documentos na prática e funções arquivísticas, tal pesquisa será realizada em literatura da área, também brasileiras e espanholas.

2 ORDENAÇÃO DE DOCUMENTOS DE ARQUIVO

De acordo com o Dicionário de Terminologia Arquivística de Camargo & Belloto (1996, p. 64), entende-se por Ordenação a “disposição dos documentos de uma série, a partir de elemento convencionado para sua recuperação”. O Dicionário de Biblioteconomia e Arquivologia, de Cunha & Cavalcanti (2008, p. 270), considera a ordenação como a “operação de arranjo metódico de documentos, segundo um plano de classificação, com a finalidade de conservá-los e, se for o caso, recuperá-los posteriormente”. Já no Dicionário Brasileiro de Terminologia Arquivística (2005, p.117) não encontramos o verbete ordenação, mas sim, como termo equivalente, métodos de arquivamento, sendo esse definido como a “seqüência de operações que determina a disposição dos documentos de um arquivo ou coleção, uns em relação aos outros, e a identificação de cada unidade”.

Ao analisarmos as definições destes dicionários especializados na área, observamos que há diferentes compreensões acerca do termo, bem como de seus significados. O dicionário de Camargo & Belloto não determina em que função arquivística, ou ainda em que fase do ciclo de vida do documento, a ordenação deve ser efetuada, afirmando apenas que é uma atividade a ser realizada junto às séries documentais, de modo a permitir que seja possível a recuperação – individual ou seriada – do documento de arquivo. Por seu turno, a definição apresentada por Cunha & Cavalcanti condiciona a ordenação à função classificação, delimitando que sua operacionalização deve ser resultado de um instrumento de classificação – o plano de classificação. Estabelece uma sinonímia entre ordenação e arranjo, coloca como finalidade a conservação do documento e, de maneira secundária, sua recuperação. Vale ressaltar que esta definição também não determina em que fase do ciclo de vida do documento a ordenação deve ser efetuada.

Por fim, de modo a contemplar as definições dos dicionários de especialidade brasileiros, analisamos a do Dicionário Brasileiro de Terminologia Arquivística. Nesta obra de referência, elaborada pelo Arquivo Nacional, não há alusão ao termo ordenação. Entretanto, observamos que o verbete métodos de arquivamento contempla definição equivalente. Nesse sentido, o significado apresentado não estabelece em que função arquivística a ordenação deve acontecer, tampouco em qual fase do ciclo de vida do documento. Diferente das definições dos dicionários acima mencionados, não faz relação direta com a recuperação do documento, afirmando apenas tratar-se de uma operação que visa distribuir os documentos de maneira que seja possível sua identificação.

Já em relação aos dicionários espanhóis, iniciamos com Antonia Heredia Herrera, em seu *Lenguaje y vocabulario archivísticos algo más que un diccionario* (2011, tradução nossa), no qual define ordenação como a operação que consiste em unir elementos ou unidades de um conjunto,

relacionando-os uns com os outros de acordo com a unidade de ordem estabelecida anteriormente. A obra de outro arquivista espanhol de referencia, o *Diccionario de Archivística*, de José Ramón Cruz Mundet (2011, p. 265, tradução nossa), apresenta como ordenação a “operação arquivística integrada na organização de fundos, que consiste em relacionar uns elementos com outros, de acordo com um critério estabelecido anteriormente, como por exemplo datas, letras do alfabeto ou números”. Já o *Diccionario de Terminologia Arquivística do Governo Espanhol* (1993, tradução nossa), define a ordenação como sendo a “operação arquivística realizada dentro do processo de organização, que consiste em estabelecer sequencias naturais croológicas e ou alfabéticas, dentro das categorias e grupos definidos na classificação. Se encontra, portanto, dentro da fase de tratamento arquivístico denominada identificação”.

Diferente das definições apresentadas pelos dicionários brasileiros, as da literatura espanhola estão mais equivalentes e em concordância umas com as outras, ainda que pequenas especificidades sejam observadas. Os significados estabelecidos por Heredia Herrera e Cruz Mundet são muito próximos. Ambos não esclarecem em que função arquivística, ou em qual fase do ciclo de vida dos documentos, a ordenação deve ser operacionalizada, como também não mencionam a recuperação do documento como um de seus elementos. Entretanto, é possível identificar que os dois autores afirmam que a unidade de ordem deve ser estabelecida antes que a ordenação seja efetuada, sendo que apenas Cruz Mundet indica exemplos destes critérios.

Tanto no dicionário de Cruz Mundet, como no do Governo Espanhol, encontramos a ordenação como sendo uma etapa da organização de documentos de arquivo, além de exemplos de critérios para a ordenação. Todavia, apenas o dicionário do Governo Espanhol afirma que a ordenação deve ser colocada em prática após a classificação, estando inserida na etapa da identificação arquivística. E, assim como os demais, não faz referencia a fase do ciclo de vida do documento na qual a ordenação deve estar presente.

Compreendidas e observadas a coexistencia de distintas definições para o termo ordenação, se faz importante analisar como a literatura da área o significa em relação a sua operacionalização prática e com as funções arquivísticas.

3 A ORDENAÇÃO E AS FUNÇÕES ARQUIVISTICAS CLASSIFICAÇÃO E DESCRIÇÃO

É fundamental, após apresentar e analisar definições acerca do significado de ordenação, estabelecer distinção sta atividade e as funções arquivísticas classificação e descrição, uma vez que há entendimentos de que entre a ordenação e a classificação exista uma relação de sinônimos, principalmente no tocante à classificação - devido à relação intrínseca que possuem, além de discutir acerca de sua aplicação prática na Arquivística.

Nesse sentido, Heredia Herrera (1991a, p. 107, tradução nossa) afirma que, enquanto atividades técnicas próprias do arquivista, a classificação é anterior à ordenação. Todavia, ainda de acordo com a autora (1991a, p. 254, tradução nossa), no *Elsevier's Lexicon of Archive Terminology*, de 1964, classificação e ordenação são consideradas como sinônimos. Além disso, Heredia Herrera (1991b, p. 109, tradução nossa), é categórica ao considerar que “classificar é separar ou dividir um conjunto de elementos, estabelecendo classes ou grupos. Já ordenar é unir todos os elementos de cada grupo, estabelecendo

uma ordem única, que pode ser data, alfabeto, tamanho ou número”. E que, “frente à organização, atividade profunda e intelectual, a ordenação responde a métodos mais mecânicos, exige menos preparação (...)”.

Para Jiménez Gonzalez (2003, p. 6, tradução nossa), “classificação e ordenação não admitem igual significado, ao passo que a organização de documentos de arquivo é um resultado, em primeiro lugar, da classificação e, em seguida, da ordenação”. Se ambos estes processos não são realizados de maneira complementar, “deve-se dizer que não existe uma organização adequada, pois pode-se manter arquivos bem classificados, mas mal ordenados (...) pode-se dizer que a ordenação começa quando a classificação for concluída, isto é, aplicada à série documental.” Aurelio Tanodi (1981, p. 8, tradução nossa), ao distinguir classificação de ordenação, afirma que esta última é a atividade de colocar em ordem os documentos e que, “o método principal, fundamental, de ordenar os documentos, é sua boa classificação”. Por fim, o autor afirma ainda que “existe certa confusão entre os conceitos de classificação e de ordenação arquivísticas” (Tanodi, 2009, p. 172, tradução nossa).

Ainda discutindo acerca da função classificação e da atividade de ordenação, para Marilena Leite Paes (2002, p. 61), a “tarefa de classificar documentos para um arquivo exige do classificador conhecimentos não só da administração a que serve, como da natureza dos documentos a serem classificados. Cada ramo de atividade exige um método diferente, adequado às suas finalidades”. Continua sua reflexão afirmando que isso se configura em um problema difícil, ou seja, “quando se quer organizar um arquivo, da escolha de um método ideal de classificação para que a finalidade precípua do arquivo, que é o acesso aos documentos, seja plenamente atingida. O método de arquivamento é determinado pela natureza dos documentos a serem arquivados e pela estrutura da entidade”. Analisando a fala da autora, pontuamos que para ela há um vínculo entre classificação e ordenação – que ela denomina de métodos de arquivamento, assim como o Dicionário Brasileiro de Terminologia Arquivística -, sendo que a operacionalização da classificação acontece através dos métodos de arquivamento (ordenação).

Esse mesmo vínculo pode ser observado quando o Conselho Nacional de Arquivos (2001, p. 15) problematiza a ordenação, afirmando que esta “consiste na reunião dos documentos classificados sob um mesmo assunto. A ordenação tem por objetivo agilizar o arquivamento, minimizando a possibilidade de erros. Além disso, estando ordenados adequadamente, será possível manter reunidos todos os documentos referentes a um mesmo assunto, organizando-os previamente para o arquivamento”.

Já com uma abordagem que diferencia a classificação da ordenação, Camargo & Bellotto (1996, p. 30), definem a classificação como sendo a *sequência de operações que, de acordo com as diferentes estruturas, funções e atividades da entidade produtora, visam a distribuir os documentos de um arquivo*. Logo, para estas autoras, no campo dos arquivos, a classificação não é realizada objetivando o acesso aos documentos, mas sim, na compreensão da gênese documental, cabendo à ordenação, juntamente da descrição quando dos arquivos permanentes, a recuperação dos documentos.

Nessa linha de pensamento, para Renato Tarciso B. de Sousa (2007, p. 85), a classificação é “a ação intelectual de construir esquemas para agrupar os documentos a partir de princípios estabelecidos”, e a ordenação é a “forma de disposição dos tipos documentais dentro das divisões estabelecidas no esquema de classificação”. O autor discute ainda, de maneira bastante interessante, que é possível “dividir o processo classificatório em duas partes: a parte intelectual e a parte física. A parte intelectual se refere à classificação propriamente dita (processo mental de estabelecimento de classes) e à ordenação (a

disposição dos documentos nas classes estabelecidas). A parte física é representada pelo arquivamento dos documentos em um local determinado pela classificação e disposto segundo uma ordem definida”.

Em relação à descrição, ainda segundo Camargo & Bellotto (1996, p. 36), entende-se como sendo *o conjunto de procedimentos que, a partir de elementos formais e de conteúdo, permitem a identificação de documentos e a elaboração de instrumentos de pesquisa*. Elio Lodolini (1991, p. 36, tradução nossa), ao falar da ordenação, explica que não está se referindo a disposição dada aos documentos em sua origem, quando estão no arquivo corrente, mas sim a ordem que devem dar, os arquivistas, quando os documentos chegam à última e definitiva fase de sua vida, ou seja, aos arquivos permanentes, o que nos leva a considerar as relações da ordenação com a função descrição.

Apresentadas as reflexões acerca da ordenação e as funções arquivísticas, destacamos que os procedimentos de classificação, ordenação e descrição, não requerem os mesmos conhecimentos para serem aplicados, ou seja, as mesmas metodologias, uma vez que não têm os mesmos objetivos. É propósito da classificação arquivística demonstrar o contexto de produção dos documentos, as funções e atividades que lhe deram origem, sendo possível, também, compreender os vínculos orgânicos deste com os demais de seu conjunto e com o órgão de proveniência. Segundo Gonçalves (1998, p. 12), o objetivo da função classificação é, fundamentalmente “dar visibilidade às funções e às atividades do organismo produtor do arquivo, deixando claras as ligações entre os documentos. Podemos entender que a classificação é, antes de tudo, lógica: a partir da análise do organismo produtor de documentos de arquivo, são criadas categorias, classes genéricas, que dizem respeito às funções/atividades detectadas”. E quanto à descrição, entendemos que “garante a compreensão ampla do conteúdo de um acervo, possibilitando tanto o conhecimento como a localização dos documentos que o integram” (Lopez, 2002, p.12). Já a ordenação tem por finalidade facilitar e agilizar a consulta aos documentos, isto é, estabelecer previamente elementos quando da necessidade de arquivamento do documento, de maneira a permitir sua efetiva recuperação. Vale ressaltar que tais elementos são retirados do conteúdo do documento, variando por série documental. Novamente de acordo com Gonçalves (1998, p. 12), o objetivo da ordenação no campo dos arquivos “é facilitar e agilizar a consulta aos documentos, pois, mesmo no que se refere a uma mesma atividade, e em relação a um mesmo tipo documental, os documentos atingem um volume significativo. A adoção de um ou mais critérios de ordenação para uma série documental permite evitar, em princípio, que, para a localização de um único documento, seja necessária a consulta de dezenas ou centenas de outros”.

O âmbito de aplicação da classificação e da descrição são os arquivos ou os fundos, ao passo que a ordenação é aplicada nas séries documentais, por isso podemos afirmar que não se ordenam fundos e nem arquivos, reforçando nossa percepção acerca de suas diferenças teóricas, metodológicas e práticas. Nesse sentido, em termos de operacionalização da ordenação, isto é, sua aplicação prática, entendemos ser uma atividade que cabe às séries documentais, seja após a classificação nos arquivos correntes, ou após a descrição nos arquivos permanentes.

4 CONSIDERAÇÕES FINAIS

Entende-se que a ordenação de documentos é tarefa fundamental para organização da informação em arquivos e, desta maneira, necessita ter seu conceito e fundamentos bem delimitados, assim como devem ser conhecidas e compreendidas sua importância e aplicação prática. A partir de tais demarcações será possível, efetivamente, inseri-la nas discussões de destaque na Arquivística. Nesse

sentido, nossas análises junto aos dicionários de especialidade indicam que a ordenação de documentos na Arquivística é (i) considerada sinônimo de arranjo; (ii) vista como semelhante aos métodos de arquivamento; (iii) entendida como uma etapa da classificação arquivística; (iv) utilizada enquanto método de classificação arquivística; (v) percebida como o arquivamento dos documentos físicos; (vi) utilizada para a recuperação de unidades ou séries documentais; (vii) uma atividade de organização de arquivos.

Em relação à compreensão de sua aplicação prática e relação com as funções arquivísticas, pode-se destacar que a bibliografia analisada sugere que, (i) diferente da função classificação, cujo papel do arquivista está em representar o contexto de produção dos documentos, e ainda da descrição, que busca garantir uma compreensão ampla do acervo, na ordenação o arquivista deverá criar modalidades a partir da necessidade de uso destes; (ii) não há consenso acerca da aplicabilidade da ordenação na tarefa de organização de documentos; (iii) há desconhecimento sobre a relação da ordenação com a classificação e a descrição; (iv) os tipos de uso e busca do documento influenciam na escolha da modalidade de ordenação; (v) a ordenação é feita com base no conteúdo dos documentos; (vi) o ato de arquivar está diretamente vinculado com a ordenação e não com a classificação ou a descrição.

Apresentados os resultados das análises, é importante ressaltar que a classificação, a ordenação e a descrição devem ser articuladas e não excluídas, pois fazem parte do processo de organização e recuperação de documentos e informações. Ademais, observa-se ausência de pesquisas e reflexões teóricas acerca do papel e importância da ordenação, bem como aproximações com as áreas de organização do conhecimento arquivístico.

Nesse contexto, cumpre estabelecer, a título inicial, a necessidade da ordenação de documentos possuir instrumento de gestão específico, tal como o plano de classificação, a tabela de temporalidade de documentos, e ainda os instrumentos de pesquisa. Afora isso, cabe à comunidade arquivística inserir a ordenação de documentos de arquivo no âmbito de suas investigações, porquanto de seu importante papel na organização da informação.

REFERÊNCIAS BIBLIOGRÁFICAS

Barros, T. (2013). Perspectivas de renovação em classificação arquivística: novos olhares velhos métodos. In F. Ribeiro & M. E. Cerveira (org), *Informação e/ou Conhecimento: as duas faces de Jano*. Actas del I Congresso ISKO Espanha e Portugal / XI Congresso ISKO España (pp. 113-128). Porto: Faculdade de Letras da Universidade do Porto - CETAC.MEDIA.

Bellotto, H. L. (2004). *Arquivos permanentes*. Tratamento documental. Rio de Janeiro: FGV.

Bellotto, H. L., & Camargo, A. M. de A. (1996). *Dicionário de terminologia arquivística*. São Paulo: Associação de Arquivistas de São Paulo.

Conselho Nacional de Arquivos (2001). *Classificação, temporalidade e destinação de documentos de arquivo*; relativos às atividades-meio da administração pública. Rio de Janeiro: Arquivo Nacional.

Cruz Mundet, J.R. (2011). *Dicionário de Archivística*. Madri: Alianza Editorial.

Cunha, M. B. da & Cavalcanti, C. R. de O. (2008). *Dicionário de Biblioteconomia e Arquivologia*. Brasília: Briquet de Lemos.

Diccionario de Terminologia Archivística (1993). Madrid: Ministerio de Cultura / Dirección de Archivos Estatales.

Dicionário brasileiro de terminologia arquivística (2005). Arquivo Nacional (Brasil). Rio de Janeiro.

Gonzalez, G. J. (2003). *Ordenación Documental*. Bogotá: Archivo General de la Nación.

Gonçalves, J. (1998). *Como classificar e ordenar documentos de arquivo*. São Paulo: Arquivo do Estado.

Heredía Herrera, A. (1991a). *Archivística general: teoria y práctica*. Sevilla: Diputación Provincial.

Heredía Herrera, A. (1991b). Clasificación y ordenación. In Gutiérrez Muñoz, C. *Archivística. Materiales de enseñanza de la Facultad de Letras y Ciencias Humanas* (pp. 90-123) Lima: Pontificia Universidad Católica del Peru.

Heredía Herrera, A. (2011). *Lenguaje y vocabulario archivísticos algo más que un diccionario*. España: Junta de Andalucía.

Jiménez Gonzalez, G.(2003). *Ordenación documental*. Colombia: Archivo General de La Nación.

Lodolini, E. (1993). *Archivística*. Principios y problemas. Madrid: Anabad.

Lodolini, E. (1991). El problema fundamental de la archivística: la naturaleza y el ordenamiento del archivo. In Gutiérrez Muñoz, C. *Archivística. Materiales de enseñanza de la Facultad de Letras y Ciencias Humanas* (pp. 30-51) Lima: Pontificia Universidad Católica del Peru.

Lopez, A. P. A.. *Como descrever documentos de arquivo: Elaboração de instrumentos de pesquisa*. São Paulo: Arquivo do Estado, Imprensa Oficial, 2002.

Orrico, E.G.D., & Silva, E.P. (2016). Knowledge Organization in Archives: the Brazilian Case. In: J. A. C. Guimarães; S. O. Milani & V. Dodebei (org.), *Knowledge Organization for a sustainable world: challenges and perspectives for cultural, scientific, and technological sharing in a connected society* (pp. 508-514). Rio de Janeiro: Ergon Verlag.

Paes, M.L. *Arquivo: Teoria e Prática*. Ed. FGV, 2002

Schmidt, C., & Smit, J. (2013). Organização da informação e arquivos: diferentes perspectivas informacionais em torno do documento de arquivo. In F. Ribeiro & M. E. Cerveira (org), *Informação e/ou Conhecimento: as duas faces de Jano*. Actas del I Congresso ISKO Espanha e Portugal / XI Congreso ISKO España (pp. 579-588). Porto: Faculdade de Letras da Universidade do Porto - CETAC. MEDIA.

Schmidt, C., & Smit, J. (2015). Organização e representação da informação em arquivos: uma análise a partir da função classificação. In J.V. Rodríguez Muñoz; I. Gil-leiva; P. Díaz Ortuño; & F.J. Y Martínez Méndez (eds.), *Organización del conocimiento: sistemas de información abiertos*. Actas del XII Congreso ISKO España y II Congreso ISKO España y Portugal (pp. 579-583). Murcia: Universidad de Murcia.

Silva, A. P. da; Moreira, W.; Guimarães, J. A. C., & Moraes, J. B. E. de. (2015). Organização do conhecimento arquivístico: um estudo terminológico comparativo (português, espanhol, francês, inglês)

sobre classificação e descrição no Multilingual Archival Terminology. In J.V. Rodríguez Muñoz; I. Gil-leiva; P. Díaz Ortuño; & F.J. Y Martínez Méndez (eds.), *Organización del conocimiento: sistemas de información abiertos*. Actas del XII Congreso ISKO España y II Congreso ISKO España y Portugal (pp. 461-469). Murcia: Universidad de Murcia.

Sousa, R.T. B. *et al* (2007). *Arquivística: temas contemporâneos: classificação, preservação digital, gestão do conhecimento*. Distrito Federal: SENAC.

Tanodi, A (1981-1982). Introducción a la ordenación y clasificación. In *Archivística*, PNUD. Capacitación a distancia. Santiago de Chile.

Tanodi, A (2009). *Manual de Archivistica: teorías y principios*. Córdoba: Brujas.

Tognoli, N. B. (2013). A origem do método diplomático como subsídio à organização da informação arquivística: um estudo a partir das obras de Diplomática. In F. Ribeiro & M. E. Cerveira (orgs), *Informação e/ou Conhecimento: as duas faces de Jano*. Actas del I Congreso ISKO Espanha e Portugal / XI Congreso ISKO España (pp. 675-687). Porto: Faculdade de Letras da Universidade do Porto - CETAC. MEDIA.

Tognoli, N. B., & Barros, T. (2015). Os processos de representação do conhecimento arquivístico: elementos históricos e conceituais da classificação e descrição. In J. A. C. Guimarães & V. Dodebei (org.), *Organização do conhecimento e diversidade cultural* (pp. 94-99). Marília: ISKO-Brasil; FUNDEPE.



OS DISCURSOS INSTITUCIONAIS SOBRE INTERDISCIPLINARIDADE ANALISADOS POR MEIO DA HERMENÊUTICA E DE CATEGORIAS PMEST

Marco Donizete Paulino da Silva¹, Luciana de Souza Gracioso², Maria da Graça Melo Simões³

¹Universidade Federal de São Carlos, marco_donizete@yahoo.com.br

²Universidade Federal de São Carlos, lugracioso@yahoo.com.br

³Universidade de Coimbra, gsimoes@fl.uc.pt

RESUMO Pauta-se pelo cenário contemporâneo de intercâmbios dinâmicos de informação, considerando a interdisciplinaridade como instrumento essencial à eficiência dos processos comunicacionais entre setores institucionais, sobretudo quando estes se relacionam diretamente com o desenvolvimento científico. Propõe-se a analisar e refletir sobre os sentidos atribuídos ao termo Interdisciplinaridade por autores que abordam o conceito no plano institucional de pesquisa (ordem teórica) e de fomento (ordem normativo-avaliativa), buscando-se identificar a valorização de um sentido do termo em detrimento de outros, e o alinhamento dessa valorização pela conciliação entre os interesses desses planos institucionais. Partiu-se de uma pesquisa exploratória, sustentada pela combinação dos métodos de análise de conteúdo e análise hermenêutica, utilizando-se das categorias PMEST, de Ranganathan, para identificação de sentidos predicativos atribuídos ao termo, que, com base na Teoria Analítica do Conceito, de Dalberg, resultou uma caracterização conceitual do mesmo. O *corpus* compôs-se de dois seguimentos do capítulo 1 da obra *Interdisciplinaridade em Ciência, Tecnologia & Inovação*, organizada por Philippi Junior e Siva Neto (2011), que apresenta discursos em relação ao conceito de Interdisciplinaridade e seu uso no contexto educacional da pesquisa (programas de pós-graduação) e fomento (Capes. [COORDENADORIA de Aperfeiçoamento de Pessoal de Nível Superior]). Os resultados identificam variações terminológicas sobre tipos de relações disciplinares por grau de aplicação (uso de prefixos, tais como multi, pluri e trans) e por qualificação (uso de termo adicional, com função complementar), assim como convergências discursivas em algumas categorias e subcategorias (facetas). Conclui-se que houve interesse dos autores em demonstrar convergência em relação aos pressupostos teóricos apresentados e os critérios de avaliação da Capes, assim como possibilidade de uso pragmático do termo em contextos diferenciados.

PALAVRAS-CHAVE *Interdisciplinaridade; Categorização; PMEST; Tecnologia, Capes.*

ABSTRACT It is guided by the contemporary scenario of dynamic exchange of information, considering interdisciplinarity as an essential instrument for the efficiency of communication processes among institutional sectors, especially when these are directly related to scientific development. It proposes to analyze and reflect on the meanings attributed to the term Interdisciplinarity by authors that approach the concept in the institutional plan of research (theoretical order) and fomentation (normative-evaluative order), seeking to identify the valorization of a sense of the term to the detriment of others, and the alignment of this valuation by the conciliation between the interests of these institutional plans. It was based on an exploratory research, supported by the combination of the methods of content analysis and hermeneutical analysis, using the categories PMEST, from Ranganathan, to identify the predicative meanings attributed to the term, which, based on the Analytical Theory of Concept, Dalberg, resulted in a conceptual characterization of it. The corpus consisted of two chapters of the Chapter 1 of *Interdisciplinarity in Science, Technology & Innovation*, work organized by Philippi Junior and Siva Neto (2011), which presents discourses

in relation to the concept of Interdisciplinarity and its use as the educational context of research (post-graduation program) and research's funding (Capes. [COORDENADORIA de *Aperfeiçoamento de Pessoal de Nível Superior*]). The results identify terminological variations on types of disciplinary relationships by degree of application (use of prefixes such as multi, multi and trans) and by qualification (use of additional term, with complementary function), as well as discursive convergences in some categories and subcategories (facets). It is concluded that there was interest of the authors to demonstrate convergence in relation to the presented theoretical presuppositions and the evaluation criteria of Capes, as well as possibility of pragmatic use of the term in different contexts.

KEY-WORDS *Interdisciplinarity; Categorization; PMEST; Technology, Capes.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Partimos da percepção de que a diversidade de sentidos atribuídos ao termo Interdisciplinaridade, no atual e complexo cenário de trocas de informação, é uma realidade que se origina dos constantes impactos do desenvolvimento tecnológico no setor comunicacional. Tais variações de sentido do termo acarretaram, segundo Pombo (2005), um desgaste do mesmo em razão de seu uso excessivo (prática comunicativa) e da alta complexidade simbólica (teorização conceitual), tornando-se passível, a nosso ver, o risco da sua utilização instrumental pelo poder político, ou institucional, sem a devida consideração de suas características ou propriedades essenciais, uma vez que a Ciência, enquanto um campo de disputa - na acepção de Bourdieu (1989) -, é construída e, contextualmente influenciada, por uma série de forças que vão além de sua estrutura, e, portanto, lhes são determinantes das condições teóricas, metodológicas e epistemológicas pelas quais se guia.

No que concerne especificamente a um projeto de Ciência contemporânea, entendemos que tais impactos nos processos comunicacionais geraram maior dinamismo e fluidez dos intercâmbios informacionais, o que não assegurou, no entanto, a adequação da estrutura societária (agências de fomento) que dá sustentação ao desenvolvimento dos campos científicos, decorrendo uma dificuldade para absorção dos insumos em relação à velocidade com que estes são produzidos, uma vez que o tempo institucional de uma agência de fomento difere do tempo institucional de um programa de pesquisa. Essa dificuldade de atualização dos mecanismos de produção e absorção, a nosso ver, origina-se pela funções diferenciadas de ambos os setores institucionais.

No que concerne ao setor de fomento (pelas agências), consideramos que há grande pressão por parte dos governos que atrelam suas políticas de desenvolvimento científico ao critério econômico, já, no setor da pesquisa, a pressão recai sobre os perfis pelos quais os programas de pós-graduação, por exemplo, são categorizados, e as consequentes cobranças de produção de acordo com um caráter disciplinar, ou interdisciplinar, estabelecido pelas agências. Uma questão preponderante é se as exigências de enquadramentos institucionais a que os campos científicos estão submetidos pelos critérios das políticas de desenvolvimento científico – desenvolvimento almejado em modalidades disciplinares ou interdisciplinares (ou ainda, Multi, Pluri e Transdisciplinares) – nem sempre apresentam coerência quanto aos critérios estabelecidos como definidores dos perfis e das metas a que um programa de pós-graduação interdisciplinar deva atender – conforme advogado por Oliveira e Almeida (2011) ao relatarem incongruências dos processos de avaliação no Brasil.

Entendemos que os padrões estabelecidos por agências como a Capes [Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior] são os orientadores da ação dos programas de pós-graduação, e achamos legítimo analisar a variedade de sentidos atribuídos ao termo Interdisciplinaridade (e suas conformações circundantes) em enunciações discursivas produzidas a partir da apropriação desses padrões. Tais sentidos foram considerados passíveis de identificação por meio da análise de predicções conceituais imputadas ao termo Interdisciplinaridade (e similares), o que nos estimulou a decompor frases e parágrafos, selecionando predicados característicos atribuídos ao conceito, objetivando estabelecer-lhe uma identidade.

Para um melhor entendimento da problemática apresentada, propusemos-nos, como objetivo geral, analisar e refletir sobre os sentidos atribuídos ao termo Interdisciplinaridade através de estratégias de apropriação desses sentidos por autores de capítulo específico, que abordam o termo tanto no plano teórico (pelos teóricos-estudiosos do tema interdisciplinaridade) quanto no plano institucional (pelos limites de aplicação do conceito de interdisciplinaridade pela Capes); e, como objetivos específicos: i) identificar o valor atribuído a um dos sentidos do termo (um grau de predileção); ii) observar e compreender o alinhamento dessa valorização pela conciliação entre os planos teórico e institucional dos próprios autores do capítulo, ou seja, o posicionamento de Alvarenga et al. (2011) em relação à realidade teórico-institucional por eles analisada.

Nosso embasamento teórico se deu pela Teoria Analítica do Conceito, de Dahlberg (1978a; 1978b), entendendo que o processo de classificação, por essa teoria, possibilita a identificação de predicados possíveis para a compreensão de determinado objeto, ou seja, que a categorização conceitual: "[...] pode valer como pressuposto para a nossa necessidade de apreensão do conhecimento [...]" (1978b, p.3), aplicado aqui, sobre a noção de Interdisciplinaridade.

Como *corpus* de análise consideramos pertinente selecionar o texto *Histórico, fundamentos filosóficos e teórico-metodológicos da interdisciplinaridade*, de Alvarenga, Philippi Junior, Sommerman, Alvarez e Fernandes (2011) que aborda tanto questões relacionadas ao âmbito teórico quanto ao âmbito normativo desses contextos institucionais (respectivamente, pesquisa e fomento), observando, por meio da análise hermenêutica, como o termo Interdisciplinaridade foi articulado em meio às enunciações discursivas apresentadas pelos autores acerca da questão teórica (representando o embasamento da pesquisa científica) e da questão normativa (representado pelas citações creditadas à Capes).

As considerações de Alvarenga et al. (2011) – os autores acima citados - sobre a visão institucional da Capes acerca da Interdisciplinaridade - visão expressa em blocos de citações apresentadas durante o texto – nos permitiram interpretar que a construção discursiva da interdisciplinaridade, de um ponto de vista institucional, se dá por meio de definições e preconizações acerca de aplicação do conceito como fator de estímulo ao progresso.

Visando sistematização das enunciações de Alvarenga et al. (2011) sobre tais citações, são apresentadas as seguintes interpretações dos blocos de citações, nas seguintes condições:

- i. (composto de três citações): página 30, citação 1, ilustração do valor “alternativo e complementar” da interdisciplinaridade na produção científica; citação 2, capacidade peculiar da Interdisciplinaridade em lidar com fenômenos complexos; e, página 31, citação 3, as duas premissas anteriores são assinaladas como elementos essenciais para superação de um desenvolvimento disciplinar/especializado limitado; e,

- ii. (composto de quatro citações): em que a sequência de citações (página 62-63) resulta a seguinte enunciação sobre a pesquisa interdisciplinar: 1) exigência de certa “flexibilidade”, por parte do pesquisador, no processo investigativo interdisciplinar, desprendimento de “esquemas definidores rígidos”, aplicação dos princípios de interdisciplinaridade em graus variáveis; 2) sentimento de identificação do critério linguístico como linha de investigação de suas próprias abordagens (jantschiana).

Segundo Alvarenga et al. (2011) a perspectiva de Jantsch (1972) é embasada numa “axiomática comum”, expressão definida pelos autores (por meio de uma paráfrase) como: “[...] determinado esquema de referência teórico-metodológico elaborado por cientistas de diferentes áreas [...] capaz de embasar a problematização de um dado tema complexo de pesquisa [...] nos vários momentos que caracterizam o processo investigativo.” (p. 63). Tais enunciações constarão como elementos norteadores de nossa conclusão.

Intuímos, a partir da análise empreendida, que a Interdisciplinaridade pode ocupar uma posição mais relacionada a um uso pragmático – um uso passível de entendimento sobre as possibilidades de significação da Interdisciplinaridade -, admitindo que o conceito possa adquirir um sentido tanto de “Tecnologia” (que possibilita, de maneira eficaz, estabelecer o processo de interação entre campos disciplinares, orientando-se pelo grau de intervenção mútua entre os agentes participantes do processo) quanto de “espaço territorial” de fronteira (espaços de articulação, ainda não explorados e potencialmente produtivos). Observamos, no entanto, que tais possibilidades devem ser ainda verificadas de maneira mais aprofundada em futuras explorações sobre o tema.

Também foi possível interpretar que dentre as abordagens teóricas analisadas por Alvarenga et al. (2011), os mesmos adotam uma constante valorização dos aspectos desenvolvimentistas típicos da ordem institucional de fomento, trazendo para o centro do debate uma ordem mais aplicada que teórica.

METODOLOGIA

Buscamos compreender o uso do conceito de Interdisciplinaridade por meio da categorização dos sentidos que lhe são atribuídos por meio de predicções ao mesmo, identificadas no texto *Histórico, fundamentos filosóficos e teórico-metodológicos da interdisciplinaridade*, de Alvarenga, Philippi Junior, Sommerman, Alvarez e Fernandes (2011, pp. 3-68).

Procuramos estabelecer os alcances e os limites deste conceito no plano discursivo dos autores; para tanto optamos pelo paradigma metodológico, de caráter exploratório e de base qualitativa, fundamentado nas preconizações de May (2004), por sua consideração do perfil sociológico atribuído aos objetos de pesquisa pelos quais propõe investigação; e Bauer e Gaskell (2008), pela ênfase desses autores na questão textual como componente principal de avaliação. Consideramos que tais referências consubstanciam as exigências que o objeto – o conceito de interdisciplinaridade – apresenta enquanto elemento discursivo em uso em documento que circula nas diversas esferas institucionais (fomento e pesquisa).

Optamos pela combinação de dois métodos, o de Análise de Conteúdo (Bauer, 2008, pp. 189-217), que pode ser descrito sucintamente como leitura e seleção de unidades textuais de interesse exploratório para a pesquisa (no caso as enunciações sobre Interdisciplinaridade), com verificação das condições sintáticas e semânticas dessas unidades, e o de Análise Hermenêutica (May, 2004, pp. 212-215), que,

sinteticamente, estabelece procedimentos de interpretação textual (numa diversidade de formatos) a partir da explicitação dos perfis envolvidos na articulação discursiva, objetivando verificar em que ponto a apropriação discursiva de argumentações e definições – no nosso caso, referenciadas pelos contextos de setores de fomento (base normativa) e pesquisa (base teórica), ambos de caráter institucional - se apresentam afetadas pelos sentidos adquiridos em plena enunciação, pela consideração dos respectivos contextos, sobretudo pela possibilidade dessa enunciação promover outros interesses de ordem não científica (política, por exemplo).

O processo de categorização foi considerado como um instrumento auxiliar da análise, pois, a nosso ver, promove uma axiomática comum que expressa tanto o valor essencial quanto o valor social do termo, passível de compartilhamento discursivo por meio de uma compreensão pragmática de seu sentido. Também arguimos que, pelo processo de categorização nos foi possível uma reconstrução conceitual do termo, lançando mão das predicções imputadas ao mesmo e de um critério de distribuição das mesmas.

Neste sentido, para a categorização do conceito Interdisciplinaridade, optou-se pelo princípio PMEST - *Personality* (P), *Matter* (M), *Energy* (E), *S* (*Space*) e *Time* (T) -, princípio pertencente à Teoria facetada de Ranganathan, teoria orientada para uma abordagem analítico sintética com vista à identificação de assuntos em documentos e que, segundo Dahlberg(1972): “[...] implica em que a classificação de cada documento exige uma análise de seu título ou um enunciado descritivo de seu conteúdo em termos de conceitos [...]” (p.4).

Em nossa proposição, o *corpus* será a base para conceituação do termo Interdisciplinaridade (e suas variantes), identificando predicções que foram categorizadas pelas facetas PMEST, permitindo-nos uma compreensão conceitual mais profunda do termo, explicitada pelas possibilidades de aplicação e uso decorrentes dessas categorias.

O *corpus*, composto de dois segmentos pertencentes ao texto de Alvarenga et al. (2011), se justifica pela consideração de que o texto completo faz parte da obra *Interdisciplinaridade em Ciência, Tecnologia & Inovação*, organizada por Philippi Junior e Siva Neto (2011), trabalho resultante do consórcio de várias agências institucionais - dentre as quais, a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior [Capes]. Os textos inclusos em Philippi Junior e Silva Neto (2011) apresentam uma discussão aprofundada das dimensões teóricas e aplicadas do conceito de interdisciplinaridade, contextualizando tais dimensões em relação a avaliação da institucionalização acadêmica, sobretudo a da pós-graduação de natureza interdisciplinar.

Importa referir que, sendo o intuito deste estudo analisar as apropriações discursivas pelos autores do capítulo selecionado (considerados porta-vozes da pesquisa, mas, ao mesmo tempo, envolvidos nos processos de avaliação da própria Capes), e não a articulação dialética de tais apropriações em relação aos textos originais, não se julgou aplicável detalhar sobre os critérios de avaliação dos Programas de Pós-graduação interdisciplinares da Capes, nem dos critérios de seleção dos teóricos apresentados em Apostel, Berger, Briggs e Michaud (1972) – Apostel et al. (1972) - considerando-se apenas a configuração e uso dos atributos teóricos do conceito de interdisciplinaridade na convergência dos contextos abordados por Alvarenga et al. (2011).

Procedeu-se à leitura e identificação de unidades textuais (frases) que abordavam questões relativas ao termo Interdisciplinar, Interdisciplinaridade, Interdisciplina (acolhendo-se a prefixação de Multi, Pluri

ou Trans quando mencionados na intenção de atribuir uma qualidade relacional ao termo raiz Disciplina).

Seguiu-se a contagem do número total de frases obtidas no capítulo 1 (472), e distribuição desse total em divisões temáticas, tais como: abordagem histórico-sociológica (214), definições conceituais (157), perspectivas de pesquisa (107). Dessas divisões, os dois subtópicos que abordavam as definições conceituais foram identificados de maior interesse para o estudo: Texto 1, subtópico apresentado na página 31 até 42 (75); e, Texto 2, subtópico apresentado na página 42 até 52 (82).

A etapa final compreendeu a divisão dessas unidades textuais em enunciados (849) relacionados ao termo Interdisciplinaridade (e termos adjacentes), estas formações textuais (expressões, afirmações e definições) foram consideradas de interesse para aplicação da análise/categorização conceitual por meio do princípio de PMEST, derivando as seguintes facetas e subfacetas (subcategorias): Tipo (como *Personality*); Base (como *Matter*), exigência, limite e origem; Aplicação (como *Energy*), efeito, estímulo e função; Lugar (como *Space*); e Período (como *Time*).

RESULTADOS

A noção inicial utilizada como parâmetro de discussão do tema Interdisciplinaridade, segundo Alvarenga et al. (2011), foi elaborada por Guy Berger (1972, pp. 23-26), e apresentada em Apostel et al. (1972). Seguindo-se a proposição de análise apresentada no presente trabalho, as categorias estabelecidas pelo princípio PMEST (como instrumento de categorização conceitual, traduzido para o Português), na faceta Personalidade (P) Interdisciplina, foram preenchidas na seguinte configuração:

Matéria (M): Foi identificada como de base disciplinar mútua, com formato grupal, composto por especialistas, compreendendo na subcategoria, Exigência, a condição de composição de sua base por áreas de conhecimento com conceitos, termos, métodos e dados próprios (as subcategorias Limite e Origem, não foram preenchidas);

Energia (E): Foi preenchida no quesito Aplicação; na subcategoria Efeito, registrando-se, como consequência de uma Interdisciplina, a comunicação de ideias, a integração de conceitos de ordem epistemológica, terminológica, metodológica (procedimentos, dados de pesquisa, ensino); não sendo registrado preenchimento do quesito Estímulo; sendo o quesito Função caracterizado pela possibilidade de trocas e empréstimos disciplinares.

Lugar (S) e Tempo (T) foram facetas excluídas dessa fase da pesquisa, uma vez que a análise tem como contexto de interesse o momento de apropriação de Alvarenga et al. (2011) como representantes da Capes, financiadora da obra que apresentou o capítulo aqui analisado, parâmetro contextual pelo qual a análise hermenêutica se efetivou: autores-proponentes e agência financiadora.

Obteve-se, como produto desta categorização, dois quadros descritivos dos tipos de relação interdisciplinar apresentadas pelas apropriações de Alvarenga et al. (2011), visando contraposição às enunciações sobre as citações da Capes, explicitadas pelos autores no decurso do capítulo analisado e apresentadas na introdução deste trabalho.

O Quadro 1 apresenta as apropriações das abordagens/proposições sobre o termo Interdisciplinaridade (Texto 1), identificadas em Alvarenga et al. (2011) e categorizadas segundo princípio PMEST (as subcategorias Origem e Efeito foram subtraídas do quadro por não apresentarem dados respectivos).

Quadro 1. Categorização PMEST dos enunciados sobre o conceito de Interdisciplinaridade por Alvarenga et al. (2011), no Texto 1.

Facetas () e subfacetas	Piaget	Jantsch
Tipo (P)	Interdisciplinar	
Base (M)	Hipótese inicial: Fragmentação das ciências; Fronteiras observáveis. Objetos disciplinares. Problemas disciplinares.	Axiomática disciplinar comum conexas. Dois níveis.
Exigência	Ciência interativa. Estruturas dedutíveis. Estruturas hierárquicas. Trocias recíprocas. Enriquecimento mútuo.	Axiomática disciplinar comum. Ponto de vista único. Objetivo em comum
Limite	Regime interdisciplinar. Restrita à organizações estruturais.	Coordenação por nível superior. Sem axiomática única.
Aplicação (E)	Relações estruturais: Múltiplas, Inteligíveis, Dedutíveis e Comparáveis.	
Estímulo	Aprofundamento da investigação	
Função	Colaboração entre disciplinas. Colaboração entre setores heterogêneos.	Introdução de noção de finalidade. Objetivos múltiplos. Coordenação.

No caso de apropriação das proposições de Heckhausen (1972, pp. 83-89) e Boisot (1972, pp. 89-97) por Alvarenga et al. (2011), o Quadro 2 apresenta mais de um tipo de denominação, uma vez que a Faceta (P) comporta a inclusão do termo Interdisciplinaridade por meio de atribuição de qualificações que são ajustadas a partir do emprego deste (como complemento ou ferramenta auxiliar, ou de caráter linear), inclusive pelo campo científico.

No restante das facetas, consideramos que o quadro apresenta a mesma configuração, nos mesmos parâmetros, do Quadro 1, cabendo informar que as subcategorias Limite e Efeito foram subtraídas do Quadro 2 por não apresentarem dados respectivos a essas categorizações.

Quadro 2. Categorização PMEST dos enunciados sobre o conceito de Interdisciplinaridade por Alvarenga et al. (2011), no Texto 2.

Facetas () e subfacetas	Heckhausen		Boisot
Tipo (P)	Interdisciplinar Auxiliar	Interdisciplinar Complementar	Interdisciplinar Linear
Base (M)		Categorias analíticas incomensuráveis por: Falta de leis dialógicas, Falta de teorias dialógicas.	Disciplinarmente relacional. Interdisciplinarmente explicável por meio de inerências disciplinares mútuas.
Exigência		Níveis disciplinares diferenciados. Regras disciplinares restritas.	
Origem		Regiões disciplinares fronteiriças.	
Aplicação (E)			
Estímulo		Criação de níveis de correspondência Teórica. Objeto em construção parcial em direção à completude.	
Função	Trocas e empréstimos de (por meio de): Técnicas de pesquisa disciplinar, Métodos de pesquisa disciplinar.	Imbricação de níveis de integração teórica.	Trocas e empréstimos.

O primeiro aspecto a ser salientado na comparação entre os Quadros 1 e 2 é a variação terminológica entre as facetas dos Tipos (P) de relações estabelecidas pelo Quadro 1 (Interdisciplinar) e pelo Quadro 2 (Interdisciplinar Auxiliar, Complementar, e Linear) que cobrem aplicações da interdisciplinaridade no campo científico por meio de graus de aplicação: no primeiro, pelo uso dos prefixos Multi, Pluri, Inter, e Trans; e, no segundo, pela atribuição de qualificações às relações do tipo Interdisciplinar (Auxiliar, Complementar), como atividade coadjuvante da ciência na apropriação de Heckhausen (1972), e formal (Linear, Restrita e Estrutural) na apropriação de Boisot (1972).

Como aspectos secundários, identificamos o enfoque de traços convergentes (concordância) das definições apresentadas sobre Interdisciplinaridade, sendo tal enfoque considerado mais produtivo para nossa discussão, razão pela qual nos limitamos a esse aspecto nas apropriações de Piaget (1972, pp. 127-139) e Jantsch (1972, pp. 97-121) (Quadro 1) e Heckhausen (1972) e Boisot (1972) (Quadro 2) por Alvarenga et al. (2011).

Quadro 1: identificamos convergência entre denominação de tipos (P), pelo uso do termo Interdisciplinaridade; na Faceta Base (M), subfaceta Exigência, entre a condição de trocas recíprocas/enriquecimento mútuo e a necessidade de ponto de vista único/objetivo comum como requisito essencial; na Faceta Aplicação (E), subfaceta Funções: pela ação de colaboração entre setores heterogêneos e o resultante de objetivos múltiplos pelo ato de coordenação;

Quadro 2: identificamos convergência na faceta Aplicação, subfaceta Função, entre qualidades funcionais de trocas e empréstimos (mais aprofundada em Interdisciplinar Auxiliar de Heckhausen que em Interdisciplinar Linear, de Boisot) e a imbricação de níveis de integração teórica que, por sua vez, resultam também da eficiência nas funções interativas propostas pela Interdisciplinaridade.

CONCLUSÕES

De maneira geral, intui-se que as apropriações de Alvarenga et al. (2011) em torno das relações interdisciplinares expressas tanto no âmbito da pesquisa - pelos teóricos que investigaram o conceito de Interdisciplinaridade – quanto no âmbito do fomento – pela consideração dos critérios definidores do que seja Interdisciplinaridade (e seus objetivos) pela Capes -, apresentam consonância discursiva, cabendo graus de convergência entre o que se institui em ambos os contextos (fomento e pesquisa). Inferindo-se, no entanto, que essa convergência discursiva estabelecida pelos autores, em torno do que a teoria propõe e do que a Capes tem interesse em promover, prejudica a imagem de imparcialidade na avaliação das bases teóricas apresentadas.

Inferese-se que tal consonância, desconsidera graus de complexidade da pesquisa interdisciplinar, uma vez que as abordagens institucionais primam pela obediência à princípios administrativos e normativos que a ciência, nem sempre, corrobora em sua pretensão de reconhecimento de verdades disciplinares estáveis, cabendo aos cientistas, mesmo quando concordantes com as exigências institucionais, apontarem possibilidades de aperfeiçoamento do sistema científico (de pesquisa e de fomento) como um todo.

Os limites com os quais os autores se deparam estão sedimentados pela obediência a um critério pré-estabelecido em suas interpretações do que a Capes propõe como Interdisciplinaridade, ou seja, Alvarenga et al. (2011) corroboram um sentido ao conceito, que, de modo geral, é valorizado como agente promotor dos avanços quantitativos na pesquisa nacional – enfatizado, sobretudo, pelo informe de aumento exponencial de cursos de pós-graduação interdisciplinares: de 46, em 1999, a 283, em 2009.

Em contraponto às críticas acima mencionadas, a consideração de Alvarenga et al. (2011) de que a escolha teórico-metodológica menos “rígida” – passível de adaptação à realidade do problema abordado na pesquisa científica – é a solução mais benéfica ao compromisso da pesquisa interdisciplinar, nos permite confirmar a hipótese de uso pragmático de um dos sentidos do conceito, sem prejuízo dos outros.

Cabe, entretanto, averiguar se tal enunciação - uma vez que os autores se reconhecem alinhados a uma das correntes teóricas (jantschiana) - resiste à confrontação dialética com os discursos originais dos teóricos apresentados em Apostel et al. (1972, pp. 83-139), e da Capes (2009), apresentado no *Documento de Área – 2009*, documentos utilizados como referência por Alvarenga et al. (2011) em suas enunciações sobre o conceito de Interdisciplinaridade.

Julgamos pertinente recomendar, complementarmente, que uma perspectiva de investigação futura seria a de identificar em que nível essa deferência à liberdade de escolha e aplicação da pesquisa interdisciplinar encontra subsídio na própria Capes, e em seus critérios de avaliação, pelo reconhecimento e valorização das adaptações exigidas pela realidade da investigação científica e da própria interdisciplinaridade.

REFERENCES

Alvarenga, A. T., Philippi Junior, A., Sommerman, A., Alvarez, A. M. S. & Fernandes, V. (2011). Histórico, fundamentos filosóficos e teórico-metodológicos da interdisciplinaridade. In: Philippi Junior,

A. & Siva Neto, A. J. *Interdisciplinaridade em Ciência, Tecnologia & Inovação* (Cap. 1, pp. 3-68). Barueri: Manole.

Apostel, L., Berger, G., Briggs, A. & Michaud, G. (1972). *Interdisciplinarity: problems of teaching and research in universities*. Paris (France): Centre for Educational Research and Innovation.

Bauer, M. W. & Gaskell, G. (2008). *Pesquisa qualitativa com texto, imagem e som: um manual prático*. Rio de Janeiro: Editora Vozes.

Bauer, M. W. (2008). Análise de conteúdo clássica: uma revisão. In: Bauer, M. W. & Gaskell, G. *Pesquisa qualitativa com texto, imagem e som: um manual prático*, (Cap. 8, pp. 189-217). Rio de Janeiro: Editora Vozes.

Berger, Guy (1972). Introduction. In: Apostel, L., Berger, G., Briggs, A. & Michaud, G. *Interdisciplinarity: problems of teaching and research in universities*, (Cap. 1, pp. 23-26). Paris (France): Centre for Educational Research and Innovation.

Boisot, M. (1972). Discipline and interdisciplinarity. In: Apostel, L., Berger, G., Briggs, A. & Michaud, G. *Interdisciplinarity: problems of teaching and research in universities*, (Cap. 1, pp. 89-97). Paris (France): Centre for Educational Research and Innovation.

Bourdieu, P. (1989). *O poder simbólico*. Rio de Janeiro: Editora Bertrand do Brasil.

Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior. (2009). *Documento de área 2009*. Retrieved from <https://www.capes.gov.br/images/stories/download/avaliacao/INTER03ago10.pdf>

Dahlberg, I. (1978a). Teoria do conceito. *Ciência da Informação*, 7(2), 101-107. Retrieved from <http://revista.ibict.br/ciinf/article/view/115/115>

Dahlberg, I. (1978b). Fundamentos teórico-conceituais da classificação. *Revista de Biblioteconomia*, 6(1), 1-13.

Dahlberg, I. (1972, setembro). Teoria da classificação, ontem e hoje. *Anais da Conferência Brasileira de Classificação Bibliográfica*, Rio de Janeiro, RJ, Brasil, 12-17. Retrieved from http://www.conexaorio.com/bit/dahlbergteoria/dahlberg_teoriam.htm

Heckhausen, H. (1972). Discipline and interdisciplinarity. In: Apostel, L., Berger, G., Briggs, A. & Michaud, G. *Interdisciplinarity: problems of teaching and research in universities* (Cap. 1, pp. 83-89). Paris: Centre for Educational Research and Innovation., pp. 83-89)

Jantsch, E. (1972). Towards interdisciplinarity and transdisciplinarity in education and innovation. In: Apostel, L., Berger, G., Briggs, A. & Michaud, G. *Interdisciplinarity: problems of teaching and research in universities* (Cap. 1, pp. 97-121). Paris: Centre for Educational Research and Innovation.

May, T. (2004). Pesquisa documental: escavações e evidências. *Pesquisa social: questões, métodos e processos* (Cap. 8, pp. 212-215). Porto Alegre: Artmed.

Oliveira, M. R.; Almeida, J. (2011) Programas de pós-graduação interdisciplinares: contexto, contradições e limites do processo de avaliação Capes. *Revista Brasileira de Pós-Graduação*, 8(15), 37-57.

Philippi Junior, A. & Siva Neto, A. J. (2011). *Interdisciplinaridade em Ciência, Tecnologia & Inovação*. Barueri: Manole.

Piaget, J. (1972). The epistemology of interdisciplinary relationships. In: Apostel, L., Berger, G., Briggs, A. & Michaud, G. *Interdisciplinarity: problems of teaching and research in universities* (Cap. 1, pp. 127-139). Paris: Centre for Educational Research and Innovation.

Pombo, O (2005). Epistemologia da Interdisciplinaridade. *Liinc em Revista*, 1(1), 3-15

.



TEORIAS DA LINGUAGEM NOS ESTUDOS DE INDEXAÇÃO: UMA ANÁLISE A PARTIR DOS PERIÓDICOS *JOURNAL OF DOCUMENTATION* E *KNOWLEDGE ORGANIZATION*

Raquel Luise Pret¹, Rosa Inês de Novais Cordeiro²

¹Professora Assistente do Departamento de Ciência da Informação da Universidade Federal Fluminense (UFF/Brasil), pretraquel@vm.uff.br

²Professora Associada do Departamento de Ciência da Informação da Universidade Federal Fluminense (UFF/Brasil), rosacordeiro@vm.uff.br

RESUMO Este artigo é resultado de um estudo analítico sobre a influência das teorias linguísticas, em especial da Semântica, nas abordagens sobre indexação. A intenção é mostrar como o Realismo e o Mentalismo enquanto paradigmas que influenciam a compreensão da linguagem no campo da Ciência da Informação. No Realismo, a linguagem expressa o mundo real. O sentido é aquilo que descreve a realidade. É possível identificar conceitos, leis, lógicas universais porque as coisas do mundo são criadas a partir da mesma substância. Dessa forma, a vocação da linguagem é declarar, informar sobre elas. Já no Mentalismo, de inspiração aristotélica, o universal é a mente. A faculdade racional subordina a linguagem a comunicar o pensamento. Estes dois paradigmas têm grande influência nas clássicas teorias linguísticas como a Semântica Formal de Frege e a Gramática Gerativa de Chomsky. A Lógica Representacional da Semântica fregeana prioriza as relações formais, ou seja, os aspectos objetivos do significado. A notação válida sentenças a partir de critérios que estabelecem o que é verdadeiro ou falso, eliminando as dúvidas e os mal entendidos da linguagem natural. De forma semelhante, a teoria da Gramática Gerativa de Chomsky compartilha do pressuposto aristotélico que a linguagem é usada como ferramenta para dizer o que pensamos. O aparato combinatório e interpretativo da gramática consiste em um conjunto de princípios e regras, inconscientes, mas que podem ser analisados precisamente porque, de algum modo, eles são análogos a operações matemáticas. Os estudos de indexação publicados pelos *Journal of Documentation* e *Knowledge Organization* apresentam em suas abordagens a influência dos paradigmas Realista e Mentalista na compreensão da linguagem que impacta diretamente na representação temática dos documentos.

PALAVRAS-CHAVE *Indexação; Realismo; Mentalismo; Journal of Documentation; Knowledge Organization*

ABSTRACT This paper is the result of an analysis on the influence of linguistic theories, especially the Semantic field, on indexing approaches. The intention is to show how the Realism and the Mentalism as paradigms influence the understanding about language in Information Science. In Realism, language expresses the real world. Meaning is what describes reality. It is possible to identify universal concepts, laws, logics because the things of the world are created from the same essence. So, the vocation of language is to declare, to inform about them. On the other hand, for The Mentalism of Aristotelian inspiration, the universal is the mind. The rational faculty subordinates language to thought. - both paradigms have great influence on classic Linguistic theories like Frege's -Formal Semantics and Chomsky's Generative Grammar. The Rrepresentational Logic in Frege's Semantics prioritizes formal relations, that is, the objective aspects of meaning. Chomsky's Generative Grammar theory shares the Aristotelian assumption that language is used as a tool to say what we think. The combinatorial and interpretative apparatus of a grammar consists of an unconscious set of principles and rules, that can be precisely analyzed - because, -, they are, somehow,

analogous to mathematical operations. The approaches of the indexing studies published by the *Journal of Documentation* and by the *Knowledge Organization* journal demonstrate the influence of the Realism and the Mentalism paradigms on the understanding of the language, which impacts directly in the subject analysis of records.

KEYWORDS *Indexing; Realism; Mentalism; Journal of Documentation; Knowledge Organization*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Este artigo é produto de um estudo analítico sobre a influência das teorias linguísticas, em especial no campo da Semântica, nas abordagens sobre indexação. A intenção é mostrar como o Realismo e o Mentalismo influenciam os autores da área da indexação a pensar este processo.

De acordo com Silva e Fujita (2004, p. 136), o conceito de indexação surgiu a partir da elaboração de índices num período em que não estava tão demarcada a separação entre representação descritiva e representação temática. Para Lancaster (2003, p. 1), os processos de catalogação descritiva identificam autores, fontes, títulos e outros elementos bibliográficos. Já a indexação é um processo que identifica o assunto de que trata um documento, ou seja, representa o seu conteúdo.

A indexação consiste em uma assembleia de referências para itens de informação. Os itens podem ser o assunto tratado por todos os documentos ou fatos gravados nos documentos, mas dessa massa de referências o pesquisador deve ser capaz de selecionar dessas referências um assunto específico (Vickery, 1950, p. 140, tradução nossa)

A necessidade de estabelecer o assunto (*subject*) ou a tematicidade (*aboutness*) de um documento está diretamente relacionada à questão da linguagem: qual é o seu sentido? Instintivamente, respondemos esta questão a partir de nossa concepção filosófica, mesmo sem perceber quais os princípios que governam a nossa decisão. Percebemos que nas abordagens sobre indexação há dois paradigmas da filosofia da linguagem frequentes: o **realismo**, aquele que identifica parcelas da realidade, e o **mentalismo**, aquele que partilha eventos mentais entre falantes e ouvintes.

No Realismo, a linguagem expressa o mundo real. O sentido é aquilo que descreve a realidade e as suas coisas. Essa díade é associada ao pensamento de Platão sobre as palavras e sua função. Em *Crátilo*, Platão apresenta sua compreensão sobre como o mundo é formado. Para ele, há um mundo ideal onde a essência de todas as coisas reside. Esta essência transcende à experiência humana. É possível identificar conceitos, leis, lógicas universais porque as coisas do mundo são criadas a partir da mesma substância e o que muda é a maneira em que a essência é configurada no mundo sensível onde vivemos. A linguagem funciona como ferramenta para descrever as coisas do mundo.

O Mentalismo, de inspiração aristotélica, partilha do pensamento platônico sobre a essência universal das coisas. Contudo, neste paradigma, a linguagem não diz o real, mas diz o que pensamos sobre o real. Para Aristóteles, não há divisão entre o mundo ideal e o mundo sensível. É possível tocar o universal sobre as coisas porque as almas humanas vêm da mesma substância. O universal é a mente. Então, a função da linguagem é dizer o que a mente pensa. A faculdade racional ocupa o lugar central no desenvolvimento do pensamento e a linguagem um lugar secundário, subordinado à mente. Estes dois

paradigmas - Realismo e Mentalismo - têm grande influência nas clássicas teorias linguísticas como a Semântica Formal de Frege e a Gramática Gerativa de Chomsky.

Frege, filósofo matemático, herdeiro da tradição aristotélica, procurou resolver problemas da Lógica criando um sistema de representação formal. Seu trabalho objetiva o controle do significado nas equações matemáticas. A partir desta perspectiva, ele escreve *Sentido e Referência* (1976), um livro sobre a necessidade de controlar os sentidos na linguagem, removendo suas imperfeições. A Lógica Representacional da Semântica de Frege prioriza as relações formais, ou seja, os aspectos objetivos da linguagem. Os sentidos são modos de apresentação das referências, a existência de objetos referenciais independem de seu observador. Em Frege (1976), as relações entre sentenças e seus elementos constitutivos poderiam ser melhor compreendidas por uma linguagem lógica. A notação valida sentenças a partir de critérios que estabelecem o que é verdadeiro ou falso, eliminando as dúvidas e os mal entendidos da linguagem natural. "É a busca pela verdade, qualquer que ela seja, que dirige o sentido à referência" (Frege, 1976, p. 69). O sentido é objetivo porque "a humanidade possui um tesouro comum de pensamentos" (Frege, 1976, p. 65). Nesta perspectiva, é importante refrear a "euforia da linguagem" para que os sistemas de representação possam ser eficientemente comunicativos. A padronização sempre opera a partir de relações binomiais entre palavras, conceitos, termos, sentenças e expressões como: certo e errado, verdadeiro e falso, válido e inválido.

Assim como na teoria da Semântica Formal de Frege, a teoria da Gramática Gerativa de Chomsky compartilha do pressuposto aristotélico que a linguagem é usada como ferramenta para dizer o que pensamos. Segundo Chomsky (1968), o aparato combinatório e interpretativo da gramática consiste em um conjunto de princípios e regras, inconscientes, mas que podem ser analisados precisamente porque, de algum modo, eles são análogos a operações matemáticas. De acordo com a teoria da Gramática Gerativa, há uma gramática comum a todas as línguas, a mente. A Gramática Gerativa fornece um esquema geral: regras e princípios universais governam o comportamento da linguagem e ajuda a gerá-la. "A diferença das línguas está no léxico" (Chomsky, 1968, p.46). Esta Gramática Universal, segundo Chomsky, é um componente biológico humano. Se "a língua é a manifestação mais ampla da nossa vida mental" (Chierchia, 2003, p. 24), é necessário compreender os padrões mentais e reproduzi-los na linguagem. É possível fazer a correspondência entre palavras de qualquer língua porque possuímos um aparelho cognitivo parecido, a formulação dos pensamentos é semelhante entre os indivíduos. Chomsky afirma que as estruturas linguísticas variam pouco de um país para outro. A maior diversidade está nas possibilidades de combinação entre as estruturas. Nesta concepção, nós possuímos um "algoritmo inconsciente" capaz de combinar estruturas finitas da linguagem, criando sequências infinitas de palavras, sentenças, expressões e textos. A maneira com que amalgamamos os significados das palavras depende exclusivamente da maneira com que as combinamos sintaticamente.

Os estudos sobre indexação foram influenciados por esta forma de pensar a linguagem: palavras declaram conceitos universais sobre o mundo. Os elementos, características e atributos transformarão os conceitos em bem sucedidos se suas combinações chegarem próximo à descrição das coisas e do pensamento. Há uma objetividade essencialista no mundo captada pela mente e representada pela linguagem. Nesta visão instrumentalista, a principal função da linguagem é informar sobre as coisas.

METODOLOGIA

Desse modo, procuramos mapear os estudos de indexação em dois importantes periódicos, reconhecidos internacionalmente no campo da Ciência da Informação - o *Journal of Documentation* e o *Knowledge Organization*.

O *Journal of Documentation* é o periódico mais antigo na área da Ciência da Informação, publicado desde 1945, consolidou-se por sua frequência em divulgar teorias, filosofias, conceitos, modelos sobre documentação, organização do conhecimento, indexação, classificação e outros estudos no mesmo campo. Este periódico possui 73 volumes e 329 números. Já o *Knowledge Organization*, periódico oficial da International Society for Knowledge Organization, publica pesquisas sobre teorias, fundamentos filosóficos do conhecimento, operações práticas associadas à indexação e à classificação, além de problemas de terminologia. Fundado em 1976, possui 44 volumes e mais de 161 números.

Assim, delimitamos como escopo para a análise dessa pesquisa às publicações referentes aos últimos 10 anos nos dois periódicos, entre 2006 e 2016. No *Journal of Documentation* (JD) analisamos todos os números a partir do volume 62, totalizando 63 números, nos quais encontramos em seus títulos e resumos 323 artigos abordando a temática da indexação (*indexing*), 22 sobre análise temática (*subject analysis*) e 63 sobre estudos de linguagem (*language*). Números consideráveis que apontam a preocupação do campo com questões referentes à indexação.

Já os artigos do *Knowledge Organization* (KO) analisados possuem diferentes características relacionadas às teorias linguísticas. O escopo desta análise também utilizou como marco os últimos dez anos de publicações (2006-2016), ou seja, todos os artigos a partir do volume 33. Portanto, analisamos 53 publicações, dentre as quais encontramos entre títulos e resumos 53 artigos sobre indexação (*indexing*), 5 artigos sobre análise temática (*subject analysis*) e 2 artigos sobre estudos de linguagem (*language*). Entendemos que a grande diferença nos números de artigos referentes a indexação e estudos de linguagem, se compararmos aos números do *Journal of Documentation*, está associada à própria questão filosófica que, de certa forma, serve de diretriz ao KO.

ANÁLISE DOS RESULTADOS

Tradicionalmente, o *Journal of Documentation* publicou importantes estudos sobre indexação como os de Vickery, Farradane, Coates, Langridge, Faithorne e Foskett entre os anos de 1950 e 1970. O Classification Research Group, formado na Inglaterra em 1952 por A. J. Wells e B. C. Vickery a partir das discussões acerca da classificação levantadas na *Royal Society Scientific Information Conference* de 1948, teve seus primeiros boletins publicados neste periódico. A proposta inicial do CRG era reunir bibliotecários interessados em discutir os sistemas de classificação de suas instituições, sobretudo a aplicabilidade de princípios classificatórios. O grupo sentia a necessidade de compartilhar suas experiências na organização sistemática de documentos por assunto, divulgando práticas e aplicações teóricas bem sucedidas. Os debates registrados em seus boletins foram importantes referenciais para os estudos de indexação. Essas teorias sobre indexação possuem em comum o fato de pensar formas de sistematização da linguagem na entrada dos documentos em um sistema de recuperação da informação. Neste período, os computadores estavam sendo inseridos nos sistemas de recuperação da informação (SRIs). Portanto, os estudos da indexação visavam buscar maneiras de padronização da linguagem, realizando a sua tradução para uma linguagem artificial que fosse compreensível para os serviços de

informação automatizados, facilitando assim a recuperação dos documentos de tais bases de dados. Essa busca aproximou os estudos acerca da indexação dos estudos de linguagem.

A construção de linguagens artificiais, controladas e padronizadas possui forte influência da lógica encontrada na semântica formal de inspiração fregeana. A premissa semelhante à lógica quantificadora de predicados e sentidos de Frege é a do controle da linguagem pela indexação a fim de realizar a correspondência entre assunto requerido e conteúdo do documento. Assim, os primeiros estudos publicados no *Journal of Documentation* associando a indexação aos serviços de informação como abordados por Vickery (1953), Saracevic & Rees (1967) e Foskett (1975) vão se aproximar dos estudos linguísticos para realizar suas proposições. Esses estudos possuem em comum a premissa de entender a indexação como uma operação que deve reconhecer os assuntos (*subjects*) mais significantes de um documento para auxiliar a recuperação da informação. Mas como precisar o que é mais significativo? Tais teorias apontam para o uso dos princípios da terminologia que procuram controlar as significações das palavras (Cabré, 1993). A proposta é estabelecer uma tradução da linguagem natural a um sistema lógico tentando criar um vocabulário em que cada palavra tenha apenas um significado, esses estudos procuram estabelecer leis para controlar as variações na indexação, expressando a relação entre o tamanho do vocabulário e a especificidade com que um conceito pode ser descrito (Gardin, 1973, p. 143).

A linguagem representacional deste modelo matemático-terminológico de indexação aproxima-se com a semântica formal proposta por Frege (1976), marcada pela influência do paradigma realista da linguagem. Tanto nessa teoria da indexação, quanto na semântica formal de Frege (1976), a linguagem é compreendida como representação da realidade. O sentido do documento estaria no conteúdo expresso no texto.

A influência deste pensamento pode ser apreendida nos artigos produzidos nos últimos anos acerca de indexação automática. Nos 36 artigos encontrados sobre indexação automática como o trabalho de Tudhope *et al* (2006), *Query expansion via conceptual distance in thesaurus indexed collections*, o estudo de Zhang e Zeng (2014), *A new similarity measure for subject hierarchical structures; Classifications and concepts: towards an elementary theory of knowledge interaction*, de Smiraglia e Heuvel (2013); *Untangling search task complexity and difficulty in the context of interactive information retrieval studies*, de Wildemuth, Freund, Toms (2014) são alguns exemplos da influência do paradigma realista na abordagem acerca da indexação e da recuperação da informação. Esses artigos apresentam uma concepção de linguagem semelhante à semântica formal. Tais estudos apresentam um princípio em comum que é procurar estabelecer possibilidades - seja na entrada dos documentos, seja na sua saída dos sistemas - para que o computador realize a melhor conexão entre os documentos armazenados nas bases de dados e a pergunta do pesquisador. Os estudos enfatizam o controle terminológico como uma condição primordial na recuperação da informação. Outrossim, a construção de modelizações, leis, padronizações, para o controle de termos, conceitos e palavras pela indexação automática filia-se à semântica formal ao procurar *traduzir* a linguagem para uma representação lógica, tentando evitar o que Frege chamava de *imperfeições* da linguagem. Problemas também identificados na indexação como a sinonímia, a polissemia e os sentidos fora do contexto eram questões que a semântica formal procurava eliminar a partir do controle da linguagem pela notação lógica e o cálculo de predicados.

Já os estudos mais recentes sobre indexação automática que propõem a utilização de métodos a partir da padronização por sintagmas como o artigo de Ibekwe-Sanjuan (2006), *Constructing and maintaining*

knowledge organization tools, de Rafferty (2011), *Knowledge Representation in the Social Semantic Web* são marcados pela influência da gramática gerativista chomskyana. Segundo Chomsky, a linguagem é uma faculdade mental, um equipamento inato e universal, ilimitado e criativo (é um repertório de unidades e princípios de combinação que oferecem ilimitadas possibilidades de expressão. Se "a linguagem é a manifestação mais ampla de nossa vida mental" (Chierchia, 2003, p. 24), é necessário que compreendamos as estruturas gramaticais procurando padronizá-las, pois a forma de pensar é semelhante.

Sendo assim, a tradução seria a simples operação de substituir a palavra em determinada língua por a sua correspondente, o mesmo se daria para as sentenças e até mesmo para os textos. O sentido não se altera porque ele é estabelecido pela mente e não pela linguagem. De acordo com Chierchia (2003, p. 38), possuímos um "algoritmo inconsciente" capaz de combinar as regras e padrões, conjunto finito e matricial, de maneiras infinitas que se desdobram em variadas sentenças, textos e línguas.

A Gramática Gerativa tem sido amplamente utilizada como suporte teórico dos sistemas de indexação automática. Alguns métodos como a indexação por extração automática sintática, a indexação por atribuição semântica, a identificação de palavras *in full context* são exemplos de operações que buscam estabelecer padrões de conceitos, de relações hierárquicas entre as palavras, termos e sintagmas, prevendo inúmeras possibilidades de combinação entre elas.

Esta abordagem da indexação a partir da gramática gerativista e, por conseguinte, do paradigma mentalista, facilitou a interface entre o indexador e o computador, possibilitando o aparecimento de diversas bases de dados com graus elevados de precisão e revocação na recuperação dos documentos.

O *Knowledge Organization* fundamenta-se no conceito de organização e representação do conhecimento, diferentemente do *Journal of Documentation* que privilegia o tratamento temático da informação e o comportamento de uso. Dessa forma, operações como a indexação estão inseridas nas discussões, teorias e problematizações da própria área da Organização do Conhecimento.

Dahlberg, ao analisar a Organização do Conhecimento enquanto campo científico em 1993, quando da mudança de nome do periódico de *International Classification* para *Knowledge Organization*, estabelece que a organização do conhecimento é um domínio da biblioteconomia e da ciência da informação comprometido em organizar, ordenar e dispor a informação (Dahlberg, 1993, p. 211). Este pensamento está em consonância com a sua teoria do conceito que estabelece as formas em que os indivíduos organizam seu modo de pensar. De acordo com a autora, os conceitos "identificam as características, os atributos das coisas designadas" (Dahlberg, 1978, p. 102). Segundo esta teoria, o processo de análise dos conceitos somente é possível "coletando-se os enunciados verdadeiros que sobre determinado objeto se podem formular" (Dahlberg, 1978, p. 102). Assim como nos pressupostos platônicos, Dahlberg indica que nos conceitos há elementos que constituem enunciados verdadeiros sobre as coisas, passíveis de serem identificados e classificados de acordo com suas características naturais. Nesta concepção, existe uma autonomia simbólica dos enunciados que compõem os textos. Os elementos, as características e os atributos vão compor os conceitos, a tessitura bem sucedida dessas combinações é que vão indicar se a representação do real será bem ou mal sucedida. A eficácia dos conceitos está relacionada diretamente à fidelidade de sua descrição do real. Encontramos, portanto, a influência do pensamento realista de tradição platônica na definição de *conceito* estabelecida por Dahlberg.

Percebemos que a tendência majoritária dos artigos publicados nos últimos dez anos no *Knowledge Organization* é de uma abordagem semelhante acerca do processo de indexação e organização do conhecimento. *Semantic relations in knowledge organization systems*, de Brascher (2014), faz alusão ao trabalho de Dahlberg sobre a teoria do conceito na construção da categoria *organização do conhecimento*. O artigo enfatiza a necessidade do mapeamento das estruturas e funções semânticas e sintáticas dos KOS a fim de facilitar a identificação e, por conseguinte, a representação dos conceitos. Esta operação visa a padronização da linguagem artificial, facilitando a recuperação da informação. Os são compreendidos como unidades universais e representativas do pensamento do indivíduo, tal como na gramática gerativista chomskiana. Nesta perspectiva, a linguagem representa sistemas de conceitos elaborados e organizados pelo pensamento, produto da cognição característica comum a todos os indivíduos.

Portanto, a representação do conhecimento é construída a partir de conceitos articulados e relacionados entre si, pois ela é produto da mente. A língua apenas expressa o conteúdo mental cuja estrutura sintática e semântica reproduzem o modo de pensar dos indivíduos em seus contextos sociais, culturais, geográficos e históricos. De maneira semelhante ao gerativismo chomskyano, percebemos nesta abordagem a influência dos paradigmas realista e mentalista, por tratar os conceitos como entidades abstratas que designam as coisas no mundo e por considerar a linguagem como expressão do pensamento influenciado por seu meio. Encontramos uma abordagem semelhante acerca da indexação no artigos *Using a semantic analysis tool to generate subject access points*, de Zeng *et al* (2014); *Method for selecting specialized terms from a general language*, de Anguiano Peña e Naumis Peña (2015); *A survey of the coverage and methodologics of schemes and vocabularies use to describe information resources*, de Hider (2015); além do artigo *Terminology web services*, de Binding e Tudhope (2010).

As publicações analisadas vão procurar tratar de questões voltadas à organização e representação do conhecimento a partir da compreensão de que existem conceitos universais que expressam o pensamento tal como vimos em Dahlberg (1978). Portanto, os artigos vão procurar refletir sobre formas de classificar, enquadrar, estabelecer hierarquias e nivelamentos do conhecimento. Há uma predominância dos estudos sobre ontologias - encontramos 56 artigos abordando esta temática nos últimos 10 anos dentre eles *Formal Ontology and the Foundation of Knowledge Organization*, de Herre (2013), *Organizing Acts and Objects*, de Budd (2014) e *Boundary Objects*, de Marchese e Smiraglia (2013). Ademais, as linguagens documentárias e taxonomias também destacam-se como temas de muitos artigos no KO. São 30 artigos abordando a questão da taxonomia na organização do conhecimento como *Towards a Taxonomy of KOS*, de Souza, Tudhope e Almeida (2012), *Concepts and Terms in the Faceted Classification*, de Broughton (2010), *Naming and Reframing*, de Gross (2015). Todos os artigos citados possuem em comum a filosofia de que o conhecimento é uma entidade autônoma com estruturas passíveis de serem classificadas, codificadas e organizadas de maneira a promover o seu acesso. Esses métodos possuem uma grande proximidade do paradigma mentalista. Os sistemas de organização do conhecimento (KOS), assim como a mente humana, possuem padrões limitados, inclusive de estruturas linguísticas tal como os sintagmas, no entanto podem ser utilizados de infinitas maneiras, de acordo com as combinações feitas de forma criativa. Os algoritmos tão utilizados na Matemática e na Ciência da Computação, e aplicados à recuperação da informação, são exemplos dessas possibilidades. A sequência finita de regras, instruções bem definidas e não ambíguas, cada uma das quais devendo ser executadas mecânica ou eletronicamente em um intervalo de tempo finito e com uma quantidade de esforço finita que podem gerar infinitas combinações. Assim como a mente é capaz de realizar conexões entre os pensamentos, os computadores por sua programações são

capazes de realizar, em muitos casos, a recuperação de documentos úteis ao usuário, eliminando o ruído e o silêncio pelo controle e padronização da linguagem natural.

CONSIDERAÇÕES FINAIS

Aproximamo-nos de nosso ponto de chegada, que talvez também seja outro ponto de partida, no sentido de suscitar questionamentos que não se encerram aqui. Nosso percurso procurou refletir sobre como o realismo e o mentalismo, enquanto filosofias da linguagem, que influenciam os estudos acerca da indexação. Escolhemos como campo empírico para tal análise os periódicos *Journal of Documentation* e o *Knowledge Organization*, mediante exame dos títulos e resumos dos artigos, que são fragmentos dos textos na íntegra, é sabemos dessa delimitação na observação dos artigos na pesquisa. Com o objetivo de refinar e apurar a pesquisa, procuramos delimitar nossa investigação usando como marco temporal as publicações dos últimos dez anos de cada jornal, analisando os artigos que problematizaram o processo da indexação. Dessa forma, procuramos compreender como complexas teorias linguísticas tal como o gerativismo chomskyano e a semântica formal de inspiração fregeana encontram-se presentes nas abordagens dos artigos analisados, sobretudo aqueles voltados à indexação e à organização do conhecimento. Percebemos o quanto a tradição da filosofia socrática influencia até hoje a nossa forma de compreender a linguagem. A concepção da linguagem como um instrumento comunicacional, de base universal, cuja função básica é designar os nomes das coisas no mundo é algo presente na forma em que abordamos a indexação dos documentos em nossa contemporaneidade.

Compreendemos que o pensamento aristotélico da língua como sistema de representação da vida mental influenciou inclusive a construção de diversos sistemas de recuperação da informação.

Destarte, desfecha-se este artigo, ressaltando-se que este foi apenas um esforço experimental, resultante de uma pesquisa em andamento para a conclusão de uma tese de doutoramento, que procurou mapear como algumas filosofias da linguagem influenciam as abordagens e estudos sobre indexação e organização do conhecimento. Nossa intenção foi apresentar algumas interações entre elas, de forma panorâmica, procurando não tratá-las como blocos teóricos indivisíveis.

REFERÊNCIAS BIBLIOGRÁFICAS

Anguiano Peña, G.; Naumis Peña, C. (2015). Method for selecting specialized terms from a general language corpus. *Knowledge Organization*, 42(3), 164-175.

Aristotle. (1990). *Metaphysics*. Madrid: Editorial Gredos.

Binding, C.; Tudhope, D. (2010). Terminology web services. *Knowledge Organization*, 37(4), 287-298.

Borko, H. (1977). Toward a theory of indexing. *Information Processing & Management*, 13 May: 355-365.

Brascher, M. (2014). Semantic relations in knowledge organization systems. *Knowledge Organization*, 41(2), 175-180.

Broughton, V. (2010). Concepts and terms in the faceted classification: the case of UDC. *Knowledge Organization*, 37(4), 270-279.

Budd, J. Organizing acts and objects: metaphysical foundations. *Knowledge Organization*. 41(6), 419-428.

Cabré, M. T. (1993). *La terminología: teoría, metodología y aplicaciones*. Barcelona: Ed. Antárdida/Empúries.

Chierchia, G. (2003). *Semântica*. Campinas: Ed. Unicamp.

Chomsky, N. (1968). *Language and mind*. New York: Harcourt, Brace & World, Inc.

Dahlberg (1993). Knowledge organization: its scope and possibilities. *Knowledge Organization*, 20(4), 211-222.

_____ (1978). Teoria do conceito. *Ciência da Informação*, 7(2), 101-107.

Farrow, J. (1995). All in the mind. *The indexer*, 19(4) Outubro: 243-247.

Foskett, D. J. (1975). *Classification and indexing in the social sciences*. London: LexisNexis Butterworths Co.

Frege, G. (1976). *Sense and Reference*. São Paulo: Cultrix/USP.

Gardin, J. C. (1973). Document analysis and linguistic theory. *Journal of Documentation*, 29(2), 137-168.

Plato (2008). *Republic*. Project Gutenberg.

_____ (2011). *Cratylus*. Cambridge: Cambridge University Press.

Farrow, J.F (1991). A cognitive process model of document indexing, *Journal of documentation*, 47(2), 149-166.

Foskett, D.J. (1970). Classification and indexing in the social sciences. *Journal of Documentation*, 22(3), 90-101.

Gross, T. (2015). Naming and reframing: a taxonomy of attacks on knowledge organization. *Journal of Documentation*, 42(5), 263-268.

Harviainen, J. T.; Hamari, J. (2015). Seek, share, or withhold: information trading in MMORPGs, *Journal of Documentation*, 71(6), 1119-1134.

Hemmig, W. S. (2008). The information-seeking behavior of visual artists: a literature review. *Journal of Documentation*, 64(3), 343-362.

Herre, H. (2013). Formal ontology and the foundation of knowledge organization. *Knowledge Organization*, 40(5), 332-339.

Hider, P. (2015). A Survey of the coverage and methodologies of schemas and vocabularies used to describe information resources. *Knowledge Organization*, 42(3), 154-163.

- Hjorland, Birger (2017). Subject (of documents). *Knowledge Organization*, 44(1), 55-64.
- Howarth, L. C. (2010). Clare Beghtol: exploring new approaches to the organization of knowledge. *Knowledge Organization*, 37(2), 95-100.
- Ibekwe-Sanjuan, F (2006). Constructing and maintaining knowledge organization tools: a symbolic approach. *Journal of Documentation*, 62(2).
- Kwaśnik, B. H. (2010). Semantic warrant: a pivotal concept for our field. *Knowledge Organization*, 37(2), 106-110.
- Marchese, C.; Smiraglia, R. P. (2013). Boundary objects: CWA, and HR firm, and emergent vocabulary. *Knowledge Organization*, 40(4), 254-259.
- Martínez-Ávila, D.; San Segundo, R. (2013). Reader-interest classification: concept and terminology historical overview. *Knowledge Organization*. 40(2), 102-114.
- Rafferty, P. (2011). Knowledge representation in the social semantic web. *Journal of Documentation*, 67(5), 896-899.
- Robson, A.; Robinson, L (2015). The information seeking and communication model: a study of its practical application in healthcare. *Journal of Documentation*, 71(5), 1043-1069.
- Sabbar, C.; Xie, I (2016). Language in the information-seeking context: A study of US scholars using non-English sources. *Journal of Documentation*, 72(1). 103-126.
- Satja, M. P.; Madalli, D. M.; Dutta, B. (2014). Modes of growth of subjects. *Knowledge Organization*, 41(3), 195-204.
- Saracevic, T.; Rees, A. M. (1967). Towards the identification and control of variables in information retrieval experimentation. *Journal of Documentation*, 23(1), 7-19.
- Smiraglia, R. P.; Heuvel, C. Classifications and concepts: towards an elementary theory of knowledge interaction. *Journal of Documentation*, 69(3), 360-383.
- Souza, R. R.; Tudhope, D.; Almeida, M. B. (2012). Towards a Taxonomy of KOS: Dimensions for Classifying Knowledge Organization Systems. *Knowledge Organization*, 39(3), 179-192.
- Tudhope, D.; Binding, C.; Blocks, D.; Cunliffe, D (2006). Query expansion via conceptual distance in thesaurus indexed collections, *Journal of Documentation*, 62(4), 509-533.
- Vállez, M.; Pedraza-Jiménez, R.; Codina, L.; Blanco, S.; Rovira, C (2015). Updating controlled vocabularies by analysing query logs. *Journal of Documentation*, 39(7), 870-884.
- Vickery, B. C. (1953). Systematic subject indexing. *Journal of Documentation*, 9(1), 48-57.
- Wildemuth, B.; Freund, L.; Toms, E. G (2014). Untangling search task complexity and difficulty in the context of interactive information retrieval studies. *Journal of Documentation*, 70(6), 1118-1140.
- Zeng, M. L., Gracy, K. F., Žumer, M. (2014). Using a semantic analysis tool to generate subject access points: a study using Panofsky's theory and two research samples. *Knowledge Organization*. 41(6), 440-451.

Zhang, J.; Zeng, M. L (2014). A new similarity measure for subject hierarchical structures. *Journal of Documentation*, 70(3), 364-391.



A EXPANSÃO DO DOMÍNIO DO ARQUIVO: O CASO DO CSAC – CENTRO STUDI E ARCHIVIO DELLA COMUNICAZIONE.

Giulia Crippa

*Faculdade de Filosofia, Ciências e Letras da Universidade de São Paulo - Brasil,
0000-0002-6711-3144, giuliac@ffclrp.usp.br*

RESUMO Propomos, aqui, uma reflexão sobre novos modelos de arquivo que se desenvolveram ao longo das últimas décadas. Acreditamos ser necessária uma discussão sobre uma tipologia específica, a dos arquivos ligados à arte e ao projeto de design e arquitetura. **Objetivos:** Trata-se de discutir projetos e experiências que ilustram algumas questões que nos parecem relacionadas ao fato do termo arquivo ter expandido seu uso, para incluir situações, lugares e coleções de documentos que, em sua constituição, não podem ser inseridos nas definições tradicionais. Acreditamos ser necessário entender que tipologias de arquivos estão se desenvolvendo no campo da arte. **Metodologia:** Procuraremos identificar o papel desenvolvido pela reflexão filosófica na ampliação do sentido do que chamamos arquivo. Nos parece, com efeito, que o termo arquivo passe, hoje a abranger um leque semântico inimaginável até algumas décadas atrás. O uso do termo arquivo inclui inúmeras “memórias” que, se de um lado ecoam algo relativo à ideia principal de arquivo, por outro descreve algo que extrapola aquilo que, tradicionalmente, a ele pertencia. Queremos, assim, indagar os mecanismos que provocaram essa expansão do termo. O trabalho se desenvolve, em um primeiro momento, através da recuperação bibliográfica das definições de arquivo tradicional e seu uso. Nesse sentido, procura-se marcar os elementos essenciais que constituem seu conceito, os materiais e práticas como desenvolvidas no tempo. Em um segundo momento, desenvolveremos um percurso em busca de uma síntese de algumas características da arte contemporânea, em suas transformações materiais e conceituais, estabelecendo as relações com as necessidades de ordenar seus materiais. **Conclusões:** Observaremos, através do estudo do caso do CSAC – Centro Studi e Archivio della Comunicazione da Universidade de Parma - as possíveis modalidades de formação e desenvolvimento desses arquivos que chamamos “híbridos”, procurando entender suas relações com a arte e com sua exposição.

PALAVRAS-CHAVE *Arquivos, Arquivos de Arte, Arte Contemporânea, CSAC.*

ABSTRACT We propose a reflection on new archival models that have been developing over the last decades. We believe that a discussion about a specific typology, that of archives related to art, design and architecture design, is necessary. **Objectives:** To discuss projects and experiences that illustrate issues related to the fact that the term archive expanded its use to include situations, places and collections of documents that, in their constitution, cannot be included in a more traditional definition of archive. We believe it is necessary to understand which typologies of archives are developing in the field of art. **Methodology:** We will try to identify the role of philosophical reflection in the expansion of the meaning of what we call archive. Actually, it seems to us that the term archive has been reaching an unimaginable semantic range up to a few decades ago. The use of the term archive includes innumerable "memories" which, if on one side it echoes something relative to the main idea of archive, on the other it describes something that extrapolates what traditionally belonged to it. We want, therefore, to investigate the mechanisms that led to this expansion of the term. The work is developed, initially, through the bibliographic retrieval of traditional archive definitions and their use.

In this sense, we look forward to mark the essential elements that constitute its concept, the materials and practices as developed along history. In a second moment, we will develop a path in search of a synthesis of some characteristics of the contemporary art, in its material and conceptual transformations, establishing the relations with the needs of ordering its materials. **Conclusions:** Through the study of the case of CSAC - Centro Studi e Archivio della Comunicazione of the University of Parma - we will observe the possible modalities of formation and development of these archives that we call "hybrids", trying to understand their relations with art and its exhibition.

KEY-WORDS *Archives, Art Archives, Contemporary Art, CSAC.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Neste estudo queremos oferecer um percurso que permita compreender os rumos do desenvolvimento da organização do conhecimento engajada na estruturação da memória cultural no âmbito da Arte Contemporânea.

Nossa proposta centraliza dois modelos de memória, um ligado ao museu e outro ao arquivo. No primeiro caso, observam-se as funções de uma instituição que propõe uma memória ligada ao conceito de monumento.

Por contraste, o arquivo desponta como novo sistema de organização e de comunicação da produção estética, em uma relação diferente daquela, já conhecida, de lugar das práticas administrativas, na medida em que uma série de mudanças investem as linguagens, os materiais e os modos de produção da arte, perante os quais o sistema do museu não se expressa de maneira eficaz na realização de sua finalidade de garante da memória. Observando os fenômenos artísticos das últimas décadas, podemos ver como, face os novos produtos artísticos, o museu se encontra, muitas vezes, na impossibilidade de incorporar as obras em suas coleções ou na reserva técnica, na medida em que se trata de obras desmaterializadas, efêmeras, registradas em suportes híbridos. Não é um acaso, assim, que grande ênfase tenha sido dada às atividades expositivas, mais que às de preservação, e que estas últimas se tornaram um problema cada vez mais próximo às discussões da documentação.

Propomos um estudo sobre novos modelos de arquivos que se desenvolveram nas últimas décadas, principalmente no campo da Arte. Através de um estudo de caso, o CSAC – Centro Studi e Archivio della Comunicazione, da Universidade de Parma -, analisando seu desenvolvimento. Observaremos as tensões que ele cria com as definições de arquivo tradicional, bem como a nova dialética que estabelece entre a ideia de usuário de arquivo e público, quando se transforma em lugar de exposição da memória cultural que organiza, através da sobreposição constante entre sua declinação de documento e de monumento, entre serialidade e possibilidades expositivas e estéticas. Busca-se:

- 1) Destacar as definições “clássicas” de arquivo, ligadas à concepção administrativa de uma entidade produtora. Recorreremos principalmente, aqui, às diretrizes da escola italiana do pensamento arquivístico ligadas à obra de Cencetti (1970), que delineou a configuração dos arquivos na Itália.
- 2) Discutir a formação de arquivos “híbridos”, coleções e lugares que surgem no entrelaçamento entre mudanças na concepção de arte, principalmente através de uma

panorâmica que abrange as mudanças no conceito de documento ao longo do século XX no campo da documentação (Otlet, Briet), no campo da História (Le Goff) e no campo das discussões estéticas (Benjamin, Foster).

- 3) Estudar o CSAC como lugar pioneiro na proposição de uma fórmula de arquivo ligada à arte e ao projeto de arquitetura e design.

A discussão, que atravessa pontos teóricos, atinge suas questões centrais no estudo de caso, pois é nele que se encontram os pontos principais da discussão colocados em prática, em ação. Através da análise de sua formação, a partir de 1968, e das diretrizes que norteiam o desenvolvimento de suas coleções, será possível observar qual sentido se imprime à expansão do uso do termo arquivo.

METODOLOGIA

Basicamente, são duas abordagens concomitantes de pesquisa que se alimentam reciprocamente:

- 1) Revisão bibliográfica da produção acadêmica na área de Arquivologia, Ciências da Informação e História da Arte, focada nos grandes temas expressos nos conceitos-chave propostos de Arquivo, Arte Contemporânea, Políticas de Memória.
- 2) Estudo de caso. A opção pelos estudos de caso fundamenta-se na longa tradição que este tipo de pesquisa possui no âmbito das Ciências Sociais. O estudo de caso típico geralmente envolve uma organização ou comunidade, e o pesquisador comumente faz uso do método de observação participante (em qualquer de suas variações), juntamente com outros métodos mais estruturados, como entrevistas, uso de questionários e, mais recentemente, a análise de redes sociais (LEMIEUX e OUIOMET, 2008). Como observa Howard Becker (1999), o estudo de caso costuma visar um duplo objetivo: chegar a uma compreensão abrangente do grupo em estudo e, ao mesmo tempo, estabelecer desenvolvimentos teóricos mais abrangentes sobre regularidades do processo e estrutura social.

A pesquisa envolve três abordagens concomitantes que se alimentam reciprocamente:

- A Um eixo “vertical”, de revisão bibliográfica da constituição das estruturas diretrizes que norteiam a identificação das práticas no âmbito dos arquivos tradicionais;
- B Um eixo “transversal”, dedicado à reflexão sobre a arte contemporânea enquanto fenômeno que se entrelaça às mudanças no uso do termo arquivo;
- C Um eixo “horizontal”, envolvendo o estudo de caso, o CSAC como exemplo “diferenciado” de proposta em termos de arquivo.

O foco de análise percorrerá o espaço da intersecção dos eixos. Procuraremos, assim, acompanhar a implementação e o desenvolvimento do CSAC nos diversos níveis das políticas elaboradas em sua existência.

RESULTADOS

1 – ARQUIVO

Costumamos definir arquivo com memória ordenada principalmente escrita. O arquivo sempre precisa de um sujeito produtor e, como memória registrada de alguém, seu pressuposto fundamental é a própria escrita, enquanto suas finalidades, praticas e funcionais, são operacionais. O arquivo não trata da transmissão das memórias pessoais de quem escreve, tanto que a formalização dos conceitos tende a se tornar extrema. A razão administrativa é tão forte que adquire valor jurídico. O lado formal do documento é prioritário em relação à sua função, e no arquivo escrito tradicional deposita-se uma memória gerada por uma entidade com finalidades práticas, administrativas e burocráticas

Os valores do documento e do arquivo tradicional são:

- 1) Práticos e administrativos,
- 2) Jurídicos – e no arquivo cartáceo isso nunca se perde,
- 3) Culturais.

O terceiro ponto representa principal problema na discussão sobre arquivos, pois requer avaliações ligadas a momentos diversos em que o documento foi produzido e nosso tempo.

Quem é a entidade produtora de um arquivo? É qualquer sujeito, do indivíduo ao organismo social mais complexo que podemos imaginar, que possui uma atividade. A memória do arquivo serve ao desenvolvimento de suas atividades. O que interessa são as finalidades da entidade produtora, seus recursos e as formas adotadas para alcançá-las. A entidade é consciente dessas necessidades e, desde o começo das atividades, sabe que precisa dessa memória. Cencetti (1970) afirma que o arquivo é a entidade, deslocando o plano existencial para o arquivo, que se torna vivo. No caso da extinção da entidade produtora, será o arquivo a substituí-la totalmente. Claro que nessa afirmação se percebe uma matriz ideológica idealista e positivista amplamente criticada.

Com efeito, se o arquivo é o duplo da entidade desaparecida, não haveria qualquer perda/subtração/mudança em relação ao que se constituiu. O que coordena é o chamado vínculo histórico – ou arquivístico. Sempre conforme Cencetti (1970), os papéis são ordenadas pela forma através da qual a entidade as produziu. Sem esse vínculo só há um conjunto de papéis, não um arquivo.

A tradição arquivística italiana e holandesa elaborou o conceito de arquivo como globalidade das coisas e dos fatos: a totalidade é o arquivo, cujas partes são conectadas pelos vínculos/nexos históricos, onde “histórico” não é conceito disciplinar, mas propriamente arquivístico, isso é, que pertence à evolução da entidade a que o arquivo pertence.

Cencetti (1970) dizia que o arquivo é a entidade produtora, ou seja, mais corretamente, o arquivo é o espelho da entidade, desenhando, assim um arquivo em relação a uma metodologia particular, derivada de uma cultura historicista, romântico/idealista, que vê o arquivo como desenvolvimento orgânico da entidade.

A memória é, porém, só uma seleção da vida da entidade. Em si, portanto, existe uma seleção no arquivo, tratando-se de uma memória especificamente funcional. A forma de seleção escolhida pela entidade é, assim, importante tanto quanto sua ordem.

2 – ARTE

O problema que queremos enfrentar é da ordem do conhecimento, pois concerne a constituição da memória cultural no campo artístico. O debate sobre a autonomia da arte se desenrola ao longo da modernidade, gerando um sistema próprio, ao qual pertence um conjunto de estruturas, além do artista: galerias, *marchands*, crítica, revistas, museus. A estes é atribuída, entre outras, a função de preservar e expor as “hierarquias” da memória artística, legitimando fronteiras e limites conceituais sistêmicos das definições de arte, desempenhando estas funções de mediação entre o sistema da arte e os públicos. Desde seu desenho conceitual iluminista, os museus são deputados à seleção, guarda e visibilidade daquela que Le Goff (1978) define “herança do passado”, isto é, monumentos. Para o autor, características do monumento são suas “capacidades [...] de perpetuar sociedades históricas”, pois são memória compartilhada, “e de remeter a testemunhas só in parte mínima escritas” (Le Goff, 1978, p.38).

Ao mesmo tempo em que se assiste à configuração do sistema da arte, os fenômenos estéticos se multiplicam com o surgimento da Indústria Cultural, com a transformação originada pela grande distribuição, com os fenômenos estéticos de massa, com o aparecimento das estéticas do mercado de massa no cinema, na fotografia, nas gravações musicais, no design, nas lojas de conveniência, na moda e, mais em geral, em todos os produtos esteticamente projetados.

Uma das principais características da arte contemporânea é a multiplicidade de suas formas e manifestações. Na criação das obras, materiais clássicos são substituídos ou mesclados a outros de uso/origem industrial, doméstica, ou a objetos de uso comum.

Dessas obras frequentemente só sobram as lembranças do projeto e da realização, através da produção de documentação escrita, fotográfica, vídeo e digital que se torna novo objeto de reflexão de natureza arquivística por sua organização, preservação e acessibilidade.

Esta operação, porém, não encontra suas razões de ser realizada no âmbito do sistema da arte, pois é construção da memória cultural através da autoridade do arquivo, e não pela autoridade do museu. Atribuindo valores estéticos e artísticos a objetos produzidos industrialmente, em série, a presença de um destes objetos dentro de uma coleção de museu não é suficiente para representar a memória cultural que, por outro lado, se torna potencialmente narrativa histórica quando esta memória é formada pelos documentos do processo inteiro de criação e produção do objeto.

A documentação de todas as fases, do projeto à realização, desenha uma lógica de arquivo, representativa do processo definido por Le Goff revolução documentária. Afirma o autor que, no final da década de 1920 os fundadores da revista *Annales* insistem na necessidade de ampliar a definição de documento. Citando Febvre, e Bloch, Le Goff ilustra esta tendência: “A história se faz com os documentos escritos, certamente, quando existem. Mas pode ser feita, deve ser feita, sem documentos escritos, se não tiver”, (Le Goff, 1978, p. 41). O “historiador das religiões [...] bem sabe: as imagens pintadas ou esculpidas nos muros dos santuários, a disposição e a decoração dos túmulos podem lhes dizer [...] pelo menos quanto muitos escritos”. Particularmente interessante a citação de Samaran de 1961: “O termo documento deve ser entendido no sentido mais amplo, documento escrito, ilustrado, transmitido mediante o som, a imagem, ou em qualquer outra maneira” (Le Goff, 1978, p.41).

No *Traité de Documentation* de Paul Otlet (1934), o autor explica que os livros são os “éléments servant à indiquer ou reproduire une pensée envisagée sous n'importe quelle forme o (OTLET, 1934, p. 9), “(...) un support d'une certaine matière et dimension, éventuellement d'un certain pliage ou enroulement sur

lequel sont portés des signes représentatifs de certaines données intellectuelles” (OTLET, 1934, p.43). Otlet (1934) propõe uma terminologia enraizada no Documento, com os termos documento, documentação, documentalista, documentar, documentário entre outras. Livros e documentos eram classificados em quatro grandes grupos:

- a) Les documents proprement bibliographiques.
- b) Les documents graphiques, autres que les publications imprimées et les manuscrits d'ordre littéraire et scientifique.
- c) Les documents qui, sans être bibliographiques ni graphiques, sont cependant des équivalents ou des substituts du livre.
- d) Les documents qui sont le résultats de l'enregistrement, sous toutes formes de données relatives à l'administration publique et privée, aux « affaires » (correspondance, notes, rapport, comptes, registres, état, listes et répertoires, etc.).(OTLET, 1934, p. 124)

Ainda, Otlet afirma que "(...) les œuvres d'art sont rattachées à la documentation, puisque les documents se définissent incorporation de données susceptibles de reproduction." (OTLET, 1934, p. 247).

Trata-se das definições de documento em volta das quais se desenvolvem os debates acerca do termo documento, definições que permitem a Le Goff, através das citações de Bloch, Fevre e Samaran, de falar de revolução documentária. Ainda, Suzanne Briet (1951) define documento “Tout indice concret ou symbolique, conservé ou enregistré, aux fins de représenter, de reconstituer ou de prouver un phénomène ou physique ou intellectuel”. (BRIET, 1951, p.7).

Para Le Goff a revolução documentaria promove um elemento crucial para a memória cultural artística: tende, com efeito, a “promover uma nova unidade de informação [...] que leva à série [...] Tornam-se necessários novos arquivos [...]. A memória coletiva [...] organiza-se em patrimônio cultural. O novo documento é armazenado e manuseado nos bancos de dados” (Le Goff, 1978, p. 42).

O deslocamento conceitual de documento tem consequências empíricas, pois permite a instituição de lugares alternativos aos museus de arte enquanto espaços onde está presente a experiência estética, em uma lógica documentalista.

Como vimos, arquivo foi, por longo tempo, instrumento ordenado de documentação na base da estrutura memorial do passado, enquanto hoje na definição de arquivo encontra-se a invenção e sua narrativa se torna a construção composta pelos materiais selecionados por grupos ou indivíduos que lhes fornecem forma e conteúdo. Desde que se tornou possível definir o arquivo, nas palavras de Deleuze, como álbum audiovisual de uma época, a visibilidade e a enunciabilidade tornaram-se os estratos das formações históricas que constituem o *corpus* destes arquivos, os quais, como alerta Foucault, não são acumulação de textos preservados por uma cultura, nem os documentos de seu passado nem os testemunhos de sua identidade. Também não são as instituições que permitem o registro e a preservação dos discursos a serem lembrados. Para Foucault, arquivo é, em princípio, a lei do que pode ser dito; o sistema que rege o aparecimento dos enunciados como acontecimentos singulares. É um arquivo em que as coisas ditas permitem que os enunciados se articulem em figuras diversas e se combinem entre si em múltiplas relações.

Foster (2016) adota o conceito de Foucault, com a finalidade de esboçar as mudanças significativas nas relações arquivísticas dominantes entre a prática da arte, o museu de arte e a história da arte no Ocidente:

a estrutura mnemônica produzida pelas três instâncias no século XX. Em seu artigo a dialética Malraux/Benjamin aparenta ser a mais eficaz na análise do CSAC, Centro di studi e Archivio della Comunicazione. Se Benjamin antecipa a ruína definitiva do museu – e do arquivo – provocada pela fotografia, Malraux propõe sua expansão infinita, mobilizada exatamente pela fotografia. O Museu Imaginário de Malraux reúne os fragmentos em um arquivo que permite um sistema de continuidades globais, capaz de transformar o caos das imagens em uma ordem documentária, entre arquivo e museu.

No caso da arte contemporânea, portanto, ao lado das necessidades de preservação da ampla tipologia de materiais é necessário pensar nas modalidades de preservação da ação do próprio artista, para além da realização final da obra, ação que deve ser amplamente documentada. Nesse sentido, um aspecto metodologicamente mais complexo no âmbito da arte contemporânea é o da documentação que, muitas vezes, substitui as obras definitivas mais efêmeras e conceituais.

Uma documentação completa da obra desde sua concepção até sua exposição permite, em teoria, que mesmo sem instruções diretas do artista, ela possa ser apresentada como percurso fiel ao original: é nessa representação de natureza arquivística que reside o nó da questão: não teremos a obra, assim como um arquivo não é a entidade produtora, mas sua representação selecionada e organizada, que permite que o arquivista/historiador da arte realize o percurso do lastro documentário, não mais unicamente cartáceo, mas constituído por objetos variados, inclusive tridimensionais.

É claro que réplicas e proposições novas de obras efêmeras, *happenings*, *performances* e instalações possuem, na maioria dos casos, somente valor documentário, de testemunha histórica, cujo valor estético precisa ser repensado. Apesar disso, essa postura arquivística” se torna importante e transfere as questões de registro não somente no plano da documentação escrita, ampliando o conceito de arquivo tradicional.

3 – CSAC

Após essas breves observações sobre arquivos e arte, identificamos uma proposta pioneira nessa lógica de arquivo. Trata-se do CSAC – Centro Studi e Archivio della Comunicazione da Università de Parma -, instituição surgida em 1968 e que, desde 2015, tornou-se arquivo-museu. Na base do CSAC encontra-se a ideia que, basicamente, era possível ver a arte pela História Cultural, fundamentada na aplicação das ideias de Le Goff (1978) e da leitura oferecida por Benjamin (1962) no âmbito da produção industrial de novos objetos, que permitem uma leitura contemporânea das artes figurativas, da escultura, do projeto arquitetônico e de design, da moda e da fotografia. A ideia do CSAC é de que existem qualidades estéticas de nível diferente na comunicação. Por exemplo, o *design* é visto como produto criativo de um projetista voltado à produção industrial: a ideia do original, nesse caso, deve ser revisada, pois não é mais ligada às peças produzidas, cada um sendo cópia e, ao mesmo tempo, original, enquanto portador da marca do autor do projeto. A escolha do fundador do CSAC, Arturo Carlo Quintavalle (2010), é de evidenciar, nas coleções desenvolvidas, o debate sobre *design*, enquanto relacionado à produção industrial, sobre as artes figurativas, sobre a produção gráfica e fotográfica.

O CSAC busca responder à necessidade de repensar o que constitui uma coleção e qual pode ser o lugar e o sistema de organização e exposição melhor para acolhe-la. Torna-se, assim, um verdadeiro laboratório de experimentação das potencialidades comunicacionais e estéticas de linguagens diferentes.

A proposta do CSAC, que desenvolve em poucos anos coleções heterogêneas de pintura, escultura, fotografia, gráfica, moda, design (graças a um sistema de doações do qual Quintavalle é o grande articulador) oferece hoje acerca de 12.000.000 “peças” de arquivo aos usuários interessados. A proposta de articular suas coleções como arquivo, isso é, escolhendo uma lógica de “série documentárias” e não, como um museu, selecionando obras na base de uma presumida unicidade, é uma proposta antagonista ao próprio museu em quanto lugar da “unicidade”. O lugar ao qual pertence, idealmente, a serialidade das “peças”, explica Quintavalle (2010), é um arquivo que pode conter a diversidade de materiais (projetos, objetos, fotografias, pinturas, esculturas, roupas). Em perspectiva, no CSAC pode ser observada a constituição de um modelo inovador na estrutura organizativa do conhecimento da contemporaneidade histórica e artística, através das articulações documentárias que abrem reflexões sobre questões relativas às fronteiras da arte, da fotografia artística e de reportagem, do desaparecimento da hierarquia entre desenho de projeto, *affiche*, ilustração, hierarquia que galerias e museus – partes do sistema da arte – estabelecem em suas políticas de vendas e aquisições.

CONCLUSÕES

O que o CSAC realiza em seu desenvolvimento é uma cisão com as tradições estéticas iluministas e pós-românticas da arte, aplicando as novidades teóricas da história antropológica de matriz francesa, proposta pelos medievalistas, sim, mas materializada no campo das expressões artísticas do século XX: imagens, documentos, monumentos impõem uma ampliação nas definições e nas coleções dos arquivos.

Quintavalle (2010) oferece um exemplo de procedimentos de análise que privilegiam abordagens de arquivo, em que todos os “documentos adquirem o mesmo valor, em detrimento da lógica hierárquica dos museus.

O arquivo do CSAC, como foi concebido, preserva – assim como os arquivos tradicionais – o princípio da serialidade. Diferente do museu, funda-se, idealmente, em concepções não hierárquica dos documentos/monumentos que acolhe, transformando-se em uma “máquina” histórica que expõe efetivamente a identidade possível entre documentos e monumentos.

Claramente, não se trata da tipologia dos materiais colecionados, aquilo que diferencia o CSAC dos museus, pois esses também colecionam objetos de produção industrial, por exemplo. O que a proposta do CSAC desenvolveu é uma questão conceitual sobre o conhecimento artístico na contemporaneidade, opondo-se ao conceito de museu como lugar da peça “único” em favor da reconstituição dos processos de elaboração artística, que graças ao arquivo pode acessada. Questão conceitual, portanto, essa que se expressa no modelo de arquivo e não de museu, em que o produto da arte é semiotizado, desarticulando (sempre idealmente, pois resta aprofundar em outro momento a eficácia da operação e o que aconteceu com sua abertura ao público como Arquivo-Museu em 2015) O que é certo é que a destituição das hierarquias do museu da arte, em favor do modelo do arquivo, redimensiona a interpretação estética, favorecendo a historicidade dos processos produtivos da arte.

REFERÊNCIAS BIBLIOGRÁFICAS

- Becker, H. (1999). *Falando da sociedade*. Rio de Janeiro, Brasil: Jorge Zahar.
- Benjamin, W. (1962). *Angelus Novus: saggi e frammenti*. Torino, Itália: Einaudi.
- Briet, S. (1951). *Qu'est-ce que la documentation*, Paris, França: EDIT.
- Cencetti, G. (1970). *Scritti archivistici*. Roma, Itália: Il Centro di ricerca (Fonti e studi di storia, legislazione e tecnica degli archivi moderni, 3).
- Deleuze, G. (2014). *Il sapere: Corso su Michel Foucault (1985-1986) – I*. Verona, Itália: Ombre Corte.
- Foster, H. (2016). *Design e crime (e outras diatribes)*. Belo Horizonte, Brasil: UFMG.
- Foucault, M. (1994). *L'archeologia del sapere: una metodologia per la storia della cultura*. Milano, Itália: BUR.
- Le Goff, J. (1978). *Documento/Monumento*. In: AA.VV.: *Enciclopédia Einaudi*, vol. V (pp. 39-48). Torino: Einaudi.
- Lemieux, V. Ouimet M. (2008). *Análise Estrutural das Redes sociais*. Lisboa: Instituto Piaget.
- Otlet, P. (1934). *Traite de Documentation: le livre sur le livre*. Bruxelles, Bélgica: Editones Mundaneum-Palais Mondial.
- Quintavalle, A. C. (2010). *Nove100*. Milano, Itália: Skira.



UM ESTUDO PRELIMINAR SOBRE A APLICABILIDADE E EXTENSÃO DO MODELO SECI COM O USO DE DISPOSITIVOS DA INTERNET DAS COISAS

Antônio José Rodrigues Neto¹, Renato Souza Rocha²

¹Universidade de Coimbra, 0000-0003-2364-0927, neto@student.uc.pt

²Fundação Getúlio Vargas, 0000-0002-1895-3905, renato.souza@fgv.br

RESUMO Até o ano 2020 estima-se que aproximadamente 24 bilhões de dispositivos serão conectados na nuvem em um ambiente digital colaborativo de dados, informação e conhecimento. A sociedade está passando por mudanças sociais, econômicas e tecnológicas e, devido à importância desse conhecimento para todas as organizações, torna-se necessário mudar nossa maneira de pensar sobre a inovação organizacional, estratégica e competitiva para todos os envolvidos. Refletindo em torno da conceptualização de dados, informação e conhecimento, o presente trabalho utiliza métodos qualitativos de pesquisa exploratória com o objetivo de discutir a influência do uso da Internet das Coisas na sociedade atual, em uma visão da máquina como novo membro da sociedade atual e seus potenciais riscos, bem como uma perspectiva global sobre os novos serviços de informação a serem oferecidos neste novo modelo de interação entre homem e máquina, bem como um novo modelo baseando-se em serviços e sub-serviços distribuídos na nuvem. Este trabalho também sugere um estudo preliminar sobre a aplicabilidade de um modelo interativo e colaborativo onde todos os componentes do modelo passam a estar interconectados numa rede distribuída de conhecimento, executando um papel ativo e contínuo de colaboração, promovendo, deste modo, a gestão (criação, troca, oferecimento e compartilhamento) de forma segura e interoperável de serviços e informações, nesta nova era da sociedade digital.

PALAVRAS-CHAVE *Modelo SECI, Internet das Coisas, Organização do Conhecimento*

ABSTRACT By the year 2020, it is estimated that approximately 24 billion devices will be connected in the cloud in a collaborative digital environment with data, information and knowledge. The society is going through social, economic and technological changes and, due to the importance of this knowledge for all organizations, it becomes necessary to change our way of thinking about organizational, strategic and competitive innovation for all the involved members of the society. Reflecting around the conceptualization of data, information and knowledge, the present work uses qualitative methods of exploratory research with the objective of discussing the influence of the use of the Internet of Things (IoT) within the current society, as a new way to view the machine as a new member of the society and its potential risks, as a new global perspective related to information services to be offered within this new model of human and machine interaction, in a new model based on distributed services and sub-services in the cloud. This work also suggests a preliminary study on the applicability of an interactive and collaborative model where all the components of the model become interconnected in a distributed network of knowledge, performing an active and continuous role of collaboration, thus promoting management (creation, exchange, offering and sharing) in a secure and interoperable way of services and information in this new era of the digital society.

KEYWORDS *SECI Model, Internet of Things, Knowledge Management*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A sociedade moderna foi-se transformando gradualmente numa sociedade do conhecimento (Nonaka & Takeuchi, 1995). Um importante fator de produtividade e competitividade para as organizações está na capacidade dos indivíduos, e das próprias organizações, criarem, processarem e transformarem informações e conhecimento. Devido à importância do conhecimento para a sociedade, torna-se necessária uma mudança no nosso pensamento no que diz respeito à inovação organizacional, estratégica e competitiva das organizações (Klaus, Schmidt, Schreiber, & Bessi, 2016).

A revolução digital, através do uso da *Internet e, de sua extensão, a Internet of Things (IoT)* (Atzori, Iera, & Morabito, 2010) e da nuvem (Buyya, Yeo, Venugopal, Broberg, & Brandic, 2009), faz com que a sociedade do conhecimento passe por uma série de mudanças económicas, sociais, organizacionais e tecnológicas, configurando assim um novo modelo sócio-técnico-económico caracterizado por uma nova era na sociedade digital (Chen & Chen, 2006; Nonaka & Takeuchi, 1995).

A Internet das Coisas (*Internet of Things, IoT*) teve seu conceito popularizado no ano de 1999, sendo definida como uma rede digital de dispositivos interligados (Figura 1). Esse conjunto de dispositivos podem ser, mas não excluindo outros, os sensores, *smartphones*, pulseiras inteligentes de monitoração (*wearable devices*), câmeras digitais, dispositivos instalados em nossos veículos, dispositivos utilizados em nossas casas, como por exemplo o *Amazon Dash Button*¹, *Amazon Echo & Alexa*², *Roomba Robot Vacuum*³, entre outros (Abashidze & Dabrowski, 2016).

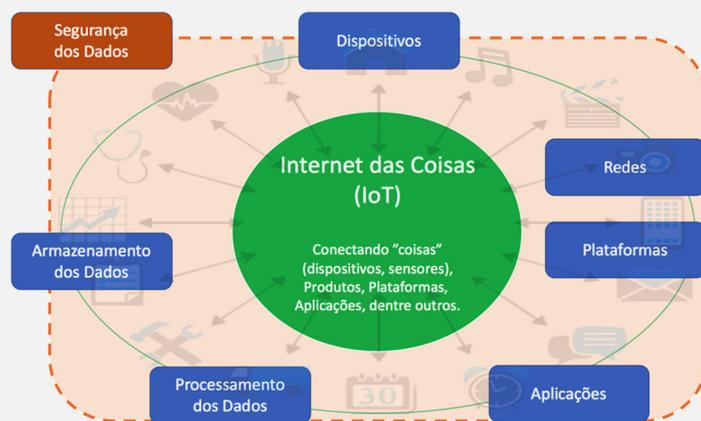


Figura 1. Uma visão geral sobre a Internet das Coisas (IoT).

Fonte: Elaboração nossa.

¹ O dispositivo *Amazon Dash Button* é um dispositivo Wi-Fi que permite ao usuário com um simples apertar de um botão, a (re)compra de produtos domésticos na empresa Amazon – <https://www.amazon.com/Dash-Buttons/b?ie=UTF8&node=10667898011>.

² Echo & Alexa são dispositivos que permitem aos usuários escolher músicas, fazer compras, controlar sua *casa inteligente*, buscar informações sobre notícias, clima, entre outros – https://www.amazon.com/b/?node=9818047011&ref=fs_ods_fs_aucc_cp_t2

³ O *Roomba Robot Vacuum* é um dispositivo destinado para limpeza de carpetes de forma automática e programável – <http://store.irobot.com/default/home>.

Richard Weber (2016) cita um ponto interessante onde Neil Gross (1999) comenta que o planeta Terra, no próximo século, estará vestindo uma pele eletrônica, e irá utilizar a Internet como um “andaime” de apoio, facilitando a transmissão de suas sensações de um ponto a outro.

Atualmente, vivemos numa sociedade onde o poder da informação e do conhecimento tem aumentado constante e exponencialmente, desempenhando cada vez mais um papel central em nossa sociedade (Hjørland & Capurro, 2007), onde a sociedade “[...] is becoming more knowledge based” (Nonaka & Takeuchi, 1995, p. 967-968).

Devido à importância do conhecimento para todos e da quantidade de informações que estão sendo produzidas e transitadas na rede atualmente, torna-se necessário uma mudança no nosso pensamento no que diz respeito à inovação social, estratégica e competitiva das organizações e de todos os seus indivíduos (Klaus et al., 2016).

Como referido em um texto de Terry Kony em 1998 (“*A Digital Dark Ages? Challenges in the Preservation of Electronic Information*”), o conteúdo digital passou a representar uma mudança na tomada de consciência no início da utilização e da expansão da Internet, onde o mundo digital se consolidava como uma realidade para indivíduos, organizações, e para a sociedade na época (Corujo, 2014).

Segundo Corujo (2014), Terry Kony já alertava a sociedade para que não entrasse em uma idade das trevas digital. Segundo Gantz e Reinsel (2012), desde o ano de 2012 o universo digital passa a dobrar de tamanho a cada dois anos, ou seja, uma analogia à “idade das trevas digitais”, mencionada por Terry Kony, que pode ser vista nos dias atuais com o crescimento exponencial e “descontrolado” da Internet, das redes sociais e da IoT.

Para uma sociedade e seus indivíduos, não basta possuir informação e conhecimento sem a possibilidade de armazená-la, organizá-la, encontrá-la e distribuí-la quando quiser. Em outras palavras, torna-se necessário a existência de uma forma eficiente de gerenciar a informação, principalmente em uma era na qual possuímos dispositivos que geram e distribuem informações, mas potencialmente, sem o nosso conhecimento e/ou consentimento.

Portanto, com a massiva quantidade de dados digitais sendo produzidos e do crescimento “descontrolado” dos dispositivos digitais em nossa sociedade, passa a ser de vital importância o entendimento de um sistema de informação e seus serviços no campo da Ciência da Informação, bem como o seu uso em uma nova versão digital da sociedade contemporânea.

Nesse contexto, este trabalho tem como objetivo cujo objetivo é contribuir para um estudo preliminar sobre a potencial aplicabilidade e extensão do modelo SECI com o uso dos dispositivos distribuídos na rede digital de dados denominada Internet das Coisas (*IoT – Internet of Things*). Ele também sugere uma proposta na adaptação e extensão do modelo SECI em um novo modelo baseando-se em serviços e sub-serviços distribuídos na nuvem, onde os seus componentes passam a estar interconectados numa rede digital de dispositivos, executando um papel ativo e contínuo de colaboração, e promovendo, deste modo, a criação, troca e compartilhamento de conhecimento de forma segura e interoperável em uma nova era da sociedade digital.

METODOLOGIA

Para responder ao objetivo proposto no presente trabalho, foi efetuada uma revisão da literatura. A pesquisa bibliográfica foi realizada nas bases de dados disponibilizadas pela Universidade de Coimbra. Os termos usados para a recuperação de informação foram “*Internet of Things*”, “*SECI Model*” e “*Knowledge Management*”.

Para o desenvolvimento de um trabalho fundamentado e sem o comprometimento da qualidade da pesquisa (Gil, 2002), tornou-se necessária a busca de fontes de pesquisa primária em diversos trabalhos que julgamos necessários e importantes para a elaboração deste, além de diversos conteúdos que não estavam disponíveis através do acesso pelo Portal b-on⁴, assim como alguns outros trabalhos científicos disponíveis através do acesso ao Google Scholar⁵, utilizado como um recurso adicional na análise de citações e relevância.

Em uma fase posterior foi feito uma análise transversal de todos os artigos reunidos tendo como base o resumo, palavras-chave, introdução e conclusão, com a finalidade de determinar e selecionar os documentos relevantes e de interesse para o estudo.

Terminada a fase de pesquisa e seleção bibliográfica, passou-se a uma análise aprofundada dos conceitos fundamentais de modo a poder refletir sobre tais conteúdos com a finalidade de responder aos objetivos de pesquisa delineados anteriormente, tendo como objetivo discutir a influência da utilização da IoT na sociedade atual e sugerir um estudo preliminar sobre a aplicabilidade de um modelo seguro, interativo e colaborativo que possibilite a gestão (comunicação e coleta) das informações dos dispositivos da IoT de forma transparente e segura para a sociedade atual.

RESULTADOS

No ano de 2007 o *International Data Corporation* (IDC) descreveu que vivemos num mundo inundados de dados digitais com 281 exabytes (2.25×10^{21} bits) estimados, equivalendo a 281 trilhões⁶ de novelas digitalizadas, mas representando somente 1% do número de Avogadro que corresponde ao número de átomos em 12 gramas de carbono (6.022×10^{23}) (Fox & Hill, 2007).

O mesmo estudo também demonstrou que, no ano de 2007, foi atingido um marco histórico denominado ano do “cross-over”, onde o mundo digital passou a ter mais informações digitais produzidas do que área de dados disponíveis para as armazenar. Estima-se ainda que, a partir do ano de 2020, a quantidade de informações digitais produzidas irá superar o número de Avogadro (Berman, 2008).

A criação do conhecimento organizacional consolida-se, nos termos e processos organizacionais, com a finalidade de auxiliar e amplificar a criação do conhecimento, através dos seus indivíduos, como parte da rede de conhecimento da organização. Através dessa linha de metodologia torna-se possível a

⁴ Portal b-on disponível em <http://www.b-on.pt>.

⁵ Google Scholar disponível em <https://scholar.google.com>.

⁶ Um trilhão (*one trillion*) corresponde a um fator de 10^{12} ou 1,000,000,000,000.

identificação e distinção dos diversos níveis de interação social, em que o conhecimento foi transformado e reconhecido.

A interação entre as duas formas de conhecimento (tácito e explícito) é a chave para a criação de um conhecimento pleno, único e competitivo pelas organizações, onde o processo de aquisição do conhecimento organizacional se baseia numa relação dinâmica e contínua entre os dois tipos de conhecimento, num processo social, de forma interativa, contínua e em espiral entre os indivíduos na organização (Nonaka, 2007).

Tendo em conta o exposto, sugerimos um estudo preliminar sobre a aplicabilidade de um modelo interativo e colaborativo sobre os serviços oferecidos pelos dispositivos da IoT, tendo como base o modelo SECI (Espiral do Conhecimento) (Nonaka & Takeuchi, 1995) onde, desta forma, todos os componentes do modelo passam a estar interconectados numa rede distribuída de conhecimento executando um papel ativo e contínuo de colaboração, promovendo, deste modo, a gestão (criação, troca e compartilhamento) das informações e dos serviços em uma nova era da sociedade digital (Figura 2).

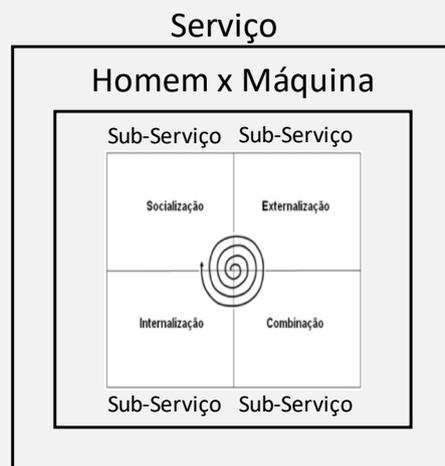


Figura 2. Modelo colaborativo de serviços entre homens e máquinas baseado no Modelo SECI.

Fonte: Elaboração nossa.

Qualquer forma de interação para com a sociedade baseia-se em comunicação, que pode ser considerada como o eixo central na interação entre indivíduos em uma sociedade moderna. Quando nos referimos especificamente sobre comunicação social e/ou digital, os dispositivos da IoT oferecem novas e diversificadas possibilidades para seus fabricantes (*marketers*) como, por exemplo, a integração desses dispositivos para com as redes sociais (Kotler, 2010; Kotler, 2017).

O comportamento social e o uso apropriado desses dispositivos, principalmente na gestão do uso e da coleta de informações, tornam-se ainda mais cruciais à medida que convivemos cada vez mais com a IoT, já que este é um ambiente interconectado que cresce exponencialmente a cada dia sendo considerado como o espaço “*cyber-physical-biological*” que liga dispositivos, sistemas, dados e pessoas (Berman & Cerf, 2017). Segundo os mesmos autores, tecnologias não possuem ética, sendo que muitos dos sistemas existentes podem ser usados para o bem e/ou mal. Faz-se aqui um questionamento sobre uma potencial necessidade no desenvolvimento de *éticas artificiais* para sistemas automatizados no auxílio das decisões. O processo para o desenvolvimento de um modelo de éticas artificiais não é um processo simples pois abrange diversos outros aspectos tais como tecnológicos,

humanos, organizacionais e sociais, além da definição de um planejamento a nível político que consiste em regulamentações visando a definição de regras na captura e no uso das informações utilizadas pelos dispositivos da IoT.

Pensando no aspecto tecnológico, propõe-se que os dispositivos, em uma rede digital, possuam uma linguagem padrão de comunicação (criação, consulta e oferecimento) para gestão das informações, além dos serviços a serem oferecidos e disponibilizados pelos dispositivos, tornando-os assim dispositivos de serviços abertos (*open service devices*) em um modelo colaborativo na sociedade digital.

Este modelo passa a permitir que os consumidores (seres humanos) ou até mesmo da interação entre máquinas, possam ter a liberdade e flexibilidade na comunicação, assim como a escolha, privacidade e o conhecimento dos serviços que estão sendo oferecidos. Desta forma, está se proporcionando um claro entendimento das regras do jogo e de como e quais informações estão sendo utilizadas e capturadas, possibilitando o poder de decisão do usuário final, ou seja, o homem, na escolha do que compartilhar ou não, visando assim garantir nossa privacidade no mundo digital. Resumindo, o *livre arbítrio digital*. (Figura 3).

Conforme pode ser visto na Figura 3, o ser humano passa a interagir com o SI (dispositivo da IoT) através de uma linguagem padrão, permitindo a gestão da informação (entrada) em todos os serviços prestados pelos dispositivos digitais (SI), permitindo que os consumidores possam escolher quais são as informações que podem ou não ser coletadas, e quais os serviços (saídas) que possam ser escolhidos (livre-arbítrio) em um determinado contexto promovendo assim a satisfação dos consumidores desse modelo.

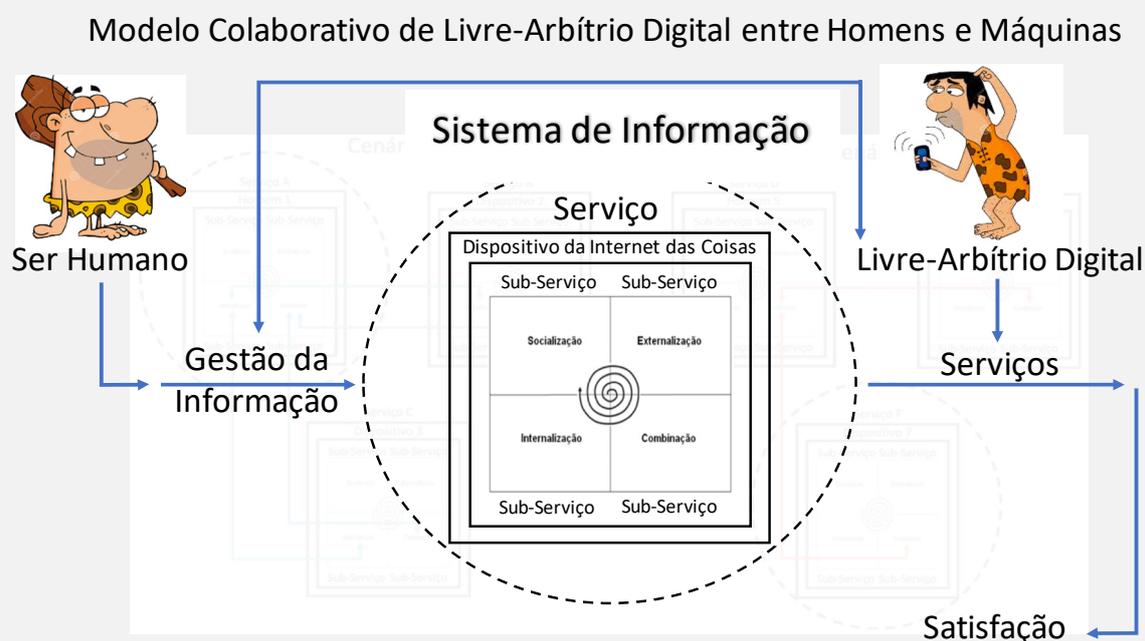


Figura 3. Modelo colaborativo de Livre-Arbítrio Digital entre Homens e Máquinas.

Fonte: Elaboração nossa.

Caso o indivíduo humano (Homem 1) necessite de uma consulta aos serviços oferecidos pelos Dispositivos 2 e 3 (Figura 4, Cenário 1), poderá se utilizar de um modelo de conexão na aquisição do conhecimento através do sub-serviço *Combinação*, que pode ser demonstrado como:

1 → 2 onde: {Homem 1: A: Combinação (Consulta)} → {Dispositivo 2: B: Combinação (Consulta)}
 1 → 3 onde: {Homem 1: A: Combinação (Consulta)} → {Dispositivo 3: C: Combinação (Consulta)}

Modelo Colaborativo de Interação entre Homens e Máquinas – baseado no Modelo SECI

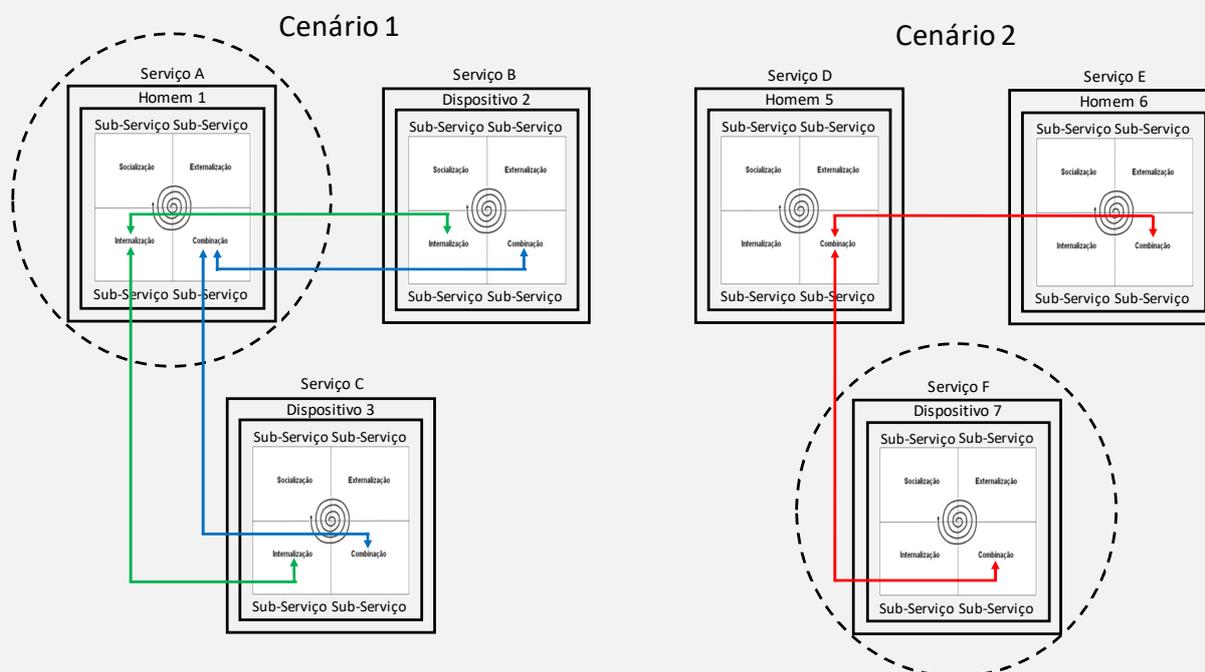


Figura 4. Modelo colaborativo – novos indivíduos da sociedade moderna – homens e máquinas.

Fonte: Elaboração nossa.

O domínio digital em um processo colaborativo na troca de conhecimento dos serviços oferecidos passa a ser baseado num modelo distribuído SECI, com a possibilidade da existência de centenas de milhares de *containers*⁷ principais representando diferentes tipos de serviços associados com os seus respectivos indivíduos (*sub containers*), onde cada *sub container* representa os processos característicos da Espiral do Conhecimento.

Um possível exemplo da utilização do modelo SECI e da sua aplicação no mundo real é:

O usuário (Homem 1) adquire um dispositivo especializado no ramo da saúde e monitoração de atividades esportivas dos seus usuários, que possuem bases de dados especializadas para o armazenamento das informações enviadas pelas pulseiras inteligentes. As pulseiras são dispositivos inteligentes que enviam dados dos usuários para Sistemas de Gestão de Informação na nuvem, contendo informações que não conhecemos ou não dominamos por completo. Assumimos então que as pulseiras inteligentes possam enviar as seguintes informações:

- 1) Qual a distância percorrida pelos usuários?
- 2) Qual a rotina dos usuários? Horário de acordar? E dormir?
- 3) Quantos passos foram dados pelos usuários?
- 4) Batimento cardíaco dos usuários?

⁷ Os *containers* são um método de virtualização de sistema operacional que permite executar uma aplicação e suas dependências em processos com recursos isolados – <https://aws.amazon.com/pt/containers>.

5) Localização do usuário (GPS).

Se o consumidor, no caso o Homem 1, não possui o controle dos dados que estão sendo enviados para o Sistema de Gestão de Informação e, conseqüentemente, quais serviços estão sendo oferecidos e consumidos, estabelecemos uma operação que podemos chamar de “versão 2.0 de fragilidade social”. Sem o conhecimento dos serviços coletados e oferecidos pela pulseira digital, cria-se a possibilidade de sermos monitorados a todo momento sem o nosso consentimento.

Com a adoção e desenvolvimento desse modelo tecnológico colaborativo, todo dispositivo passa a possuir uma linguagem padrão, permitindo a gestão da informação: ou seja, a criação, alteração, consulta e interação de todos os serviços prestados pelos dispositivos digitais. Este modelo tem como objetivo permitir que os consumidores possam escolher quais são as informações que podem ou não ser coletadas, e quais os serviços que possam ser escolhidos em um determinado contexto; por exemplo, como não autorizar um serviço que envia os itens 2 e 5 quando estamos falando de dispositivos utilizados por adolescentes, onde o poder de decisão de enviar as informações ou não seria do consumidor, e não de quem oferece o serviço.

Cada processo e serviço passam, assim, a atuar como um serviço conhecido e público na nuvem interconectada com diferentes membros da sociedade, sendo a principal finalidade a de desenvolver um ambiente colaborativo, interoperável e contínuo para a criação, troca e consulta dos serviços oferecidos de forma equitativa entre todos em uma nova sociedade.

CONCLUSÕES

A Internet das Coisas (Brynjolfsson, 2014) passa a ter uma importante influência no desenvolvimento dos serviços e na disseminação do conhecimento que, além de proporcionar um ponto central de acesso para o domínio do conhecimento, disponibiliza uma plataforma de recursos e serviços colaborativos entre homens e máquinas (Rafiq & Bashar, 2014).

A problemática surge com o desconhecimento dos serviços oferecidos ou utilizados pelos dispositivos em uma sociedade digital que, conforme Umberto Eco, nos coloca ainda em um mundo selvagem e perigoso. Segundo Sugahara, Fuentes, e Oliveira (2003), se não existir uma preocupação com o interesse dos consumidores (clientes), de nada adiantará o avanço tecnológico. Os produtos e seus respectivos serviços automatizados precisam entender e principalmente respeitar as necessidades e exigências de qualidade e confiabilidade em relação ao contexto social que se aplicam: a de seus usuários.

Conforme mencionado anteriormente, o processo para o desenvolvimento de um modelo de interação humano-digital com o auxílio de éticas artificiais não é um processo simples. Ele abrange diversos outros aspectos não somente tecnológicos, mas humanos, organizacionais e sociais, consistindo em regulamentações que visam o desenvolvimento e a definição de políticas na interação dos dispositivos das IoT para com a sociedade, com a finalidade de estabelecer uma convivência saudável e confiável entre homens e os dispositivos da IoT. Com a adoção e desenvolvimento desse modelo tecnológico interativo e colaborativo, todo dispositivo passa a possuir uma linguagem padrão, permitindo assim a gestão da informação (criação, alteração, consulta e interação) e de todos os serviços disponíveis e prestados pelos dispositivos digitais para seus usuários e/ou consumidores.

Esse modelo também tem como finalidade a implementação do livre-arbítrio, de quais informações podem ou não ser coletadas e distribuídas, e quais são os serviços disponíveis e acessíveis. Isso tudo sendo baseado em um determinado contexto social, económico, tecnológico ou qualquer outro, com o objetivo de aplicar a menção de Umberto Eco onde “conhecer é cortar, é selecionar”.

Sem um modelo de gestão de informação dos dispositivos digitais da IoT, que seja acessível, controlado e gerenciado pelos verdadeiros indivíduos com mentes, corações e espíritos, a sociedade passa a ser cada vez mais gerida ou dominada pela inteligência artificial, caracterizando-se assim, não somente como uma nova versão digital da sociedade contemporânea, mas com o surgimento de um novo e assustador apocalipse social.

Deste modo, a principal finalidade deste trabalho foi a de propor um estudo preliminar sob a aplicabilidade e extensão do modelo SECI com o uso dos dispositivos da Internet das Coisas em serviços e sub-serviços distribuídos na nuvem, onde os componentes do modelo possuem características flexíveis, colaborativas e de fácil adaptação às necessidades da organização do conhecimento das organizações.

Os componentes no modelo (serviços e sub-serviços) passam a estar interconectados numa rede digital do conhecimento, passando a executar um papel ativo e contínuo de colaboração, e promovendo, deste modo, a criação, troca e compartilhamento, de forma segura e interoperável, dos conhecimentos adquiridos no domínio do conhecimento entre os diferentes membros numa nova era da sociedade digital.

REFERÊNCIAS BIBLIOGRÁFICAS

- Abashidze, I., & Dabrowski, M. (2016). Internet of Things in marketing: opportunities and security issues. *Management Systems in Production Engineering, nr 4 (24)*. <https://doi.org/10.12914/MSPE-01-04-2016>
- Atzori, L., Iera, A., & Morabito, G. (2010). The Internet of Things: A survey. *Computer Networks, 54(15)*, 2787–2805. <https://doi.org/10.1016/j.comnet.2010.05.010>
- Berman, F. (2008). Got Data? a Guide to Data Preservation in the information age. *Communications of the ACM, 51(12)*, 50–56. <https://doi.org/10.1145/1409360.1409376>
- Berman, F., & Cerf, V. G. (2017). Social and ethical behavior in the internet of things. *Communications of the ACM, 60(2)*, 6–7. <https://doi.org/10.1145/3036698>
- Brynjolfsson, Erik. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W. W. Norton & Company. Kindle Edition. ASIN: B00D97HPQI
- Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J., & Brandic, I. (2009). Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. *Future Generation Computer Systems, 25(6)*, 599–616. <https://doi.org/10.1016/j.future.2008.12.001>
- Corujo, L. M. N. (2014). *Repositórios Digitais e Confiança Um exemplo de repositório de Preservação Digital: o RODA*.

Fox, R. F., & Hill, T. P. (2007). An Exact Value for Avogadro's Number. *American Scientist*, 95(2), 104–107. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&site=eds-live&db=a9h&AN=24171597>

Gantz, J., & Reinsel, D. (2012). THE DIGITAL UNIVERSE IN 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East. *IDC*.

Gil, A. C. (2002). Como elaborar projetos de pesquisa. Editora Atlas, 4. ed. São Paulo, 2002

Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, 29(7), 1645–1660. <https://doi.org/10.1016/j.future.2013.01.010>

Hjørland, B., & Capurro, R. (2007). O Conceito da Informação. *Perspectivas Em Ciência Da Informação*, 12(1), 148–207. Retrieved from <http://bogliolo.eci.ufmg.br/downloads/CAPURRO.pdf>

Klaus, V., Schmidt, S., Schreiber, D., & Bessi, V. G. (2016). Criação de conhecimento em uma empresa de software. *Pensamento & Realidade*. Retrieved from <http://revistas.pucsp.br/index.php/pensamentorealidade/article/view/24155>

Kotler, P. (2010). Marketing 3.0: From Products to Customers to the Human Spirit. Wiley. Kindle Edition. 978-0-470-60979-8

Kotler, P. (2017). Marketing 4.0: Moving from Traditional to Digital. Wiley. Kindle Edition. ISBN 9781119341147

Nonaka, I. (2007). The Knowledge-Creating Company. *Harvard Business Review*, 162–171.

Nonaka, I., & Takeuchi, H. (1995). The knowledge-creating company: How Japanese Companies Create Dynamics of Innovation. Oxford University Press. ISBN-13 978-0-19-509269-1

Rafiq, M., & Bashar, A. (2014). Innovative Trends in Knowledge Management: A Cloud Computing Perspective www.globalbizresearch.org. *Proceedings of the First Middle East Conference on Global Business, Economics, Finance and Banking*.

Sugahara, C. R., Fuentes, L. F., & Oliveira, S. M. (2003). Marketing : uma ferramenta fundamental para o profissional da informação. *TransInformação*, 15(1), 83–88. <https://doi.org/http://dx.doi.org/10.1590/S0103-37862003000100007>

Weber, R. M. (2016). Internet of Things Becomes Next Big Thing. *Journal of Financial Service Professionals*, 70(6), 43–46. Retrieved from <http://content.ebscohost.com/ContentServer.asp?T=P&P=AN&K=118873042&S=R&D=bth&EbscoContent=dGJyMNxb4kSep7A4v%2BbwOLCmr0%2BepZSr6e4SK6WxWXS&ContentCustomer=dGJyMOzprkiuqLdIuePfgeyx43zx>



PERSPETIVAS EPISTEMOLÓGICAS EM ORGANIZAÇÃO DO CONHECIMENTO

ORGANIZAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO: INTERFACES INFORMACIONAIS, COMUNICACIONAIS E LINGUÍSTICAS

Nair Yumiko Kobashi

Universidade de São Paulo, nairkobashi@gmail.com

RESUMO Discussão das relações entre os conceitos “informação-comunicação-linguagem” sob a perspectiva da Organização do conhecimento. Este estudo baseia-se na análise de literatura selecionada sobre o tema produzida na área da Ciência da informação e da Comunicação. Nesse sentido, foi analisada de forma específica a literatura francesa sobre os conceitos acima para promover melhor compreensão e distinção entre informação e comunicação. O reconhecimento das diferenças entre informação e comunicação mostra ser crucial para o campo da Organização do conhecimento. O termo informação, ao ser definido, neste campo, como representação de conteúdos de documentos, requer tratamento específico para ser comunicável. Semelhante compreensão permitiu lançar mais um olhar sobre as ferramentas e dispositivos de organização e difusão de informação como os sistemas do organização do conhecimento (SOCs), ferramentas que remetem tanto ao conteúdo quanto ao processo comunicacional.

PALAVRAS-CHAVE *Informação-comunicação-linguagem; Organização do conhecimento; Linguagens documentárias.*

ABSTRACT Discussion of the relations among the concepts "information-communication-language" from the perspective of the knowledge organization domain. This study is based on the analysis of selected literature on the subject produced in Information Science and Communication. In this sense, the French literature on the above concepts was specifically analyzed to promote a better understanding and distinction between information and communication. The recognition of the differences between information and communication proves to be crucial for the field of Knowledge Organization. The term information whether defined, in this field, as representation of contents of documents, requires specific processing to be communicable. This understanding enabled us to take a closer look at the tools and devices for information organization and dissemination such as Knowledge organization systems (SOCs), tools that refer to both the content and the communication process.

KEY-WORDS *Information-communication-language. Knowledge organization; Documentary languages.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

1. INTRODUÇÃO

Este texto sistematiza reflexões sobre os aspectos informacionais, comunicacionais e da linguagem sob o olhar da Organização do Conhecimento (OC). A análise dos termos foi realizada em um *corpus*

selecionado de textos da Ciência da Informação e Comunicação. São levantados, nesses textos, questionamentos sobre essa tríade e seus impactos nos processos informacionais.

Inicia-se a discussão sobre a referida tríade revisitando a Teoria Matemática da Comunicação, de Shannon & Weaver (1949), teoria que se popularizou e foi apropriada por inúmeras disciplinas científicas (ESCARPIT, 1976; WOLF, 2003), dentre as quais a Comunicação, a Ciência da informação e as Ciências da Computação.

A Teoria Matemática da Comunicação, ou Teoria da informação, foi formulada com o objetivo de obter precisão e eficácia no fluxo comunicativo-informativo, independentemente das características e contextos de aplicação (ESCARPIT, 1976). Desenvolvida na área das telecomunicações, foi aceita como modelo aplicável à comunicação entre duas máquinas, dois seres humanos ou entre uma máquina e um ser humano (WOLF, 2003).

Uma representação simplificada do modelo aparece em diversas publicações das áreas da Comunicação, da Ciência da Informação e mesmo dos Estudos da linguagem, como ilustrado abaixo.



Figura 1 – Modelo E→M→R de Comunicação

Adaptado de Wolf (2003).

Semelhante simplificação promoveu a neutralização dos problemas da linguagem, embora esse aspecto esteja implícito no modelo de Shannon e Weaver (1949), como mostra a Figura 2.

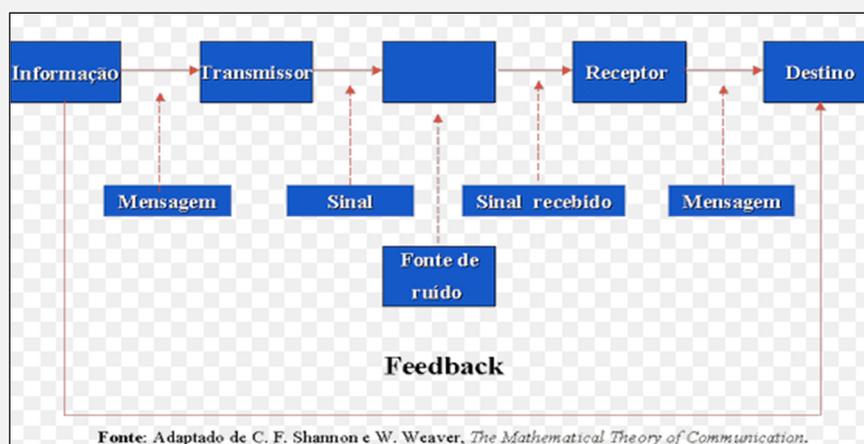


Figura 2 – Modelo de Sistema de Comunicação de Shannon e Weaver (1949)

As primeiras interpretações do modelo enfatizavam a relação Transmissor e Receptor, processo passível de cálculo. A Teoria da informação é, nessa perspectiva, um método para calcular a transmissão ou transferência de sinais, porém, pouco adequado para calcular significados (ECO, 1972).

O modelo de Shannon e Weaver (1949) pode ser decomposto: informação, mensagem, sinal, canal, receptor, destino, ruído, *feedback*. Em primeiro lugar, é possível observar a distinção entre Informação, Mensagem e Sinal, entre Informação e Transmissor e entre Receptor e Destinatário. Segundo Eco

(1972), a proposta modela um sistema comunicacional complexo no qual está presente uma fonte de informação que elabora Informação ou **Mensagem**; esta última é codificada em sinais por meio de um aparelho **Transmissor**; a mensagem codificada viaja por um **Canal** que pode ser fonte de **Ruídos**. Após sair do canal, o sinal é recebido por um **Receptor**, que o converte novamente em **Mensagem** compreensível pelo **Destinatário**. Para contornar possíveis perturbações ou ruídos no sistema, Shannon previa a necessidade de codificação perfeita ou seja, um modo econômico, veloz e seguro de elaborar mensagens por meio de linguagem específica de transmissão e recepção (ECO, 1972, p. 10)

Na interpretação de Eco (1972), os problemas do significado se encontram na relação Informação-Mensagem-Destinatário, relação que comporta, necessariamente, as questões da linguagem e da significação. Nesse sentido, as mensagens para serem produzidas e compreendidas, requerem um sistema de significação compartilhado. No entanto, esse esquema foi adotado para justificar epistemologicamente a formalização e a quantificação das relações de interdependência entre os constituintes do processo informacional-comunicacional, quer entre máquinas, quer entre seres humanos (ABRIL 1997), sem problematizar as questões das significação.

A significação coloca desafios ainda intransponíveis às tentativas de cálculo. A pragmática torna mais desafiadora a questão, ao adicionar a noção de contexto à interpretação. São os problemas relativos à recepção/interpretação de mensagens – portanto a linguagem- que torna o modelo significativo e problemático, ao mesmo tempo, para as pesquisas em Ciência da Informação e Comunicação. É o problema da significação que se desenvolveu no interior das Ciências da Informação e Comunicação, aglutinadas na tríade informação-comunicação-linguagem.

No presente trabalho, recortamos, os conceitos relativos essa tríade, de itens selecionados da literatura produzida no campo da Ciência da Informação e Comunicação, com especial destaque para a produção francesa.

A análise da literatura francesa sobre o tema justifica-se pela longa tradição de intercâmbio científico entre Brasil e França, no campo das Ciências Humanas e Sociais. Com efeito, às instituições francesas de ensino e pesquisa ainda hoje se dirige significativo contingente de pesquisadores para obter formação pós-graduada. Tem colaborado para dar continuidade a essa tradição a Rede MUSSI (Rede Franco-Brasileira de Pesquisadores em Mediações e Usos Sociais de Saberes e Informação), criada em 2004 sob a coordenação de Viviane Couzinet e Regina Marteleto (COUZINET e MARTELETO, 2010).

Pode-se acrescentar ainda outra justificativa: no Brasil, as pesquisas sobre o par informação-comunicação-linguagem têm lugar, majoritariamente em Programas de pós-graduação em Ciência da Informação, avaliados dentro da área Comunicação e Informação da Capes e CNPq. Assemelha-se, dessa forma à avaliação institucional da pesquisa científica sobre o campo, realizada na França.

A apresentação do estudo segue a seguinte sistemática: no item 1 apresenta-se a pesquisa, seus objetivos e procedimentos metodológicos; no item 2 discute-se o marco teórico que justifica a abordagem da tríade conceitual informação-comunicação-linguagem; no item 3 apresenta-se o corpus empírico que caracteriza as pesquisas sobre informação-comunicação-linguagem e no item são tecidas as considerações finais os resultados da investigação.

2. INFORMAÇÃO-COMUNICAÇÃO-LINGUAGEM.

Informação e comunicação, tanto quanto outros conceitos a eles associados, tais como significação, representação, recuperação de informação, recepção, estão presentes nos processos de compartilhamento de saberes, porém, nem sempre definidos de forma explícita. Três áreas - Ciências da Comunicação, Ciência da Informação e Ciências da Computação -, ao menos, reconhecem que defini-los não é tarefa trivial, em particular para operacionaliza-los. São estes problemas que motivaram o desenvolvimento da presente reflexão.

A associação do termo linguagem à dupla comunicação-informação deve-se à importância atribuída à significação nas reflexões sobre os processos infocomunicacionais. Pode-se, desse modo, acoplá-los para constituir a tríade informação-comunicação-linguagem como conceitos solidários na Ciência da Informação e Comunicação, em particular no subcampo da Organização do Conhecimento (OC).

A Organização do conhecimento tem estabelecido diálogos com diversas áreas que se preocupam com a elaboração de artefatos sociais de compartilhamento de conhecimentos. A Museologia e a Arquivística, por exemplo, têm adotado diversos conceitos e teorias tanto da OC quanto da Comunicação para dar sustentação aos métodos de elaboração de seus instrumentos de trabalho. Por outro lado, as atividades de construção e disponibilização de memórias sociais, como os diferentes tipos de repositórios têm aproximado essas áreas das Ciências da Computação.

Convém lembrar, porém, que disponibilizar não é sinônimo de comunicar. Para esta função é imprescindível compreender as distinções semióticas entre sinal e signo. A compreensão mais ampla das questões sobre a linguagem, tem concorrido para o estabelecimento de diálogos mais profícuos e pesquisas colaborativas entre a Computação, a Ciência da Informação e a Comunicação. O conceito de Web semântica é, certamente, sinal de incorporação da linguagem pelo campo da Computação.

3. PESQUISAS SOBRE INFORMAÇÃO-COMUNICAÇÃO-LINGUAGEM.

A natureza polissêmica dos termos informação e comunicação tem gerado discussões fecundas. Nos Estados Unidos da América, as pesquisas realizadas em instituições de Information and Communication Studies privilegiam o par informação-comunicação. Pode-se afirmar que o eixo nucleador das pesquisas realizadas nesse contexto são os problemas da Recuperação da Informação (RI), tradição que se consolidou desde os experimentos de Vannevar Bush (1945). As questões sobre a linguagem são raras nessas pesquisas. Contudo, podem ser identificados estudos que foram além das abordagens funcionalistas dominantes na área de RI. Borko (1969), por exemplo já questionava as questões semânticas e comunicacionais da documentação, como pode ser visto no artigo ‘Subject analysis from a communication point of view’.

As relações entre a Computação e a Ciência da Informação já foram destacadas por Saracevic (1995), que afirma que as principais relações interdisciplinares da C.I. ocorrem com a Ciência da Computação, a Ciência Cognitiva e a Comunicação. Sobre a Comunicação, o autor destaca: “Se há uma palavra com mais conotações, mais usos nos mais diferentes contextos, mais motivo para confusão do que a palavra ‘Informação’, a palavra é ‘Comunicação’”. (Saracevic, 1996, p. 39). Prossegue o autor:

The evolving relation between information science and communication has a number of dimensions: a shared interest in human communication, together with an increase in realization that

information and communication need to be studied together, a confluence or certain streams of research, some exchanges of faculty, and a potential for professional cooperation. (SARACEVIC, 1996, p. 39)

No registro mais contemporâneo, podem ser destacadas as contribuições de Day (2014), Buckland (1991a, 1991b) e Frohmann (2004), importantes pontos de clivagem nas pesquisas norte-americanas da Ciência da informação. Foram por eles lançados olhares críticos sobre a abordagem funcionalista das pesquisas de RI, fato que introduziu importantes modificações no panorama teórico dos estudos da informação. Esses autores, com base em teorias foucaultianas, revisitam as contribuições de Otlet e Briet (DAY, 2014), como também sobre a relação entre informação e documento (BUCKLAND, 1991a, 1991b; FROHMANN, 2004), questões já tematizadas anteriormente por Meyriat e Briet.

Na França, Escarpit (1976,197), Meyriat (1983), Couzinet (2000), Miège (2004, 2014), Bougnoux (1999), Jeanneret (2011), entre outros, são responsáveis pelas pesquisas e formação de quadros científicos e profissionais em cursos e programas de Ciências da Informação e da Comunicação (CIC). Trata-se de área com conformação particular, na qual os programas de Comunicação e de Ciência da Informação compartilham um mesmo espaço de avaliação, como ocorre no Brasil.

A complexidade do conceito de informação e a necessidade de adotar um ponto de vista adequado para abordá-lo é discutido por Couzinet e Marteleto (2010, p. 22):

A tripla acepção da noção de informação (dados: *data*, notícias: *news*, conhecimentos: *knowledge*) fornece um primeiro quadro analítico que permite ordenar esses discursos [informação e seus discursos]. Seja ela especializada (científica e técnica) ou midiática, a informação se desdobra no espaço das ciências da informação e da comunicação, assim como em sua periferia. Uma segunda abordagem oferece a alternativa de focalizar de forma mais precisa os discursos da informação (considerada como *significante*) e os discursos sobre a informação (apreendida como *significado*). Trata-se assim de contribuir para a definição dessa noção complexa a partir de sua ancoragem teórica ou de suas aplicações documentárias.

Em texto conhecido, Bougnoux (1999, p.13-14), confessa a dificuldade de estabelecer distinção entre informação e comunicação. Propõe, assim, colocá-los sob tensão, “numa dialética fecunda para as Ciências da Informação e da Comunicação” (BOUGNOUX, 1999, p. 125).

Tal tensão é considerada necessária no interior das Ciências da Informação e Comunicação, por serem campos que examinam as condições práticas (instrumental midiático, institucional e simbólico) da produção e circulação de informações. Em sua concepção, Ciências da Informação e Comunicação são uma “disciplina desconfortável”, que circula criticamente entre distintos saberes, em face da amplitude das questões a que se dedica (BOUGNOUX, 1999, p.14). Quanto à institucionalização da disciplina, o autor observa que:

As CIC [Ciências da Informação e da Comunicação] respondem a uma exigência pedagógica e teórica. Nasceram nas universidades, do desejo de adaptar seus cursos a perspectivas inéditas e à rápida expansão de novas profissões; no campo intelectual, a disciplina surgiu de uma interrogação antropológica sobre a redefinição da cultura, identificada com as diferentes maneiras de comunicar e, de início, centrada, nos anos sessenta, na troca e na formalização linguísticas (com as pesquisas ‘estruturalistas’ de Lévi-Strauss, Barthes ou Jakobson. (BOUGNOUX,1999, p. 13-14)

Para Bougnoux (1999), comunicar supõe sempre dois níveis: primeiramente, mensagens-quadros, e sobre esta base, a construção e compreensão de conteúdo, ou informação propriamente dita. Bougnoux

(1999) reafirma, assim, a ideia de que a comunicação requer códigos. O código é, de fato, um sistema que permite elaborar mensagens e filtrá-las.

Se definirmos, portanto, nossos fenômenos de comunicação como a esfera das atividades pragmáticas de tratamento de **mensagens entre sujeitos**, vê-se que um dos critérios de reconhecimento dessas ações reside em seu *fracasso* sempre possível (BOUGNOUX, 1999, p. 17).

Em um sistema de informação/comunicação convivem a esfera semiótica, responsável pela circulação de conteúdos, e a esfera propriamente técnica, de transmissão de sinais. É nesse contexto que Bougnoux se pergunta: “É preciso opor comunicação a informação?” (BOUGNOUX, 1999, p. 125). O autor responde a essa interpelação definindo conceitos e suas relações de interdependência, como segue: “[...] a informação vale e mede-se no campo do conhecimento, e a comunicação no campo da ação e da organização. Desta divisão decorre que a segunda precede e condiciona necessariamente a primeira (BOUGNOUX, 1999, p. 125-126).

O reconhecimento das diferenças entre informação e comunicação é crucial também para o campo da Organização do conhecimento. Com efeito, o termo informação é definido, de forma geral como representação de conteúdos de documentos, construídos segundo códigos e gramáticas específicas.

A codificação é um processo de representação necessário para reduzir ruídos. Por meio desses filtros (códigos), procura-se garantir a pertinência da informação. Decorre, daí, a ideia de que a informação pede tratamento específico para ser comunicável, sendo o par informação/comunicação uma dualidade que remete tanto ao conteúdo quanto à relação ((BOUGNOUX, 1999).

Na discussão em pauta, vale lembrar Peirce (1978): para conhecer é necessário reconhecer códigos. Na visão peirceana, o interpretante designa o código, isto é, a convenção que permite relacionar um signo a um objeto. As noções de código e signo são, assim, fundamentais nas tarefas de construir linguagens de codificação de informações. Em outros termos, a indexação de informações requer um sistema de codificação específico, compartilhado pelo sistema e usuários. São esses sistemas, denominados de forma genéricas como Linguagens documentárias ou Sistemas de Organização do Conhecimento, que subsumem tipos específicos: sistemas de classificação, tesouros, taxonomias e ontologias. Esses sistemas as entidades e as relações entre entidades significam segundo o quadro de referência considerado, muitas vezes caracterizadas como compromissos ontológicos (CAMPOS, L. M.; CAMPOS, M. L. A.; CAMPOS, M. L. M. (2010). Nesse sentido, é possível associar o conceito de ‘compromisso ontológico’ ao interpretante peirceano. Pode-se, da mesma forma, aproximar as operações de codificação e decodificação de informações do conceito de jogos de linguagem (WITTGENSTEIN, 1988), questão que Bougnoux (1999) enuncia como segue:

Se nossa apreensão da informação morre por excesso de desordem ou de ruído, observa-se, inversamente que ela se extingue na redundância ou previsibilidade pura: uma repetição, como tal, não traz informação. [...] O grande jogo da informação desenvolve-se, portanto, entre obstáculos. O jogo é certamente um modelo inspirador, bem apreendido por Wittgenstein, ao afirmar que a significação é construída por meios de jogos de linguagem. Há, portanto, regras estáveis que governam a produção e a recepção de mensagens (conteúdos), contexto no qual as jogadas podem ser imprevisíveis, variáveis (BOUGNOUX, 1999, p. 139).

As relações entre informação-comunicação e linguagem são apresentadas, também, por Jeanneret (2011). Seu texto problematiza o contexto informacional contemporâneo, mediado por dispositivos de comunicação de informações, observando que a “pressão profissional e a inquietude dos estudantes

tendem a conceber as ciências da documentação como uma pura engenharia de redes informáticas (JEANNERET, 2011, p. 15). Para evitar tal redução, propõe que [...] “as relações entre os dispositivos técnicos e as práticas sociais de informação, ou de compartilhamento do saber, poderão ser definidas no interior das questões mais amplas da cultura” (JEANNERET, 2011, P. 17). Trata-se de olhar os dispositivos tecnológicos de difusão de mensagens como artefatos culturais de inscrição e memorização de objetos culturais, sujeitos a inúmeras mediações.

As mídias informatizadas são dispositivos técnicos não apenas de memorização. Sua importância reside também na capacidade de perenizar o acesso à informação por meio da neutralização do espaço e do tempo. São dispositivos complexos construídos para a circulação social de informações. A circulação e a troca de informações, devidamente compreendidas e interpretadas, permitem produzir novos conhecimentos. (JEANNERET, 2011).

Como observa o autor, as inscrições e processos não são neutros:

Não, há, nesse sentido, representação de informação em si, mas somente para o outro, porque a informação é uma relação que se estabelece entre um objeto e um olhar; a constituição de um documento ou de um conjunto de documentos não é mais do que pura representação do mundo. [...] São proposições ou mais exatamente uma implicação de comunicação e de leitura que adota um ponto de vista, que procede de reescritura e adaptação. (JEANNERET, 2011, p. 87).

A organização da informação, por suas vez, não é mera operação de tratamento de sinais. Nesse processo, é necessário olhar analiticamente os documentos para, em seguida, construir representações contextualizadas. Estas últimas – as informações documentárias – inscrevem-se, portanto, em um regime informacional em que informação é conteúdo e comunicação é processo de relacionar informações e sujeitos em contextos.

Na discussão sobre a codificação e decodificação de informações é imprescindível considerar os sistemas de organização do conhecimento, ou linguagens documentárias. Estas últimas são artefatos pragmáticos de mediação de informações. Jakobson (1975) enunciou, da perspectiva estruturalista, o que podemos, de um certo modo, considerar como os aspectos pragmáticos da linguagem, isto é, as suas principais funções: informativa (referencial), emotiva, conativa, fática, poética, metalinguística. Nesse modelo, as linguagens documentárias exercem função informativa, mais especificamente de codificação de mensagens documentárias que remetem a um referente. Como todo código, as linguagens documentárias promovem efeitos de sentido, que precisam ser previstos para promover interação adequada.

A codificação de informações, como afirma Escarpit (1976, p. 159), “deve dispor de dois léxicos: um léxico geral que corresponde ao estado da língua em seu tempo e um léxico especializado que corresponda à prática que serviu à produção do documento de partida”. Além disso,

Para ser utilizável em computadores, esses léxicos devem tomar a forma do que chamamos tesouros. O tesouro é uma lista de palavras-chave [descritores, na terminologia contemporânea], isto é, palavras com alto grau informativo e pertinência estreita com assuntos, como os índices que figuram ao final de livros. Mas esta lista deve ser completada com a adição de outras palavras que possuam um certo número de relações com as palavras-chave: sinonímia, paronímia, analogia, inclusão antonímia, condição, etc. Um tesouro, assim compreendido é, portanto, um excelente instrumento de reescritura porque ele pode ser inteiramente codificado e indefinidamente aumentado pela inclusão de novas palavras-chave e de novas relações (ESCARPIT, 1976, p. 159).

As noções de código e de reescritura, acima expostas, mostram a complexidade das linguagens documentárias: elas devem comportar dois sistemas de significação compatibilizados para significar de modo específico. Infere-se daí, que as linguagens documentárias codificam a partir de um quadro geral estabelecido previamente. A noção de interpretante, de Peirce, como se pode ver, ressurgiu sempre como questão básica na discussão sobre os códigos e a codificação.

Ainda no contexto da codificação, decodificação e recuperação de informações, ganham importância crescente os objetos denominados ontologias. Uma ontologia, na definição clássica de Gruber (1995) é uma especificação explícita de uma conceitualização compartilhada. Infere-se, desta definição: a) que as ontologias são compostas de conceitos submetidos a definições; b) os conceitos derivam de conhecimento consensual construído em contextos pragmáticos. Na perspectiva semiótica, portanto, uma ontologia é um código com fins informacionais pragmáticos. O significado de cada conceito é submetido a restrições explícitas comuns, tanto aos construtores de ontologias quanto aos usuários, que fazem uso das ontologias para codificar, recuperar e descobrir conhecimentos.

As ontologias apresentam, além dos conceitos, uma estrutura relacional expressa por operadores lógicos (inclusão hierárquica, partitiva, sequencial, entre outros). Tanto os conceitos quanto as relações têm seu valor de verdade condicionado pelo sistema de significação do campo de conhecimento, ou domínio, considerado.

Portanto, uma ontologia pode ser olhada como um artefato informacional e comunicacional, no qual a informação é conteúdo composto de conceitos e relações entre conceitos logicamente formalizados por meio de outro sistema de significação - uma linguagem lógica. Nessa medida, as terminologias, são fontes imprescindíveis de conceitos e relações entre conceitos para a construção de ontologias. Dito de outro modo, as entidades das ontologias não significam individualmente. Elas requerem conceitos definidos consensualmente, que podem ter como fontes privilegiadas as terminologias de domínios.

Do que foi dito, infere-se que os tesouros e as ontologias, por serem códigos, podem compartilhar teorias e procedimentos. Nessa medida, a aproximação mais estreita entre a Organização do conhecimento, do campo da CI, e a criação de ontologias, do campo da Computação, pode trazer benefícios mútuos.

4. CONSIDERAÇÕES FINAIS

Apresentamos aqui um estudo exploratório sobre as questões informacionais, comunicacionais e linguísticas no interior das Ciências da Informação e Comunicação. Semelhante incursão foi realizada pela importância dessas discussões para pensar/repensar as teorias e métodos da Organização da Informação e do Conhecimento. Como constatou Saracevic (1996) a ambiguidade das palavras informação e comunicação torna necessária a sua explicitação, em cada contexto.

Embora sejam distintos os objetos de pesquisa das Ciências da Comunicação e da Ciência da informação, ambas têm refletido sobre os conceitos informação-comunicação-linguagem. Esses conceitos são igualmente importantes para as áreas da Ciência da Computação para conceber e construir artefatos infocomunicacionais como as bases de dados e as ontologias.

A análise empreendida indica que, nos estudos da organização do conhecimento, é necessário aprofundar as reflexões sobre a linguagem, dando atenção especial aos complexos problemas da

codificação de mensagens. De fato, diversos aspectos da tríade informação-comunicação-linguagem podem ser melhor compreendidos à luz das teorias semióticas de codificação.

A compreensão das questões da significação pode promover diálogos produtivos com as diversas áreas que lidam com o armazenamento, o tratamento e a comunicação de informações. Com efeito, informação, comunicação e linguagem são conceitos que atravessam as áreas que organizam informação para recuperar. A construção de bases de conhecimentos, de ontologias, de planos de classificação, de catálogos de pesquisa e a documentação museológica não prescindem da compreensão desses conceitos.

As pesquisas sobre a relação informação-comunicação-linguagem, foram desenvolvidas, de forma bastante visível, na França, pela forte influência da Semiologia de Saussure (1973) e da Semiótica francesa (GREIMAS, 1976). Esta última, elucidou inúmeros aspectos das ações discursivas, tendo proposto importantes métodos para realizar operações analíticas de desvelamento da enunciação e recepção de informações. Tais teorias foram fortemente incorporadas às pesquisas realizadas nas Ciências da Informação e Comunicação, na França. Este fato motivou a escolha do *corpus* para desenvolver a presente pesquisa.

Esperamos, que o presente texto, suscite debates que possam contribuir para aprofundar a reflexão crítica e de ação no campo da Organização do Conhecimento, como também dar densidade ao diálogo com as áreas afins, dedicadas à produção, circulação e recepção de informações e conhecimentos.

Em tempos de uso ampliado das mídias informáticas, é necessário que a formação de pesquisadores e profissionais da informação esteja ancorada na compreensão aprofundada dos problemas de pesquisa próprios da Ciência da Informação. Parece ser imprescindível que os formadores compreendam que as operações concretas de disponibilização da informação não são meras questões técnicas. Ao contrário, as atividades informacionais são práticas sociais de compartilhamento de saberes que requerem aportes teóricos consolidados, muitos deles já explorados no campo das Ciências da Informação e da Comunicação. É importante destacar esta questão. Ela remete diretamente à discussão dos currículos de algumas instituições brasileiras de formação de bibliotecários. Há, no quadro docente desses cursos, profissionais que, por desconhecerem a abrangência e a profundidade das questões da documentação, reduzem o conteúdo de suas disciplinas a meras questões técnicas do campo da informática. Essa visão estreita e equivocada, tem resultado em propostas de extinção de campos de saber como a Biblioteconomia e a Ciência da Informação, na ilusão de que as atividades humanas de tratamento de informação e comunicação de saberes possam ser inteiramente realizadas por algoritmos.

Acompanhamos Jeanneret, na convicção de que é urgente combater as concepções tecnocráticas. Pode-se combatê-las evidenciando as complexas relações entre informação-comunicação e linguagem no funcionamento dos dispositivos informacionais. Mesmo porque os algoritmos são produtos do engenho humano. Sua eficácia não é dependente apenas do aprimoramento das ferramentas informáticas. Ao contrário, os algoritmos dependem da compreensão fina dos processos comunicacionais mediados pela linguagem.

Como afirmaram os autores aqui apresentados, a informação vale e mede-se no campo do conhecimento, e a comunicação no campo da ação e da organização. Desta divisão decorre que a segunda precede e condiciona necessariamente a primeira. Inscrevem-se, ambas, em um regime informacional em que informação é conteúdo e comunicação é o processo de relacionar informações e sujeitos em contextos.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- ABRIL, G. (1997) **Teoría general de la información**. Madri: Catedra.
- BORKO, H. (1969) **Subject analysis from a communication point of view**. Paper presented at the American Library Association, Atlantic City, N.J.
- BOUGNOUX, Daniel. 1(999) **Introdução às Ciências da Comunicação**. Bauru, EDUSC.
- BUCKLAND, M. (1991) **Information and information systems**. New York: Praeger.
- BUCKLAND, M. (1991) Informations as thing. **Journal of the Americam Society of Information Science**, v. 42, n.5, p.351-360.
- BUCKLAND, M. (1997) What is a document. **Journal of the Americam Society of Information Science**, v. 48, n. 9, p. 804-809.
- CAMPOS, L. M. ; CAMPOS, M. L. A. ; CAMPOS, M. L. M. (2010) Diretrizes para a definição de domínio no reuso de ontologias biomédicas: uma abordagem baseada na análise do compromisso ontológico. In: XI Encontro Nacional de Pesquisa em Ciência da Informação. **Anais eletrônicos**. ANCIB, Rio de Janeiro. Organização e Representação do Conhecimento, **CD-ROM**
- CAMPOS, M. L. A. (2010) O papel das definições na pesquisa em ontologia. **Perspectivas em Ciência da Informação**, v. 15, p. 220-238.
- COUZINET, Viviane, MARTELETO, Regina. (2010)Médiations documentaires: entre réalités et imaginaires. **Actes de la première journée scientifique internationale du réseau**. MUSSI. Toulouse: Université de Toulouse, p. 15-25.
- ESCARPIT, Robert. (1978) **L'écrit et la communication**. Paris: PUF.
- ESCARPIT, Robert. (1976) **Théorie générale de l'information et de la communication**. Paris: Hachette.
- FROHMANN, B. (2004) Documentation redux: prolegomenon to (Another) Philosophy of Information. **Library Trends**, v. 52, n. 3, p. 387-407.
- GRUBER. T.R. (1993) A Translation Approach to Portable Ontologies. **Knowledge Acquisition**, v.5, n. 2, p.199–220.
- GRUBER,T.R. (1995) Toward Principles for the Design of Ontologies Used for Knowledge Sharing. **International Journal of Human Computer Studies**, v.43, v.5–6, p.907–928.
- GUARINO, N. ((2001) Formal ontology in information systems. (2001) In: **Proceedings of the internat. conference on formal ontology in information systems**, 2001, Buffalo. Nova York University at Buffalo.
- GUARINO, N.; GIARETTA, P. (1995) Ontologies and Knowledge Bases: Towards a Terminological Clarification. In N. Mars (ed.) **Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing 1995**. IOS Press, Amsterdam: 25-32.

JEANNERET, Yves (2011). **Y-a-t-il (vraiment) des Technologies de l'information?** Villeneuve d'Ascq.

MEIRYAT, Jean. (1983) Por une classification des sciences de l'information et de la communication. **Schéma et schématisation**. N. 19, p. 61-64.

MIÉGE, Bernard. (2014) A circulação do conhecimento e a construção das CICs (Ciências da informação e comunicação na França. Questões Transversais: **Revista de epistemologia da Comunicação**. v. 2, n.4, jul/dez.

MIÉGE, Bernard. (2004) **L'information-communication, objet de connaissance**. Bruxelles: De Boeck.

MIÉGE, Bernard. (1995) **La pensée communicationnelle**. Grenoble: PUG, 1995.

PEIRCE, Charles. (1978) **Écrits sur le signe**. Paris: Seuil.

SARACEVIC, Tefko. (1996) Ciência da Informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**. v. 1, n. 1, p. 41-62, jan./jun. 1996.

WITTGENSTEIN, Ludwig. (1988) **Investigaciones filosóficas**. Barcelona: Grijalbo.



ANÁLISE TEÓRICA CONCEITUAL DA TERMINOLOGIA, GARANTIA CULTURAL, MULTICULTURALIDADE E TRANSCULTURALIDADE NA CONSTRUÇÃO DE LINGUAGENS DOCUMENTÁRIAS

Ricardo Biscalchin¹, Érica Fernanda Vitorini², Walter Moreira³

¹Doutorando - Universidade Estadual Paulista - Programa de Pós-Graduação em Ciência da Informação,
ricardo_biscalchin@hotmail.com

²Doutoranda - Universidade Estadual Paulista - Programa de Pós-Graduação em Ciência da Informação,
erica.vitorini@marilia.unesp.br

³Professor Doutor - Universidade Estadual Paulista - Programa de Pós-Graduação em Ciência da
Informação, *walter.moreira@marilia.unesp.br*

RESUMO Este artigo aborda a organização do conhecimento para representação e disseminação da informação segundo as perspectivas da Terminologia, em particular com a Teoria Comunicativa da Terminologia de Cabré, e dos conceitos da garantia cultural, da multiculturalidade e da transculturalidade para auxílio na construção de linguagens documentárias. O objetivo deste artigo é demonstrar a importância da aplicação dessas teorias na construção de linguagens documentárias para contemplar os usuários de acordo com o contexto sócio-histórico-cultural que se encontram. A metodologia aplicada consiste na análise teórica conceitual das temáticas Terminologia, Multiculturalidade, Transculturalidade, Garantia Cultural e Linguagem Documentária, definindo cada uma delas. Os resultados obtidos por meio da análise teórica conceitual indicam que o uso da Teoria Comunicativa da Terminologia, da multiculturalidade, da garantia cultural e da transculturalidade, auxilia na construção de linguagens documentárias possibilitando uma representação e recuperação da informação mais próxima a realidade sócio-histórico-cultural dos usuários, justificando sua aplicabilidade no momento da representação e organização da informação. Conclui-se que a aplicação desses conceitos como auxílio na construção de linguagens documentárias, possibilita que minorias sociais, raciais e culturais sejam representadas de maneira igualitária em relação à cultura da grande massa.

PALAVRAS-CHAVE *Linguagem Documentária, Teoria Comunicativa da Terminologia, Multiculturalidade, Transculturalidade, Garantia Cultural.*

ABSTRACT This paper deals with the organization of knowledge for the representation and dissemination of information according to the perspectives of Terminology, in particular with the Communicative Theory of Terminology by Cabré, and the concepts of cultural warranty, multiculturality and transculturality to aid in the construction of documentary languages. The purpose of this article is to demonstrate the importance of the application of these theories in the construction of documentary languages to contemplate the users according to the socio-historical-cultural context that they are. The applied methodology consists of the theoretical conceptual analysis of the Terminology, Multiculturality, Transculturality, Cultural Warranty and Documentary Language, defining each one of them. The results obtained through the theoretical conceptual analysis indicate that the use of the Communicative Theory of Terminology, multiculturality, cultural warranty and transculturality, helps in the construction of documentary languages, enabling a representation and retrieval of information that is closer to socio-historical-cultural context of the users, justifying their

applicability when representing and organizing information. It is concluded that the application of these concepts as an aid in the construction of documentary languages allows social, racial and cultural minorities to be represented in an egalitarian way in relation to the culture of the great mass.

KEYWORDS *Documentary Language, Communicative Theory of Terminology, Multiculturalism, Transculturality, Cultural Warranty.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Considerando-se a grande quantidade de informações produzidas diariamente em diferentes meios, suportes e culturas, a Organização do Conhecimento, enquanto campo de pesquisa teórica e aplicada dedica-se a estudos voltados para a organização e representação desse conhecimento produzido pela sociedade. Neste sentido, a organização e a representação do conhecimento requer diálogos com a Terminologia, a qual “[...] estuda, teoricamente, os termos e seus respectivos conceitos, os sistemas de conceitos e sua representação” (Boccatto, 2005) e que possui entre suas vertentes a Teoria Comunicativa, conforme sua apresentação em Cabré (1993). Considerando-se ainda a diversidade e a pluralidade de objetos e suportes de informação, de usuários, de idiomas, de contextos documentários e de contextos culturais, a organização e a representação do conhecimento demandam discussões que envolvam as premissas da Garantia Cultural, da Multiculturalidade e da Transculturalidade. A observância de tais fundamentos auxiliará na construção de instrumentos de representação e recuperação da informação mais precisos, e ao mesmo tempo, mais flexíveis.

Nesse artigo apresenta-se uma análise teórico-conceitual relativamente aos conceitos de “garantia cultural”, “multiculturalidade” e “transculturalidade” visando demonstrar a importância de sua aplicação integrada para a construção de linguagens documentárias representativas do contexto sócio-histórico-cultural de seus usuários. A discussão apresentada orienta-se também por um dos princípios fundamentais da Organização do Conhecimento que diz respeito à disseminação da informação a todos os usuários de maneira ética e sem distinção.

As discussões sobre terminologia, garantia cultural, transculturalidade e multiculturalidade são importantes para a área de organização do conhecimento. Dentre um conjunto de autores, podemos destacar os trabalhos de Cabré (1993, 1999, 2003) relativos à Terminologia, Begthol (2002) e Barité (2010) que abordam a garantia cultural e a garantia literária, Nicolescu (2005), Neiva, Alonso e Ferneda (2007), Lucchesi e Malanga (2011) e García Gutierrez (2014) que abordam a transculturalidade e a transdisciplinaridade e Lazzarin (2008) que aborda a multiculturalidade.

METODOLOGIA

A metodologia utilizada nessa pesquisa consistiu de uma revisão teórica conceitual de Terminologia, com enfoque na Teoria Comunicativa da Terminologia, idealizada por Maria Teresa Cabré, na Garantia Cultural, na Multiculturalidade, na Transculturalidade e nos conceitos e normas para a construção de Linguagem Documentária, definindo cada uma delas conceitualmente.

Para esse levantamento teórico conceitual, foram realizadas buscas nas principais bases de dados em Ciência da Informação, utilizando como critério de busca, os termos: terminologia, garantia cultural, multiculturalidade, transculturalidade, linguagem documentária e vocabulário controlado, além de seus equivalentes em língua inglesa e espanhola.

Após o levantamento teórico conceitual, foi realizada a leitura dos textos buscando a intersecção dos conceitos. Por meio dessa análise, buscamos apresentar a importância e a potencialidade do uso desse conjunto de conceitos para atender critérios que contemplem distintos usuários em diferentes realidades sócio-histórico-culturais no momento da recolha terminológica, objetivando auxiliar na construção de linguagens documentárias que atendam ao mais amplo espectro de usuários.

A metodologia busca por meio da intersecção, apresentar como resultado preceitos que possibilitem a linguagem documentária compreender o contexto e a realidade cultural de seus usuários reais ou potenciais, realidade essa que impacta diretamente em suas necessidades de busca.

RESULTADOS E DISCUSSÃO

A Terminologia em conjunto com a Ciência da Informação, serve como instrumento colaborativo na representação e disseminação do conhecimento. A terminologia segundo a norma ISO (1087, 2000, p.2) é "a ciência que estuda a estrutura, a formação, o desenvolvimento, o uso e a gestão das terminologias nos diferentes domínios." Por meio dos recursos metodológicos da Terminologia é possível descrever e identificar os termos e os conceitos que lhes são inerentes e aplicá-los na organização, representação e disseminação da informação.

A Terminologia possui distintas abordagens, sendo elas a Teoria Geral da Terminologia, a Teoria Comunicativa da Terminologia, a Socioterminologia e a Teoria Sociocognitiva da Terminologia. Krieger (2005, p.1) define Terminologia como

[...] uma área de conhecimentos e de práticas, cujo principal objeto de estudos teóricos e aplicados são os termos técnico-científicos. Em sua face teórica, a Terminologia ocupa-se da descrição da gênese e dos modos de constituição e funcionamento das unidades lexicais especializadas. Estas são assim denominadas porque se constituem e são utilizadas no âmbito de atividades profissionais especializadas, cumprindo a missão de veicularem conceitos próprios de cada área do conhecimento.

No plano aplicado são desenvolvidos, basicamente, trabalhos de produção de glossários de diferentes áreas temáticas, de dicionários especializados e de bancos de dados de terminologias.

Bocato (2005, p.51) afirma que "[...] a Terminologia estuda, teoricamente, os termos e seus respectivos conceitos, os sistemas de conceitos e sua representação". Assim podemos dizer que a terminologia atua nas questões lógico-semânticas, dando consistência na relação entre termos da linguagem natural e da linguagem de especialidade.

Cabré (1993) afirma que os termos representam os conceitos próprios de cada disciplina especializada. Ou seja, os termos representam o conceito inerente ao contexto onde se encontram inseridos. Em relação à interdisciplinaridade Cervantes (2004, p.56) afirma que

[...] tanto a Terminologia como a Ciência da Informação têm como finalidade atender os usuários em suas necessidades de informação. Desse modo, as duas ciências possuem diversos instrumentos comuns ou similares para a realização de suas atividades, tais como os dicionários, glossários, enciclopédias, sistemas de classificação, tesouros, normas, fontes literárias, as informações dos profissionais e especialistas da área, entre outros.

Dentre as vertentes da Terminologia, considera-se mais adequada aos propósitos desta investigação a Teoria Comunicativa da Terminologia (TCT), que dá ênfase na linguagem e seus aspectos comunicativos, para fins de disseminação da informação, não fazendo uso de termos fixos, possibilitando que estes sejam modificados conforme o contexto em que se apresentam. Essa possibilidade de alteração dos termos para representar os conceitos é essencial para garantir os preceitos culturais na representação da informação.

Termo, de acordo com a ISO 1087 (2000, p.6) é "a designação verbal de um conceito geral em um domínio específico". O contexto, é possível depreender, é determinante na escolha do termo e a relação de representação que estabelece com o conceito. Na fase de construção de representações documentárias, seja na indexação ou na busca, é extremamente importante a escolha do termo coerente e autorizado pela área de conhecimento que envolve a representação. Desse modo é possível aproximar-se mais da garantia de que o termo expresse o conceito de maneira coerente e compreensível ao conjunto de usuários da terminologia.

Assim, uma unidade lexical é geral por padrão e adquire significado especial ou terminológico quando ativado pelas características pragmáticas do discurso (Cabré, 2003).

Toda unidade lexical pode se tornar uma unidade terminológica dependendo do discurso onde se encontra inserida. Ferini (2006, p.30) apresenta que,

[...] os termos não pertencem a um âmbito específico, mas são utilizados nesse âmbito, o que justifica seu caráter especializado. Dessa forma não existiria, a priori, nem palavras nem termos. É o contexto de uso que vai atribuir a uma determinada unidade linguística dotada de referência essa função.

O contexto cultural e a área de especialidade onde os termos se apresentam, são essenciais para seu uso, pois interferem diretamente no sentido, no conceito representado por esse termo, ou seja, a representação do conceito ocorre por meio do termo, que adquire significado conforme o contexto em que é aplicado. Distintos contextos podem implicar no uso de distintos termos para a representação do mesmo conceito, cabendo nesse caso a análise de qual é o contexto em que se encontra inserido o vocabulário, para que sua representação atenda as expectativas dos usuários.

A escolha da TCT se justifica, pois a escolha dos termos é variável conforme o contexto e cultura que está representando, ou seja, ela é compatível com os conceitos da garantia cultural, da multiculturalidade e da transculturalidade.

A TCT possibilita

[...] dar conta dos termos como unidades ao mesmo tempo singulares e similares a outras unidades de comunicação, dentro de um esquema global de representação da realidade, admitindo a variação conceitual e denominativa, e tendo em conta a dimensão textual e discursiva dos termos. (Cabré, 1999, p.136).

A terminologia busca a representação do conhecimento por meio de conceitos representados por termos, segundo o contexto e especificidade da área a ser representada. Acrescenta-se ao contexto de definição do conceito a preocupação com o contexto cultural no qual o vocabulário se encontra inserido.

Na perspectiva do contexto cultural, apresentamos a garantia cultural, a multiculturalidade e a transculturalidade. O termo garantia cultural foi primeiramente utilizado em 1976 por Joel M. Lee, que o descreveu como "[...] uma aplicação prática dos valores culturais expressa na literatura produzida por uma cultura determinada" (p.111-112).

Begthol (2002, p.511) apresenta o conceito de garantia cultural afirmando que "Qualquer sistema de representação e organização de conhecimento pode ser maximamente apropriado e útil para os indivíduos em uma cultura somente se for baseado nos pressupostos, valores e predisposições dessa mesma cultura (tradução nossa)."

Barité (2010) apresenta cinco fatores para integrar o princípio da garantia cultural em linguagem documentária: 1) ênfase nas culturas ou nos aspectos locais; 2) respeito pelas minorias; 3) não utilização de termos discriminatórios ou pejorativos; 4) equilíbrio de gênero nos assuntos relacionados com homens e mulheres; 5) equidade na apresentação de religiões, ideias, correntes e teorias.

A garantia cultural busca representar a informação da maneira mais apropriada e útil à cultura a qual busca atender. Seus valores buscam a representação mais próxima e mais ética possível da realidade desse conjunto de usuários, levando em considerações questões sociais, religiosas e culturais, inclusive das minorias sociais.

A garantia cultural desta forma contribui com a integração cultural, reafirmando ao mesmo tempo a identidade das culturas locais, e inclusive neutralizando nos sistemas de organização do conhecimento os efeitos de aculturação associados com a globalização e com os processos políticos e econômicos de exclusão social. (Barité, 2011, p.9)

Barité (2011) reafirma a importância da integração cultural, a neutralidade em sua apresentação e a inclusão de todas as vertentes culturais, independente de serem de grande massa ou de minorias sociais.

A garantia cultural é muito útil e recomendável para a aplicação em linguagens documentárias, devendo no caso das linguagens multilíngues ser aplicada em cada língua representada, de modo a atender às especificidades linguísticas e culturais.

A multiculturalidade "[...] designa a característica de sociedades formadas por múltiplas comunidades culturais, que convivem entre si" (Lazzarin, 2008, p.122), assim reconhece e busca representar com a mesma importância as múltiplas culturas existentes dentro de um mesmo grupo de indivíduos. Sua aplicação é muito apropriada em linguagens documentárias monolíngues e multilíngues, pois cada língua tem especificidades regionais e nacionais.

A multiculturalidade reconhece a pluralidade presente em grupos sociais. No Brasil, por exemplo, temos múltiplas culturas convivendo, com características próprias, marcadas geralmente pelo regionalismo, que se apresentam na fala, nos costumes, crenças e tradições.

O conceito de transculturalidade

[...] consiste em perceber o todo (a população mundial [...] globalizada e às redes de informação) e suas relações com as partes, que são as culturas, respeitando-as e permitindo que elas dialoguem

com o todo, na construção de um mundo mais equilibrado em termos de troca de conhecimentos. (Lucchesi & Malanga, 2011, p.75).

Ao contrário do multiculturalismo, que vê a presença de múltiplas culturas coexistindo, a transculturalidade observa a interação das culturas de modo a se misturarem e se tornarem transculturais, a ponto de se tornarem integradas e de ser difícil diferenciá-las. A transculturalidade vê a fusão de distintas culturas formando uma cultura supranacional.

Nicolescu (2005) apresenta a transculturalidade como algo que

[...] vai além da interpretação de uma cultura por outra cultura e também não se restringe a fecundação de uma cultura por outra cultura, mas a mesma assegura a tradução de uma cultura para qualquer outra cultura, através do sentido que une as mais diferentes culturas, mesmo que as ultrapassando (Neiva, Alonso & Ferneda, 2007, p.6)

Conjunto a esses quatro conceitos apresentados, temos a Linguagem Documentária, que para Lancaster (2002) visa melhorar a representação tanto na entrada do sistema, para os indexadores como na sua saída, no momento da busca realizada pelos usuários. O controle do vocabulário permite agrupar os sinônimos e quase sinônimos, diferenciar os homógrafos, entre outros casos, realizando a padronização que facilita e torna mais preciso o acesso.

Referindo-se às linguagens documentárias, Tálamo et al.(1994, p.18) afirmam que estas "funcionam como instrumentos intermediários, através dos quais se traduzem, de forma sintética, as informações contidas em textos, ou as perguntas dos usuários, para a linguagem do sistema documentário". A linguagem documentária atua, deste modo, como uma "ponte" entre o conteúdo do documento e o usuário do sistema.

Para Gardin et al. (1968 *apud* Cintra, 2002, p.35-36) uma linguagem para ser consistente precisa possuir três elementos básicos em sua estrutura: a) um léxico, identificado com uma lista de elementos descritores, devidamente filtrados e depurados; b) uma rede paradigmática para traduzir certas relações, entre os descritores; c) uma rede sintagmática destinada a expressar as relações entre os descritores.

Por meio da linguagem o conteúdo do documento pode ser representado o mais fielmente possível e posteriormente recuperado pelo usuário. Neste caso, a linguagem documentária assume função intermediária e requer, mais uma vez, o controle de vocabulário para que esta comunicação ocorra de modo preciso.

A norma ANSI/NISO Z39: 19 (2005) apresenta quatro princípios relativos ao controle do vocabulário: a) eliminação de ambiguidades; b) controle de sinônimos; c) estabelecimento relações apropriadas entre os termos; d) teste e validação de termos.

A aplicação conjunta da TCT e dos conceitos de multiculturalidade, transculturalidade, garantia cultural, para a construção de linguagem documentária, possibilita a representação da informação por meio de termos que apresentem os conceitos inerentes as distintas realidades culturais de usuários reais e potenciais. "Isso posto, ressaltamos que usuários em diferentes culturas necessitam de diferentes tipos de informação; a representação da informação deve ter um vínculo com um acervo documentário e com o universo em que o usuário está integrado." (Bocato, 2009, p.19).

Pautados nos preceitos de Bocato notamos que os usuários em distintas culturas e falantes de distintas línguas, necessitam de diferentes tipos de informação, conforme o contexto sócio-histórico-cultural ao

qual se encontram inseridos. Portanto, a necessidade de informação do usuário está relacionada ao meio ao qual ele se encontra inserido.

Para que no momento da busca essa necessidade de informação seja atendida é necessário que o profissional da informação no momento da representação tenha considerado o perfil dos usuários. Portanto a maneira de representar a informação é influenciada pelo contexto social do usuário.

A integração dos conceitos da TCT, garantia cultural, multiculturalidade e transculturalidade na construção de linguagens documentárias, irá permitir uma visão mais ampla e conceitual na recolha terminológica para fins de representação e recuperação da informação. O Quadro 1 resume os conceitos:

Quadro 1. Integração entre os conceitos da TCT, garantia cultural, multiculturalidade, transculturalidade e Linguagem Documentária

TCT	Garantia Cultural	Multiculturalidade	Transculturalidade	Linguagem Documentária
Dá ênfase à linguagem e seus aspectos comunicativos, para fins de disseminação da informação, não fazendo uso de termos fixos, possibilitando que estes sejam modificados conforme o contexto em que se apresentam.	Representação da informação conforme o contexto cultural dos usuários.	Representação da informação com enfoque em múltiplas culturas.	A transculturalidade se fundamenta na interação, isto é, no intercambio acelerado, global, descentralizado e atemporal de valores e categorias desarraigadas e procedentes de geoculturas em dissolução e de cada vez mais complexa rastreabilidade. (García Gutiérrez, 2014, p.14).	Princípios: a) eliminação de ambiguidades; b) controle de sinônimos; c) estabelecimento relações apropriadas entre os termos; d) teste e validação de termos. ANSI/NISO Z39: 19 (2005)
Admite a variação conceitual e denominativa, e tendo em conta a dimensão textual e discursiva dos termos. (Cabré, 1999).	Equidade na representação de assuntos religiosos, ideais e correntes teóricas.	Reconhecimento da existência de distintas culturas dentro do mesmo grupo social	Interação entre diversas culturas de modo a formar uma cultura supranacional, que transpassa fronteiras	A linguagem documentária atua como uma “ponte” entre o conteúdo do documento e o usuário do sistema.
A terminologia busca a representação do conhecimento por meio de conceitos representados por termos, segundo o contexto e especificidade da área a ser representada.	A representação só é apropriada e útil se for baseada em pressupostos, valores e predisposições da cultura dos indivíduos (usuários). (Begthol, 2002)	Designa a característica de sociedades formadas por múltiplas comunidades culturais, que convivem entre si" (Lazzarin, 2008, p.122).	Reconhecimento da integração e tradução de uma cultura para qualquer outra cultura, através do sentido que une as mais diferentes culturas, mesmo que as ultrapassando (Neiva, Alonso & Ferneda, 2007, p.6)	Uma linguagem para ser consistente precisa possuir três elementos básicos em sua estrutura: a) um léxico; b) uma rede paradigmática; e c) uma rede sintagmática. Gardin et al. (1968)

	Equidade de gênero, respeito a minorias sociais, evitar termos pejorativos e/ou discriminatórios	Reconhecimento de culturas regionais e nacionais distintas.	Nova forma de vivenciar a cultura, de relacionamento com o mundo, sem aparentes dominadores nem dominados com uma comunicação horizontal ilimitada (García Gutiérrez, 2014, p.15).	
--	--	---	--	--

Elaborado pelos autores (2017).

A combinação da TCT com os conceitos da garantia cultural, da multiculturalidade e da transculturalidade possibilitam e dão embasamento teórico para auxiliar na construção de linguagens documentárias. Esse conjunto de conceitos é plenamente compatível com os conceitos referentes à construção de linguagens documentárias, que buscam o controle terminológico, o teste e validação de termos, o controle de sinônimos, homógrafos e ambiguidades, objetivando uma representação da informação de maneira mais precisa e eficaz ao usuário.

A TCT permite a variação dos termos para a representação dos conceitos, reconhecendo a diversidade terminológica para a representação da informação. Essa variedade se dá em função do contexto social e cultural dos usuários, o que nos leva até os conceitos da garantia cultural, multiculturalidade e transculturalidade, que reconhecem a diversidade cultural, e a importância da representação dos termos de maneira distinta para diferentes culturas e realidades.

Desse modo, a disseminação da informação se dará de maneira mais igualitária e justa, alcançando diferentes usuários em distintas realidades culturais. Uma linguagem documentária que atenda as necessidades de sua comunidade usuária possibilita um acesso pleno e eficaz à informação, objetivo maior da Ciência da Informação e da Organização do Conhecimento.

Por meio da análise teórica conceitual nota-se que o uso conjunto desses conceitos na construção de linguagens documentárias irá possibilitar que o usuário se identifique com os termos e com a estrutura hierárquica representada segundo o seu contexto cultural, as especificidades de sua língua e cultura. Acreditamos que a não aplicação desse conjunto, implica em representações que podem diferir da realidade dos usuários, gerando equivalências e hierarquias estranhas a alguns grupos, o que reforça a importância de sua aplicação, pois resgatando Beghtol (2002) a representação só é apropriada e útil quando é baseada em pressupostos, valores e predisposições da cultura dos usuários.

Considerando-se os princípios e fundamentos relativos à transculturalidade, a multiculturalidade, a garantia cultural e com base na Teoria Comunicativa da Terminologia, apontam-se na sequência os requisitos mínimos a serem observados na construção e no uso de linguagens documentárias.

A linguagem documentária então deverá ser padronizada de acordo com as necessidades do usuário, apresentando termos representativos de sua realidade sócio-histórico-cultural, de modo a aperfeiçoar os resultados de representação e disseminação da informação.

A aplicação desses conceitos irá auxiliar no cumprimento desses objetivos, possibilitando a compreensão e especificidade do conjunto de usuários.

CONCLUSÕES

A aplicação desse conjunto TCT, garantia cultural, multiculturalidade e transculturalidade no auxílio para a construção de linguagens documentárias, com enfoque na estrutura hierárquica e na recolha terminológica, possibilita atender a distintos usuários em diferentes realidades sócio-histórico-culturais. O uso desses conceitos no auxílio na construção de linguagens documentárias irá possibilitar que minorias sociais, raciais e culturais sejam representadas de maneira igualitária em relação à cultura da grande massa, de modo a

[...] facilitar a comunicação inter-cultural em uma sociedade da informação cada vez mais global. [...] permitindo a igualdade de representação de todas as línguas [...], e em parte porque as estruturas semânticas tornam o contexto conceitual de cada termo e suas traduções mais explícita do que as listas ordenadas de forma aleatória” (Jorna & Davies, 2001, p.285)

O reconhecimento da multiculturalidade, da diversidade e da importância da inclusão de todos os usuários reais e potenciais na representação da informação justifica esse cuidado no momento da representação e organização do conhecimento, contemplando o objetivo maior das unidades de informação, que é atender a todos os usuários de maneira igualitária, ética e justa, disseminando a informação e auxiliando na produção de novos conhecimentos para a sociedade.

Conclui-se, portanto que a aplicação desse conjunto de conceitos possibilita que a linguagem documentária se aproxime mais do usuário e de suas perspectivas, pois com esse auxílio, no momento da construção da linguagem, os termos, sua organização e estruturação hierárquica, serão os mais próximos da realidade dos usuários, estejam eles incluídos em maiorias ou minorias sociais.

REFERÊNCIAS BIBLIOGRÁFICAS

ANSI/NISO Z39.19 (2005). Guidelines for the construction, format, and management of monolingual controlled vocabularies. Disponível em: http://www.niso.org/apps/group_public/download.php/6484/Guidelines%20for%20the%20Construction,%20Format,%20and%20Management%20of%20Monolingual%20Controlled%20Vocabularies.pdf

Barité, M. et al. (2010). Garantia literária: elementos para uma revisão crítica após um século. *Transinformação* (Campinas) 2(22), 123-138. Disponível em: http://www.scielo.br/scielo.php?pid=S0103-37862010000200003&script=sci_abstract&lng=pt

Barité, M. (2011). La garantía cultural como justificación en sistemas de organización del conocimiento: aproximación crítica. *Palabra Clave* (La Plata) 1(1), 2-11.

Begthol, C. (2002). A proposed ethical warrant for global knowledge representation and organization systems. *Journal of Documentation*, 58(5), 507-532.

Bocato, V. R. (2005). Avaliação de linguagem documentária em Fonoaudiologia na perspectiva do usuário: estudo de observação da recuperação da informação com protocolo verbal. Disponível em: http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/bocato_vrc_me_mar.pdf

Bocato, V. R. (2009). Avaliação do uso de linguagem documentária em catálogos coletivos de bibliotecas universitárias: um estudo sociocognitivo com protocolo verbal. Disponível em: http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/bocato_vrc_do_mar.pdf

Cabré, M. T. (1993). La Terminología: teoría, metodología, aplicaciones. Traducción castellana de Carles Tebé. Barcelona: Antártica/Empúres.

Cabré, M. T. (1999). La terminología: representación y comunicación: elementos para una teoría de base comunicativa y otros artículos. Barcelona: IULA.

Cabré, M. T. (2003). Theories of terminology: their description, prescription and explanation. *Terminology*, 9 (2) 163-200.

Cervantes, B. M. N. (2004). Contribuição para a Terminologia do Processo de Inteligência Competitiva: estudo teórico e metodológico. Disponível em: http://repositorio.unesp.br/bitstream/handle/11449/93723/cervantes_bmn_me_mar_prot.pdf?sequence=1&isAllowed=y

Cintra, A. M. M. et al. (2002). Para entender as linguagens documentárias. São Paulo: Polis.

Currás, E. (1995). Tesauros: linguagens terminológicas. Tradução de Antônio Felipe Corrêa da Costa. Brasília: IBICT.

Ferini, V. A. (2006). Dicionário terminológico bilíngüe francês português de termos jurídicos: tratamento terminográfico e reflexões sobre terminologia bilíngüe. Disponível em: http://repositorio.unesp.br/bitstream/handle/11449/86610/ferini_va_me_sjrp.pdf?sequence=1&isAllowed=y

García Gutiérrez, A. (2014). La organización del conocimiento en el nuevo orden transcultural: del totalitarismo a la desclasificación (obstáculos epistémicos, regencia de la transcultura y desclasificación). *Brazilian Journal of Information Science: research trends*, 8(1/2), 1-34. Disponível em: <http://www2.marilia.unesp.br/revistas/index.php/bjis/article/view/4450/3265>

Guimarães, J. A. C. (1990). Recuperação temática da informação. *Revista Brasileira de Biblioteconomia e Documentação* (São Paulo) 23(1/4), 112-130.

ISO 1087 (2000). Terminology work-vocabulary. Disponível em: http://www.iso.org/iso/catalogue_detail.htm?csnumber=20057

Jorna, K. & Davies, S. (2001). Multilingual thesauri for the modern world: no ideal solution? *Journal of Documentation*, 57(2), 284-295.

Krieger, M. G. (2005). Terminologias em construção: procedimentos metodológicos. *Termisul-Ufrgs, Unisinos*, 1-6. Disponível em: http://termisul.pedrorivera.me/wp-content/uploads/2015/04/artigo_ABECAN_2005_KRIEGER.pdf

Lancaster, F. W. (2002). El control del vocabulario en la recuperación de información. València: Universitat de València.

Lazzarin, L. F. (2008). Multiculturalismo e multiculturalidade: recorrências discursivas na educação musical. Revista da ABEM, 19, 121-128. Disponível em: http://abemeduacaomusical.com.br/revista_abem/ed19/revista19_artigo13.pdf

Lee, J. M. (1976). E. Wyndham Hulme: a reconsideracion, in Rayward, W. B. (Ed.), The Variety of Librarianship, Library Association of Australia, Sydney, 101-113.

Lucchesi, M. A. & Malanga, E. B. (2011). Interculturais e identidades nacionais:

Transculturalidade e transdisciplinaridade. Visão Global. Joaçaba, 14(1), 73-88. Disponível em: http://editora.unoesc.edu.br/index.php/visaoglobal/article/view/868/pdf_287

Neiva, R., Alonso, L. & Ferneda, E. (2007). Transculturalidade e Tecnologias da Informação e Comunicação. RENOTE – Revista Novas Tecnologias na Educação, 5(2). Disponível em: <http://seer.ufrgs.br/renote/article/view/14373/8270>

Nicolescu, B. (2005). O Manifesto da transdisciplinaridade. Tradução de Lucia Pereira Souza. 3. ed. São Paulo: Triom.

Tálamo, M. de F. G. M. et al. (1994). Informação: do tratamento ao acesso e utilização. Comunicação e Educação (São Paulo), 1.



PERSPETIVAS EPISTEMOLÓGICAS EM ORGANIZAÇÃO DO CONHECIMENTO

O DOCUMENTO CIENTÍFICO COMO PRODUTO E COMO OBJETO DA ORGANIZAÇÃO DO CONHECIMENTO

Luciana Corts Mendes

Doutoranda do Programa de Pós-Graduação em Ciência da Informação da Escola de Comunicações e Artes da Universidade de São Paulo, 0000-0002-9546-6465, luciana.corts.mendes@usp.br

RESUMO O trabalho objetiva discutir o documento enquanto conceito integrante dos fundamentos teórico-conceituais da Organização do Conhecimento (OC), seja esta entendida de modo geral ou específico. A investigação tem como objeto o documento científico, sendo realizada sob a perspectiva da hermenêutica crítica por meio de levantamento, revisão e análise bibliográficos das áreas dos Estudos de Ciência e Tecnologia e OC. Entende-se o documento científico como um produto da OC geral, pois ele é a representação material do conhecimento organizado de um ou mais cientistas, e compreende-se que esse mesmo documento se torna objeto da OC específica. Conclui-se que o documento é o principal elo de conexão entre a OC geral e a OC específica, devendo ser integrado aos conceitos fundamentais da OC.

PALAVRAS-CHAVE *Documento, Organização do Conhecimento, Epistemologia.*

ABSTRACT The work aims to discuss the document as an integral concept of the theoretical-conceptual foundations of Knowledge Organisation (KO), be this understood in its broad or narrow sense. The investigation has as its object the scientific document, being executed under the perspective of critical hermeneutics through bibliographical survey, literature review, and analysis of the areas of Science and Technology Studies and KO. The scientific document is understood as a product of broad KO, since it is the material representation of the organised knowledge of one or more scientists, and it is understood as the object of narrow KO. It is concluded that the document is the main link of connection between broad and narrow KO, and that it should be integrated to the fundamental concepts of KO.

KEYWORDS *Document, Knowledge Organisation, Epistemology.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A Organização do Conhecimento (OC) é uma área afiliada à Ciência da Informação que se dedica a descrever, representar, arquivar e organizar documentos e suas representações (Hjørland, 2016). No cerne da área se encontram questões relacionadas à organização do conhecimento em sistemas de classificação e em sistemas conceituais, seus dois principais aspectos sendo os processos de organização do conhecimento – processos de catalogação, análise de assunto, indexação e classificação – e os sistemas de organização do conhecimento – sistemas de classificação, listas de cabeçalho de assunto, tesouros, ontologias e outros sistemas de metadados, sistemas que resultam da seleção de conceitos e da indicação das relações semânticas entre esses conceitos (Hjørland, 2016).

Hjørland (2008, 2016) afirma que a OC pode ser entendida de modo geral e de modo específico. A OC geral se refere 1) à organização do conhecimento em teorias dentro disciplinas acadêmicas específicas, isto é, à organização intelectual do conhecimento; e 2) à organização do conhecimento na sociedade, ou seja, à divisão social do trabalho intelectual em disciplinas e profissões, à organização de universidades e outras instituições de pesquisa e ensino superior e à disseminação do conhecimento (Hjørland, 2008, 2016). Por sua vez, a OC específica lida com “a descrição de documentos, a indexação e classificação realizada em bibliotecas, bases de dados bibliográficas, arquivos e outros tipos de ‘instituições de memória’” (Hjørland, 2008, p. 86); ou seja, a OC específica é a área que se afilia diretamente à Ciência da Informação.

A OC específica se centra na organização do conhecimento *registrado*, isto é, na organização do conhecimento *representado* em um *documento*. A OC geral, apesar de não se centrar apenas no conhecimento registrado, entende o documento como um elemento fundamental dos processos de produção e disseminação do conhecimento, ainda que sua importância na dinâmica de constituição e desenvolvimento de uma área possa passar despercebida.

Pelas razões anteriormente descritas, objetivamos discutir o documento enquanto conceito integrante dos fundamentos teórico-conceituais da OC, seja esta entendida de modo geral ou específico. O foco desta investigação é o documento científico, notadamente o documento produzido no âmbito das ciências naturais, de maneira que se faz relevante a análise de documentos de caráter diverso em pesquisas subsequentes.

Neste trabalho abordamos a OC específica enquanto área afiliada à Ciência da Informação e a OC geral na perspectiva dos chamados Estudos de Ciência e Tecnologia (ECT), surgidos a partir da década de 1960 com o aumento no interesse pelas dimensões sociais da ciência. Os ECT buscam por meio da sociologia e da história compreender o fazer científico contextualmente e através da análise de pessoas, práticas padronizadas, instrumentos materiais, narrativas científicas e suas inter-relações, compreender o processo de construção do conhecimento científico.

Destarte, realizamos uma análise crítico-hermenêutica da literatura das áreas da OC e dos ECT para explanarmos o lugar do documento em cada uma das instâncias nas quais o conhecimento é organizado. Por meio da revisão de literatura das áreas da OC e dos ECT entendemos o documento científico como um produto da OC geral, pois ele é a representação material do conhecimento organizado de um ou mais cientistas, sendo também objeto da OC específica, que organiza o conhecimento registrado nas instituições de memória. Concluímos então que o documento pode ser entendido como o principal elo de conexão entre a OC geral e a OC específica, de modo que defendemos a integração desse conceito nos fundamentos teórico-conceituais da área e, portanto, que mais atenção lhe seja dada por parte dos acadêmicos da OC.

METODOLOGIA

De modo a cumprir nosso objetivo de pesquisa, desenvolvemos uma investigação de natureza teórica e exploratória por meio de levantamento, revisão e análise bibliográficos. Ou seja, realizamos a pesquisa e seleção de textos na bibliografia da área que atendessem a nosso objetivo de investigação e procedemos à leitura e análise dos mesmos, apropriando-nos dessa bibliografia em consonância com as ideias expostas pelos autores e cotejando-as entre si mediante a hermenêutica crítica.

A hermenêutica crítica é uma metodologia de interpretação dos significados dos fenômenos investigados que se preocupa com a historicidade do conhecimento, isto é, pressupõe que o entendimento somente pode ser atingido ao se ter como referência o momento histórico e o contexto sociocultural no qual esses fenômenos tomaram lugar (Hjørland, 1998; Radford, 1991). A hermenêutica crítica é uma metodologia que objetiva a teorização analítica e avaliativa, buscando não apenas descrever os fenômenos investigados e o mundo, mas também os criticar e modificar (Radford, 1991). Nossa adoção da hermenêutica crítica como metodologia de análise resulta do entendimento de que explorar o documento como conceito integrante dos fundamentos teórico-conceituais da OC pressupõe o entendimento e a interpretação desse conceito dentro dos contextos específicos de organização intelectual e social do conhecimento, bem como no contexto da organização do conhecimento em instituições de memória. Ademais, como objetivamos desenvolver a ideia de que o documento é parte essencial dos fundamentos teórico-conceituais da OC, assumimos um posicionamento crítico que visa modificar ou ampliar os pressupostos teórico-metodológicos da área.

Para explicar o documento científico como elemento fundamental da OC, iniciamos a investigação pela análise da literatura dos ECT, de modo a compreender o lugar do documento na OC geral. A seguir, analisamos a literatura da OC específica para discutimos o documento dentro dessa área. O trabalho é finalizado explicitando-se de que modo o documento pode ser entendido como o principal elo de conexão entre OC geral e OC específica.

O LUGAR DO DOCUMENTO NA ORGANIZAÇÃO DO CONHECIMENTO

Nossa investigação sobre o papel do documento na OC parte da proposição de Hjørland (2008, 2016) de que a OC apresenta duas subdivisões, uma geral e uma específica. A OC geral se refere 1) à organização do conhecimento em teorias dentro de disciplinas acadêmicas específicas, isto é, à organização intelectual do conhecimento; e 2) à organização do conhecimento na sociedade, ou seja, à divisão social do trabalho intelectual em disciplinas e profissões, à organização de universidades e outras instituições de pesquisa e ensino superior e à disseminação do conhecimento (Hjørland, 2008, 2016). Por sua vez, a OC específica lida com “a descrição de documentos, a indexação e classificação realizada em bibliotecas, bases de dados bibliográficas, arquivos e outros tipos de ‘instituições de memória’” (Hjørland, 2008, p. 86).

Ao tratarmos a OC geral na perspectiva dos ECT, assumimos que o conhecimento científico navega pela sociedade juntamente com pessoas, práticas padronizadas, instrumentos materiais e narrativas científicas. Centradas no experimento, as práticas científicas dependem de tipos padronizados de comportamento, assim como de espaços da ciência organizados (Schaffer, 1999). Como coloca Kropf (2009, p. 56), o laboratório é o “espaço por excelência de produção das verdades científicas”, onde são estabelecidas “regras experimentais rigorosas para verificá-los”. Dos laboratórios e outros espaços de produção da ciência saem propostas de regras e procedimentos a serem compartilhados pela comunidade científica, algo necessário para que os fatos científicos possam ser reproduzidos indiscutivelmente e para que se construa um sistema uno de circulação da ciência. A perpetuação desse sistema científico vincula-se ao sistema de ensino, de modo que, através da educação formal e do treino, os estudantes se tornam aptos a realizar experimentos científicos (Licoppe, 1999; Schaffer, 1999).

O hiato entre a experiência individual do cientista e a publicitação do conhecimento é sempre preenchido pela persuasão. Desse modo, a sociologia do conhecimento científico possui um grande

número de pesquisas relacionadas às “técnicas de exposição científica”, ou seja, os meios pelos quais os cientistas efetuam o convencimento de seus pares de que suas afirmações são verdadeiras e significativas (Shapin, 1995, p. 305). As experiências laboratoriais, as performances públicas e os instrumentos científicos são fundamentais para que se entenda como o conhecimento científico é produzido e perpetuado. Contudo, a produção de literatura científica possui um papel de destaque na construção do conhecimento científico e no convencimento da comunidade científica de que o discurso de um único indivíduo se trata efetivamente de conhecimento. Apesar da importância de todas as diferentes práticas científicas, “é indiscutivelmente sobre a narrativa que repousa todo o peso da construção do facto experimental novo e inesperado a partir do qual os cientistas poderão, em seguida, filosofar à vontade” (Licoppe, 1999, p. 221).

O documento científico é construído por meio da tecnologia literária do testemunho virtual, uma técnica de exposição científica que produz “na mente do leitor uma imagem naturalística da cena do experimento” (Shapin, 2013, pp. 98-99). A tecnologia literária do testemunho virtual é capaz tanto de tornar o conhecimento individual de um cientista ou de um grupo de cientistas em uma certeza da área em que atuam, como de torná-lo um enunciado vazio, isto é, desprovido de valor para a ciência e, portanto, não entrando para o corpo teórico-conceitual de uma área. Essa tecnologia é bem-sucedida quando os artifícios de representação utilizados – uma linguagem que se pretende funcional, ou seja, básica, puritana e sem ornamentos – levam a comunidade científica a crer que o conhecimento registrado é válido e objetivo (Frohmann, 1999, 2001; Knorr-Cetina, 1999; Latour, 2000; Shapin, 2013).

A tecnologia literária do testemunho virtual é entendida por Shapin (2013, p. 98) como mais importante do que a realização de experimentos perante testemunhas presenciais ou da facilitação de sua reprodução real, pois como coloca Licoppe (1999, p. 224), “a reprodutibilidade das experiências pode ser desejável e desejada”, mas “não é, no entanto essencial”, pois por meio da literatura científica “é possível [...] legitimar um facto acontecido uma única vez”.

Shapin (2013, p. 93) afirma que os textos científicos são uma tecnologia literária por se tratarem de muito mais do que relatórios, configurando-se como um meio de estender e validar a experiência científica. Isto se dá na medida em que a tecnologia literária se configura como um *testemunho virtual*:

A tecnologia de testemunho virtual envolve a produção na mente do leitor de uma imagem de uma tal cena de realização de experimento que diminuía a necessidade tanto da testemunha direta quanto da reprodução do experimento. Por meio do testemunho virtual a multiplicação de testemunhas podia ser limitada em princípio. Esta era, portanto, a tecnologia mais poderosa para a constituição de matérias de fato. A validação de experimentos, e o crédito dado aos seus resultados como matérias de fato, implicavam necessariamente sua realização no laboratório da mente e também o olhar da mente. O que se requeria era uma tecnologia de confiança e assecuração de que tudo havia sido feito, e feito do modo que havia sido proposto (Shapin, 2013, p. 98).

Somente através da tecnologia do testemunho virtual seria possível acabar com a “suspeita natural e legítima dos que não eram nem testemunhas diretas nem faziam reproduções dos experimentos” e validar os resultados de uma experiência (Shapin, 2013, p. 98-99). Em outras palavras,

caso se escrevesse um relatório sobre um experimento de forma adequada, o leitor podia imbuir-se da confiança de que aquilo ocorreu. Seria como se os leitores tivessem estado presente às atividades. Eles seriam convocados como testemunhas e lhes seria dada a condição de poder validar os fenômenos experimentais como sendo matérias de fato. A atenção para com a escrita de relatórios

sobre experimentos tinha uma importância prática que era pelo menos igual à própria realização dos experimentos (Shapin, 2013, p. 100).

Apesar de contextual, o conhecimento técnico-científico permanece estável e é tratado como verdadeiro contanto que seja um ponto fulcral no trabalho de muitos indivíduos. Ou seja, “intercalar a ciência ou a tecnologia em redes de ação cada vez maiores é o que as torna duráveis” (Shapin, 1995, p. 308). Desse modo, “quando todos os elementos de uma rede agem juntamente para proteger um item de conhecimento, então aquele conhecimento é forte e nós passamos a chamá-lo de científico” (Shapin, 1995, p. 308). Desse modo, é possível afirmar que o valor de uma afirmação depende dos debates posteriores que serão realizados pela comunidade científica sobre esse enunciado. Assim, ocorre uma “transformação retrospectiva do valor de verdade” (Latour, 2000, p. 51) de uma sentença e é possível concluir que a construção de um fato científico é coletiva.

Este é um processo onde são acrescentados diversos “recursos para fazer o equilíbrio de forças propender a seu favor”, de modo que quando um cientista é desafiado possui como base de sustentação outros recursos “organizados em camadas e prontos para dizer a mesma coisa” (Latour, 2000, p. 150). Por essa razão o texto científico é estratificado, organizado em camadas. Cada afirmação é alicerçada por referências dentro e fora do texto, figuras, tabelas, legendas, gráficos e etc. Nenhuma parte de um texto é independente do seu restante (Latour, 2000). Ou seja, cada novo texto transforma “*a literatura anterior para atender às suas necessidades*” (Latour, 2000, p. 68, grifos do autor). Através desse processo aquilo que foi declarado por um único indivíduo vai aos poucos se transformando em fato, de modo que “a força da afirmação original não reside em si mesma, porém deriva de qualquer dos textos que a incorporam” (Latour, 2000, p. 72).

Os recursos anteriormente descritos são parte do processo de comunicação da ciência, que lhe é intrínseco:

a comunicação é dita ser intrínseca à ciência pelo facto de a ciência moderna ser um empreendimento colectivo que depende de os resultados obtidos por cientistas individuais serem retomados por outros cientistas que se fundam neles e os desenvolvem. A ciência projecta-se a si mesma no futuro através da comunicação. Uma ciência privada é tão impensável como uma linguagem privada. Apenas os resultados que são comunicados podem contar, tanto na ciência quanto na aplicação científica e na prática tecnológica (Knorr-Cetina, 1999, p. 378).

É através da comunicação que os consensos científicos podem ser atingidos, que se dá o processo de retenção seletiva e a sobrevivência das descobertas científicas, assim como o prosseguimento de negociações e controvérsias (Knorr-Cetina, 1999). Ao circular de modo amplo, o conhecimento científico de determinadas culturas se torna o padrão de produção e aplicação do conhecimento (Shapin, 1995).

Assim, é possível dizer que documentos organizam comunidades epistêmicas e manufacturam conhecimento científico, conhecimento este de autoria múltipla e produzido por práticas literárias socialmente disciplinadas, práticas documentárias institucionalizadas que tornam um documento informativo e inteligível (Frohmann, 2004, 2007). Ou seja, o significado de um documento não está nele em si mesmo, mas é construído por uma comunidade de interpretação que existe em torno do texto ou documento em consideração, de modo que documentos são um recurso para a construção e negociação do espaço social, um meio para criar e manter grupos sociais com um forte senso de identidade compartilhada (Brown, & Duguid, 1996). Nas palavras de Drucker (em linha), “não vejo um fato material simples, positivo quando olho para um documento, vejo campos de relações mutantes

momentaneamente estabilizadas em um artefato que existe em um contínuo de dimensões temporais, espaciais e quânticas, constituído somente através de atos enquadráveis de intervenção”.

Assim, o documento científico pode ser entendido como a representação do conhecimento de um ou mais cientistas, um conhecimento acadêmico que resulta da divisão social do trabalho intelectual. Esta representação é produzida por meio de uma tecnologia de organização do conhecimento que visa sua disseminação, configurando o documento como um produto da OC geral.

As características primárias do documento são a mobilidade e a imutabilidade, de modo que ele deve ser autossuficiente, ou seja, ser capaz de comunicar o conhecimento registrado sem que seu autor intervenha pessoalmente para o clarificar (Brown, & Duguid, 1996). Disto resulta o desafio de que o documento seja reconhecido como pertinente aos interesses do público ao qual se destina, desafio que aumenta quanto mais documentos são produzidos e quanto maior o seu alcance espaço-temporal (Brown, & Duguid, 1996).

À OC específica cabe buscar uma solução para esse desafio, isto é, a ela cabe determinar quais documentos são pertinentes aos interesses da comunidade usuária de uma instituição de memória específica. É assim que o documento se torna objeto da OC específica, pois esta visa organizar registros do conhecimento dentro de uma instituição de memória, uma estrutura que organiza os documentos relevantes às necessidades e aos problemas de uma comunidade específica.

Ridenour (2016, p. 44), baseando-se em Wilson (1978), afirma que o “conhecimento deve ser codificado [...] para ser explorado e controlado. Uma vez registrado, ele pode ser explorado de diversas maneiras; algumas servem para facilitar o controle e outros que podem resultar na criação de novo conhecimento”. Assim, a OC específica estabelece relações entre documentos, gerando um sistema de documentos interdependentes (Olsen, Lund, Ellingsen, & Hartvigsen, 2012). Essas relações são estabelecidas por meio de processos e sistemas de organização do conhecimento através de duas operações, 1) a representação do conteúdo documentos e 2) sua organização por meio de seleção, classificação e agrupamento (Buckland, 2012). Destarte, entender como documentos são criados e representados, bem como de que maneira se relacionam e como suas relações são comunicadas é fundamental para o desenvolvimento de instituições de memória (Hjørland, 2013; Olsen et al., 2012).

Os sistemas de organização do conhecimento “tiram uma entidade de seu contexto original, recriam-na como um substituto e a inserem no contexto muito diferente do sistema de organização do conhecimento” (Olson, 2012, p. 113), por essa razão o processo de seleção e incorporação da informação registrada a um sistema de armazenamento, organização e recuperação da informação depende dos discursos que as instituições de memória perpetuam, de modo que possa haver comunicação. Ou seja, é necessário que a comunicação ocorra dentro de um quadro contextual compartilhado, ou universo do discurso, universo este que impõe limitações, ou condições, ao entendimento (Olson, 2012; Peirce, 1994; Thellefsen, Sørensen, & Thellefsen, 2011). Por este motivo a OC específica “precisa considerar as teorias e paradigmas fundamentais dos domínios que organiza”, ou seja, a organização do conhecimento registrado depende da organização intelectual do conhecimento (Hjørland, 2013, p. 174). Assim, sistemas de organização do conhecimento e a organização específica dos documentos de uma instituição de memória revelam um dado paradigma disciplinar (Hjørland, 2013).

Isso ocorre pelo fato de que a própria instituição faz parte de uma comunidade epistêmica, suas atividades sendo social e institucionalmente disciplinadas para que seus usuários possam recuperar e entender a informação ali armazenada e organizada. Portanto, o estabelecimento de um sistema de

armazenamento, organização e recuperação da informação pressupõe o entendimento do universo discursivo de seu grupo alvo para que haja comunicação entre o sistema e seu usuário, de modo que o processo de seleção e incorporação de informação registrada em um sistema de armazenamento, organização e recuperação da informação, depende de um contexto cultural e o reflete (Buckland, 2012; Smit, 2003, 2012).

Frohmann (2001) considera estas instituições como instituições discursivas, do mesmo modo que a tecnologia literária do testemunho virtual é uma prática discursiva que disciplina a escrita científica. O autor conclui então que “os maquinários de produção, organização, recuperação e classificação de documentos, estão entre os ‘fragmentos e pedaços’ que constituem a matéria prima das práticas científicas” (Frohmann, 2001, pp. 20-21). Portanto, a OC específica é fundamental para a OC geral, seja intelectual ou socialmente, e a OC geral produz a matéria prima da OC específica.

Concluímos então que o documento científico se encontra no cerne da OC, seja como produto da OC geral ou objeto da OC específica, configurando-se como seu elo fundamental de conexão.

CONCLUSÕES

Analisar o papel do documento científico na OC nos permitiu verificar que ele é um elemento central da área, seja sob uma perspectiva geral ou sob uma perspectiva específica. Também foi possível identificar que o documento pode ser entendido como o principal elo de conexão entre a OC geral e a OC específica, de modo que esse conceito deve integrar os fundamentos teórico-conceituais da área e, portanto, mais atenção lhe deve ser dada por parte dos acadêmicos da OC.

Como o foco desta investigação foi o documento científico, entendemos que a análise de documentos de outro caráter deve ser objeto de novas pesquisas, de maneira que se compreenda como as conclusões desta investigação se aplicam ou não a outros tipos de documento e para que seja melhor entendido o papel desses outros tipos de documento na OC e como eles atuam entre as subdivisões geral e específica da OC.

AGRADECIMENTOS

Os resultados deste trabalho provem de investigação em andamento apoiada financeiramente com bolsa de doutorado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) do Ministério da Educação (MEC) do Brasil.

REFERÊNCIAS BIBLIOGRÁFICAS

Brown, J. S., & Duguid, P. (1996). The social life of documents. *First Monday*, 1(1). Retrieved March 17, 2017 from <http://firstmonday.org/ojs/index.php/fm/article/view/466/387>

Buckland, M. (2012). What kind of science can Information Science be? *Journal of American Society for Information Science and Technology*, 63(1), 1-7.

- Frohmann, B. (1999). The role of the scientific paper in science information systems. In Bowden, M. E., Hahn, T. B., & Williams, R. V. (Eds.). *Proceedings of the 1998 Conference on the History and Heritage of Science Information Systems* (pp. 63-73). Medford, NJ: Information Today.
- Frohmann, B. (2001). Discourse and documentation: some implications for pedagogy and research. *Journal of Education for Library and Information Science*, 42(1), 12-26.
- Frohmann, B. (2004). Documentation redux: prolegomenon to (another) philosophy of information. *Library Trends*, 52(3), 387-407.
- Frohmann, B. (2007). Multiplicity, materiality, and autonomous agency of documentation. In Sklare, R., Windfeld Lund, N., & Vårheim, A. (Eds.). *A document (re)turn: contributions from a research field in transition* (pp. 27-39). Frankfurt: Peter Lang.
- Hjørland, B. (1998). Theory and metatheory of information science: a new interpretation. *Journal of Documentation*, 54(5), 606-621.
- Hjørland, B. (2008). What is Knowledge Organization (KO)? *Knowledge Organization*, 35(2/3), 86-101.
- Hjørland, B. (2013). Theories of knowledge organization – theories of knowledge. *Knowledge Organization*, 40(3), 169-181.
- Hjørland, B. (2016). Knowledge Organization (KO). *Knowledge Organization*, 43(6), 474-484.
- KROPF, S. P. (2009). *Doença de Chagas, doença do Brasil: ciência, saúde e nação, 1909-1962*. Rio de Janeiro: Fiocruz.
- Knorr-Cetina, K. (1999). A comunicação na ciência. In GIL, F. (Coord.). *A ciência tal qual se faz* (pp. 375-393). Lisboa: João Sá da Costa.
- Latour, B. (2000). *Ciência em ação: como seguir cientistas e engenheiros sociedade afóra*. São Paulo: UNESP.
- Licoppe, C. (1999). Barómetros e termómetros em França no século XVIII: modalidades históricas da coordenação de medições feitas à distância. In GIL, F. (Coord.). *A ciência tal qual se faz* (pp. 215-245). Lisboa: João Sá da Costa.
- Olsen, B. I., Lund, N. W., Ellingsen, G., & Hartvigsen, G. (2012). Document theory for the design of socio-technical systems: a document model as ontology of human expression. *Journal of Documentation*, 68(1), 100-126.
- Olson, H. A. (2012). Distilling essence, enforcing shibboleth. In: Guimarães, J. A. C., & Dodebei, V. (Org.). *Desafios e perspectivas científicas para a organização e representação do conhecimento na atualidade* (pp. 107-121). Marília: ISKO Brasil.
- Peirce, C. S. (1994). *The electronic edition of the collected papers of Charles Sanders Peirce*. Cambridge, MA: Harvard University Press.

Radford, G. P. (1991). Hermeneutics: an intellectual tradition for communication studies. *Occasional Papers in Communication, Information, and Library Studies*, 1, 6-27. Retrieved January 12, 2017 from <http://www.theprofessors.net/hermen.html>.

Ridenour, L. (2016). Boundary objects: measuring gaps and overlap between research areas. *Knowledge Organization*, 43(1), 44-55.

Schaffer, S. (1999). As instituições científicas: a geografia histórica dos laboratórios. In GIL, F. (Coord.). *A ciência tal qual se faz* (pp. 415-436). Lisboa: João Sá da Costa.

Shapin, S. (1995). Here and everywhere: sociology of scientific knowledge. *Annual Review of Sociology*, 21, 289-321.

Shapin, S. (2013). *Nunca pura: estudos históricos de ciência como se fora produzida por pessoas com corpos, situadas no tempo, no espaço, na cultura e na sociedade e que se empenham por credibilidade e autoridade* (Vol. 21, Scientia). Belo Horizonte: Fino Traço.

Smit, J. W. (2003). Arquivologia/biblioteconomia: interfaces das ciências da informação. *Informação & Informação*, 8(2). Retrieved July 5, 2014 from <http://www.uel.br/revistas/uel/index.php/informacao/article/download/1713/1464>.

Smit, J. W. (2012). A informação na Ciência da Informação. *InCID: Revista de Ciência da Informação e Documentação*, 3(2), 84-101.

Thellefsen, T., Sørensen, B., & Thellefsen, M. (2011). The significance-effect is a communicational effect: introducing the DynaCom. *Sign Systems Studies*, 39(1), 209-223.



ACHADOS TEÓRICO-EPISTEMOLÓGICOS EM TRABALHOS DO GT2-ENANCIB: UMA ANÁLISE BASEADA EM CATEGORIAS

Hildenise Ferreira Novo¹, Bruna Bomfim Lessa dos Santos², Carolina de Souza Santana Magalhães³, Daniel Almeida Marins⁴, Ivana Aparecida Borges Lins⁵, Luís Carlos Batista de Jesus⁶, Maíra Salles⁷

¹UFBA, 0000-0003-0986-7164, [hnovo@ufba.br](mailto:hново@ufba.br)

²UFBA, 0000-0003-4485-203X, brunalessa@ufba.br

³UFBA, /0000-0002-2391-3393, carolina.magalhaes@ufba.br

⁴UFBA, 0000-0002-2253-7410, danielmarinss@hotmail.com

⁵UFBA, 0000-0003-0422-4135, Ivana.lins@gmail.com

⁶UFBA/UFBA, 0000-0002-8071-3494, luisjesusb@bol.com.br

⁷UFBA, 0000-0003-4151-7690, maira.salles@ufba.br

RESUMO O estudo apresentado tem como objetivo central compreender a configuração do domínio da Organização do Conhecimento em nível nacional através dos seus aspectos teórico-epistemológicos. Para tal, realizou-se um levantamento dos trabalhos apresentados nos últimos cinco anos no Grupo de Trabalho 2 (GT2 – Organização e Representação do Conhecimento) do Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB) realizado anualmente no Brasil. A seleção priorizou trabalhos com discussões teóricas e conceituais e possíveis aplicações dos resultados. A escolha se deu pelo fato desses trabalhos apresentarem elementos relevantes para a compreensão do domínio da Organização do Conhecimento. A pesquisa de natureza qualitativa utilizou-se do método dedutivo e indutivo como base para construção das seguintes categorias: títulos; base epistemológica; teorias; metodologia e resultados e que serviram para análise de conteúdo dos textos selecionados. As categorias eleitas foram estruturadas de maneira não hierárquica buscando contemplar a complexidade das teorias e bases epistêmicas apresentadas. Através do mapeamento realizado foi possível identificar a existência de teorias e bases epistemológicas recorrentes que são utilizadas para fundamentação das pesquisas apresentadas no GT2. Ao final da análise, verificou-se a escassez de apresentação de resultados efetivos na aproximação da teoria e prática, ou seja, é incipiente o entendimento sobre a importância da Organização do Conhecimento além do plano das ideias, de forma a estabelecer princípios que norteiem a sua concretização para atender as reais demandas da sociedade.

PALAVRAS-CHAVE *Organização do Conhecimento, Análise de Conteúdo, Análise de Domínio.*

ABSTRACT The research presented is mainly targeted at furthering the understanding about the structure of the domain of Knowledge Organization in Brazil, focusing on its theoretical and epistemological aspects. Aiming such purpose, we conducted a review of the scholarly works presented over the past five years within Working Group 2 (GT2 – Knowledge Organization and Representation) at ENANCIB (the National Meeting of Research about Information Science), yearly held in Brazil. Our sample covered works which addressed both the theory and concepts underlying the subject, as well as the application of outcomes. The identification of elements necessary for understanding the domain of Knowledge Organization was the criterion for selection. The inductive and deductive method was applied in this qualitative research to serve as the framework to establish the following categories: titles; epistemological basis; theories; methodology and outcomes, which were ultimately applied in the content analysis of the papers selected. The categories applied were

structured in a non-hierarchical manner, with the purpose of thoroughly covering the theories and epistemic bases presented, with due regard to their intricate nature. Such mapping allowed identifying recurrent theories and epistemological bases in the reasoning of research presented within GT2. By the end of such review, we have come to the conclusion that actually there were few findings presented which were successful making a connection between theory and practice, thus, little was elucidated about the relevance of Knowledge Organization beyond the realm of ideas, providing principles on how it should be put in practice to meet the society's demands.

KEYWORDS *Knowledge Organization, Content Analysis, Domain Analysis.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Os estudos em Organização e Representação do Conhecimento (OC e RC) no contexto da Ciência da Informação no Brasil vem se ampliando, principalmente nas últimas décadas, impulsionados pela aceleração do uso das tecnologias digitais e sociais. Além disso, estudiosos do domínio da OC preocupam-se em buscar teorias e metodologias que contribuam para a fundamentação de modelos para Sistemas de Organização do Conhecimento (SOC).

Estudos frequentes sobre taxonomias facetadas, navegacionais e dinâmicas, mapas tópicos, indexação automática, ontologias de domínios, dentre outros apontam para bases epistemológicas e teorias que fundamentam os estudos em OC e envolvem o processo mental e individual e os saberes que se desenvolvem continuamente nos variados campos interdisciplinares. Como produtos de representação e organização da informação e do conhecimento, as ferramentas geradas possuem um grau de complexidade bastante desafiador, pois declarar as ideias expressas nos objetos informacionais e individualizá-las, sem contudo se afastar de um conjunto de pertencimento, não é tarefa de fácil compreensão e execução.

A Organização do Conhecimento envolve o trabalho com assuntos dos mais variados domínios de conhecimento, dentre as questões abordadas está a sua concretização, da representação do conhecimento. Esta tarefa envolve um trabalho interdisciplinar com as áreas da Computação, Filosofia, Linguística, Psicologia, Sociologia, Ciência da Informação, dentre outras. Algumas definições sobre os conceitos aqui apresentados estão registradas nas obras de Hjørland (2008) que entende a OC no campo da descrição, indexação e classificação. Diferentemente de Dalberg (2006) que aponta o domínio da Organização do Conhecimento como uma ciência que trabalha com estrutura e organização sistemática de conceitos, considerados pela autora como unidades de conhecimento. No Brasil, autores contemporâneos como Brascher e Café (2008) esclarecem sobre as definições de OC, RC e evidenciam a diferença entre Organização e Representação da Informação e Organização e Representação do Conhecimento.

Este trabalho não objetiva esclarecer ou debater possíveis distorções aos termos, apenas citar tais autores na intenção de delimitar que o interesse dos autores deste trabalho está na Organização e Representação do Conhecimento sob a perspectiva conceitual e visa buscar caminhos para o desenvolvimento da pesquisa em sistemas de OC que efetivamente contribuam para o acesso e representação social da informação visando o acesso e uso da informação na sociedade pela pessoa humana.

Segue-se uma linha de autores contemporâneos como Campos (2003), Lima (2015), Maculan (2014), onde os princípios teóricos e as técnicas empregadas para organização da informação e do conhecimento enfrentam o desafio de atender as reais necessidades dos usuários das redes de acesso a informação, pois segundo Currás (2010, p. 19) “quando se pretende estudar um tema, é geralmente recomendável concretizar as bases sobre as quais deverão se apoiar os raciocínios elaborados”.

Acompanhando esse pressuposto, o mapeamento dos trabalhos eleitos para compor o panorama atual das pesquisas encontram-se apresentados no quadro 1, com categorias de análise pré-estabelecidas. Esta pesquisa privilegiou trabalhos que evidenciassem, em seus títulos e resumos, as propostas que coadunassem com a intenção das pesquisas realizadas pelos autores deste artigo, ou que fornecessem caminhos possíveis a serem explorados no domínio da OC.

A pesquisa aqui exposta está inserida no contexto do projeto *Organização e Representação do Conhecimento: análise conceitual e cognitiva, vocabulários semânticos, taxonomias e ontologias de domínios de conhecimento para Sistemas de Organização do Conhecimento (SOC)* realizado no período de 2015 a 2016, momento em que foi executada a sua fase inicial. Este projeto encontra-se ancorado no âmbito de um projeto maior do grupo de pesquisa *Difusão do Conhecimento e Apropriação de Saberes: Acesso e Uso da Informação na Sociedade (DIFUSAS)*.

O DIFUSAS agrega pesquisas voltadas às Políticas Públicas de Informação e Tecnologias de Informação e Comunicação em espaços culturais e sociais com o objetivo de que a difusão do conhecimento alcance de forma igualitária as pessoas que frequentam esses espaços sociais e, não apenas a comunidade de práticas. Espera-se que domínios como a OC efetivem processos que atendam as variadas comunidades e seus discursos ultrapassem as fronteiras da academia.

Desse modo, o objetivo central deste artigo é demonstrar a compreensão e a configuração do domínio da Organização do Conhecimento em nível nacional, através dos seus aspectos teórico-epistemológicos encontrados nos artigos de autores pesquisadores que apresentam resultados de pesquisas no Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB) do Grupo de Trabalho (GT2) – Organização do Conhecimento. Por meio desta compreensão buscar caminhos para o desenvolvimento de pesquisas em OC e RC que atendam demandas da sociedade visando a efetivação da apropriação de saberes.

Entender as dimensões sociais da Representação do Conhecimento, para além do simples conceito, se tornou uma tarefa e preocupação dos autores deste artigo, há algum tempo, pelo menos por 14 anos, tempo que o DIFUSAS existe enquanto Grupo de pesquisa certificado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ).

As dimensões sociais da Ciência da Informação e a sua relação com a sociedade brasileira vai além do conhecimento científico e encontra-se com a produção do conhecimento humano. E suas representações são de importância fundamental, visto que registram os discursos de sujeitos que constroem estruturas sociais tais como as instituições e mais ainda os seus instrumentos de regulação.

Compreender as instituições e seus instrumentos é também refletir sobre os sujeitos atores da sociedade e as inúmeras facetas que compõem a construção do conhecimento humano. A representação da informação e do conhecimento podem ser reconhecidas como mecanismo que tem a função de desatar as amarras que regulam o desenvolvimento do sujeito, quando viabilizam o acesso e uso da informação, de forma democrática para tornar o sujeito independente em suas escolhas.

O conhecimento humano passa e perpassa pelo contato dos indivíduos e sua co-construção sócio cultural, um estado social que coabita com o estado biológico, caracterizando o sujeito socialmente instituído, sendo assim a interdependência entre os seres humanos é um fator a ser entendido quando se pensa em construir e organizar instrumentos, artefatos e equipamentos que serão responsáveis pela difusão do conhecimento, dando oportunidade a pessoa humana de ressignificar a informação.

A ação de ressignificação se concretiza no movimento da informação nos espaços formais e informais compreendidos pelo conhecimento registrado e o conhecimento adquirido ao longo da vida do sujeito que percebe, e assim transforma, altera ou modifica o conhecimento (NOVO, 2015, p.13)

Sabe-se que conhecimento pressupõe um sujeito que conhece intuitivamente e a ideia que possui de um objeto, provocando uma individualidade advinda da vontade e da necessidade. Para Foucault (1999, p.72) “[...] não há conhecimento verdadeiro senão pela intuição, isto é, por um ato singular da inteligência pura e atenta, e pela dedução que liga entre si as evidências.”. A individualidade por vezes pode prejudicar a ação da difusão. Para socializar o conhecimento e torná-lo representado é preciso entender a manifestação nítida dos objetos, dando-lhes um sentido evidente para o indivíduo.

A partir dessa concepção pode-se entender os sistemas de organização e representação do conhecimento, como instrumentos de mediação ou ainda como processos de comunicação que estabelecem a ponte entre a informação e os usuários que dela necessita, onde só é possível obter a comunicação quando um conteúdo é internalizado. Portanto um SOC deve ser projetado para e pela voz do sujeito e seus discursos enquanto membros de uma comunidade a ser ouvida.

Assim sendo, como objetivo específico, pretendeu-se verificar a configuração da organização intelectual e cognitiva do conhecimento baseado na utilização de conceitos, investigando bases teóricas de fundamentação, métodos e metodologias de construção de ferramentas para (SOC) nos trabalhos apresentados nos últimos cinco anos compreendendo (2011 a 2015) e que ofertassem resultados práticos de aplicação.

Para a análise dos conteúdos trabalhados nos artigos publicados nos anais do evento ENANCIB foram determinadas algumas categorias (base epistemológica; teorias; metodologias e resultados) para perceber o sentido empregado no texto pelo autores e que evidenciarão a intenção do trabalho. Algumas evidências puderam ser percebidas já no resumo dos artigos, outras na introdução sem necessitar da leitura integral do texto, assim através dessa primeira “leitura” foram selecionados artigos que atendiam ao propósito da pesquisa e que serviriam de repositório para outras pesquisas no âmbito da OC e do projeto mencionado anteriormente.

A metodologia se configurou como pesquisa aplicada, com abordagem qualitativa. Demonstra-se um método analítico, a partir da comparação dos conteúdos evidenciados nas categorias (base epistemológica; teorias; metodologias e resultados). Através do resultado coletado pode-se observar que embora os autores citem bases epistemológicas e teóricas em seus trabalhos, ainda é tímida a explanação dos resultados do uso destas teorias em metodologias ou modelos para OC e ou ainda de aplicação prática. Acredita-se ainda que o domínio da Organização do Conhecimento necessita demonstrar explicitamente a aplicação de suas teorias em práticas construídas para atender aos anseios da pessoa humana na atualidade.

A pesquisa foi iniciada com o processo de seleção dos textos do GT2 do ENANCIB atendendo as categorias estabelecidas para a análise. A escolha deste evento se justifica por se tratar de um ambiente

onde grande parte das produções da área da OC são apresentadas nacionalmente, embora se reconheça que alguns trabalhos de OC que atendam aos objetivos do DIFUSAS podem ser encontrados em Grupos de Trabalho com enfoque inter e transdisciplinares a exemplo do GT8 – Informação e Tecnologia, afinal os processos de OC e seus produtos são objeto de investigação de pesquisadores da área de Tecnologia, bem como de outras áreas.

A primeira etapa configurou-se da seleção de textos que abordassem epistemologias, teorias de fundamentação para metodologias e métodos para SOC e que apresentassem resultados práticos, as evidências dessas categorias foram observadas nos títulos dos artigos e/ou nos resumos. Em alguns artigos foi necessária a leitura da introdução do trabalho e, por vezes, realizou-se a leitura integral dos textos.

A compreensão do campo da Organização do Conhecimento através de seus aspectos teórico-epistemológicos é fundamental para os estudos da área por proporcionar uma visão ampliada de sua configuração. O destaque das teorias acaba por demonstrar a maneira com a qual elas vem sendo aplicadas nas práticas de construção de Sistemas de Organização do Conhecimento (SOC), no vasto domínio da OC que inclui disciplinas da ciência da Computação, Ciências Cognitivas, Psicologia, Linguística, Semiótica, Filosofia dentre outras.

Trata-se de uma pesquisa a nível descritivo onde pretendeu-se “[...] descrever as características de determinada população ou fenômeno ou o estabelecimento de relações entre variáveis.” (Gil, 1991, p. 31). O método de pesquisa empregado foi o documental, a abordagem do problema, teve o caráter qualitativo, pois este tipo de estudo favorece a interpretação dos dados ou fenômenos, bem como a atribuição de significado e a compreensão de aspectos subjetivos dos documentos analisados.

Como técnica de coleta de dados utilizou-se a análise de conteúdo através das categorias estabelecidas nos artigos selecionados. Segundo Bardin (2010), análise de conteúdo é:

[...] um conjunto de técnicas de análise das comunicações visando obter, por procedimentos, sistemáticos e objetivos de descrição do conteúdo das mensagens, indicadores (quantitativos ou não) que permitam a inferência de conhecimentos relativos às condições de produção/recepção (variáveis inferidas) destas mensagens. (Bardin, 2010, p.44)

Visando a compreensão da ideia do autor e de sua intenção para além do discurso que está explícito, a análise de conteúdo é uma metodologia empregada para entender o que é essencial, selecionando-os em torno das ideias principais e produzindo inferências.

[...] produzir inferências sobre o texto é a razão de ser da análise de conteúdo, significa, não somente produzir suposições subliminares acerca de determinada mensagem, mas em embasá-las com pressupostos teóricos de diversas concepções de mundo e com as situações concretas de seus produtores ou receptores. (Campos, 2004, p. 613).

Entretanto, convém esclarecer que esta metodologia visa ainda, como dito por Minayo (1996) ultrapassar o subjetivismo da interpretação e buscar uma interpretação crítica dos documentos analisados para fins de atender a expectativa da pesquisa.

Seguindo os passos metodológicos, neste estudo, as informações retiradas dos artigos foram organizadas no Quadro 1 onde encontram-se as categorias de análise e os dados sistematizados em uma matriz que teve por objetivo demonstrar visualmente o resumo dos achados representativos das

categorias. A partir desses achados é que se estruturou a análise e a interpretação da mensagem expressa nos discursos dos pesquisadores.

RESULTADOS

Os conteúdos dos trabalhos para compor as categorias propostas foram extraídos e analisados no intuito de compor uma síntese que represente a respectiva categoria de maneira mais completa possível, fazendo com que o resultado de cada uma delas se configure como a expressão do que cada autor utilizou de maneira central para o domínio proposto e ou para resultados de aplicações de ferramentas e conteúdos de construção de SOC e ainda novas bases teóricas.

Quadro 1 Categorias estabelecidas para análise dos artigos

ENANCIB ANO/EDIÇÃO	TÍTULO	BASE EPISTEMOLÓGICA	TEORIAS	METODOLOGIA	RESULTADO
2011/12°	A importância dos pressupostos Ontológicos como base para o uso articulado de ontologias no contexto da web semântica	Uma mescla de abordagens no intuito de fundamentar as ontologias – Através de estudos de autores contemporâneos	Teoria do Conceito e Teoria da Terminologia.	Descrição das ontologias fundamentais.	Descrição sumária dos conceitos básicos que estão nas ontologias.
2011/12°	A teoria clássica de categorização e os princípios de Ranganathan: uma análise teórica	Aristóteles; Ranganathan	Teoria clássica de categorização (Aristóteles); princípios categoriais (Ranganathan)	Análise teórica da teoria clássica de categorização e os princípios categoriais de Ranganathan mediante estudo bibliográfico.	Apesar da proposta distinta entre as duas teorias, a percepção do que constitui a teoria ranganathiana é visivelmente inspirada nas categorias aristotélicas.
2011/12°	Tesaurus conceituais e ontologias de fundamentação: análise comparativa entre as bases teórico-metodológicas utilizadas em seus modelos de representação de domínios	Princípios da Filosofia Ciências cognitivas	Teoria do conceito (Dahlberg) Teoria da Classificação facetada (Ranganathan) Unified Foundational Ontology (GUIZZARDI)	Análise comparativa das bases teóricas para a construção de tesaurus e elaboração de ontologias de fundamentação.	A Ciência da Informação dispõe de base teórica e metodológica próprias para a construção de instrumentos terminológicos. Deixa explícito as diferenças entre modelos e instrumentos de representação de domínio.

2012/13 ^o	Prototipagem de banco de dados: o uso da teoria da classificação facetada na modelagem de dados	Ranganathan	Teoria da classificação facetada (TCF)	Pesquisa aplicada; exploratória; bibliográfica	Conclui-se que os subsídios para o desenvolvimento do Sistema Facetado foi a TCF e a modelagem de dados em BD.
2012/13 ^o	Modelo e teorias para representação: uma teoria ontológica sobre o sangue humano	Diversos autores que trabalham o conceito de modelos teóricos.	Teoria ontológica	Estudo na literatura sobre os conceitos de modelo e teoria.	Proposta de teoria ontológica sobre o sangue humano no contexto de um projeto de OI biomédica.
2012/13 ^o	Sobre o processo de construção social da “linguagem natural” e suas implicações para o domínio de organização do conhecimento	Bakhtin	Teoria do conceito; teoria geral da terminologia discutidos por autores contemporâneos.	Infere-se que seja pautado em revisão de literatura.	Análise da terminologia específica de uma especialidade pode revelar as correntes de pensamento, as quais tem se recorrido predominantemente e para construir a base teórica de um domínio científico.
2013/14 ^o	Interconexões entre a teoria da classificação facetada (TCF) de Ranganathan e o modelo entidade-relacionamento (MER) de Peter Chen	Ranganathan Peter Chen	Teoria da classificação Facetada; Modelo Entidade Relacionamento (Peter Chen)	Abordagem qualitativa e aplicada, usando métodos canônicos da CI e CC a exemplo dos sistemas de recuperação da informação (SRI)	Caso a TCF e o MER forem usados conjuntamente, a modelagem de um domínio terá um nível maior de excelência
2013/14 ^o	Conversão de tesouros em ontologias: um estudo exploratório	Dahlberg e autores contemporâneos	Teorias da Representação; Teoria da Recuperação da Informação; Teoria da Classificação Facetada; Teoria do Conceito; Teoria Geral da Terminologia.	Análise de conteúdo	A natureza das relações e entidades é semelhante, e a condução do processo pode ter como princípios norteadores as indicações de relacionamentos, nos manuais e padrões de construção de tesouros e de ontologias.
2013/14 ^o	Interfaces entre ontologias e conceitos seminais da	Autores contemporâneos de ontologies e OC	Ontologias de domínio; Ontologias genéricas.	Ontologia aplicada Infere-se uso de	Caracterização das ontologias oferecendo alternativas das

	ciência da informação: em busca de avanços na organização do conhecimento			revisão de literatura.	ontologias na OC.
2014/15 ^o	Pragmática na organização do conhecimento	Não explícita	Pragmática Linguística.	Análise na literatura; Análise documentária	As relações semânticas podem desempenhar um papel crucial nessa comunicação, ao guiar o usuário a partir do seu ponto de vista inicial até aqueles defendidos pela frente de pesquisa
2014/15 ^o	Proposta de modelagem conceitual a partir da implementação da teoria da classificação facetada de Ranganathan em banco de dados	Ranganathan	Teoria da classificação Facetada (Ranganathan)	Pesquisa Aplicada, Exploratória e Bibliográfica sobre a TCF de Ranganathan na modelagem conceitual de dados Método Dissecção e Desnudação; técnica Laminação.	O modelo de Mapa categorial Proveniente da Laminação Apresentou uma redução na quantidade de conceitos; O modelo gerado pela Desnudação, ao oferecer Diferentes pontos de vista (focos) sobre a mesma ocorrência apresentada nos resultados de busca, teoricamente, permite ao usuário maior autonomia de navegação.
2014/15 ^o	Teoria da terminologia e teoria da ontologia: um comparativo baseado na criação de definições para termos	Ontologias Wüter e autores contemporâneos	Teoria da terminologia; teoria da ontologia aplicada	Infere-se um estudo comparativo entre as teorias da ontologia.	Espera-se contribuir para um melhor entendimento da aplicação das duas teorias, bem como fornecer subsídios para auxiliar ao Cientista da

					Informação no uso das mesmas.
2015/16 ^o	Web semântica: introdução a recursos de visualização de dados em formato gráfico	Não identificado	Não identificado	Procedimentos exploratórios da terminologia para analisar tais tecnologias.	Verificou-se há poucos recursos disponíveis para apresentação de dados em formato gráfico de ambientes baseados em tecnologias da Web Semântica
2015/16 ^o	Mapeamento temático das comunicações orais do GT2: uma análise por instituições	Não identificado	Não identificado	Análise conceitual; Análise de conteúdo.	Mapeamento conceitual do domínio com alto índice de dispersão temática por instituições, identificação das instituições mais produtivas e maior concentração de pesquisas na região Sudeste do país.
2015/16 ^o	Organização do conhecimento na perspectiva do modelo dinâmico de análise conceitual	FOUCAULT – As palavras e as coisas – Representação. Análise do discurso representativo.	Teoria do conceito; teoria da classificação facetada	Análise cognitiva e conceitual.	Modelo Dinâmico de Análise Conceitual (Modac)

Fonte: Elaborado pelos autores.

ACHADOS DA PESQUISA

Dos cinquenta (50) trabalhos levantados a partir do que foi estabelecido, ou seja, pesquisas que evidenciavam bases epistemológicas, teóricas, metodologias e apresentavam evidências de resultados a partir das categorias de análise citadas (quadro1) foram selecionados quinze (15) artigos, sendo três (3) por ano analisado, compreendidos entre 2011 a 2015 para efeitos de discussão neste trabalho.

O conteúdo das mensagens dos produtores de conhecimento da OC no Brasil registrados nos artigos apresentados nos ENANCIB, chamados neste trabalho de achados da pesquisa, foram analisados conforme exposto na figura 1 – Categorias Analisadas, a seguir:

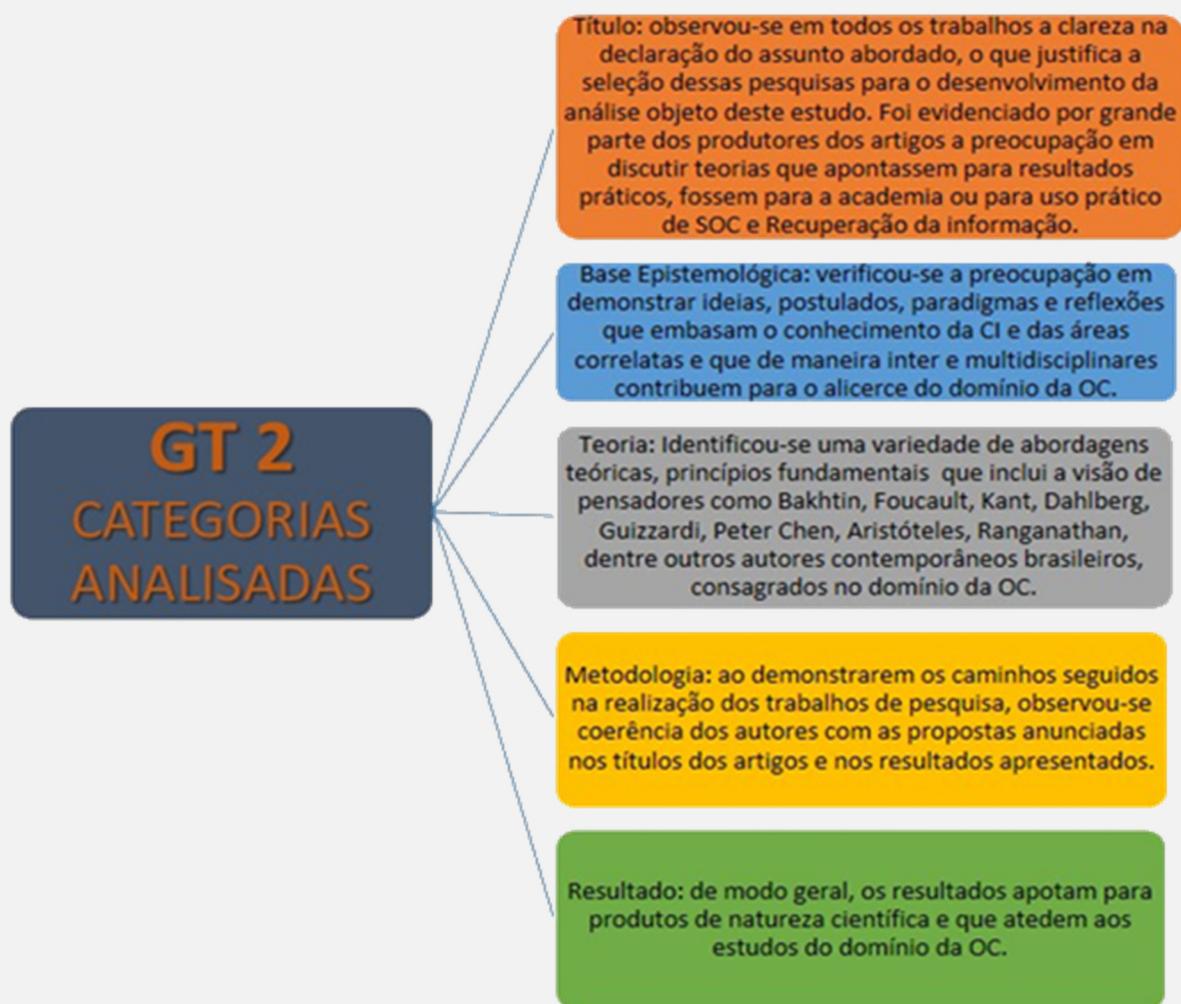


Figura 1. Categorias analisadas

Fonte: Elaborada pelos autores

A pesquisa buscou encontrar evidências de discussões na Ciência da Informação que contribuam para aprimorar e ou criar ferramentas, métodos, metodologias, modelos e resultados que viabilizem construção de SOC. Através de uma análise mais detalhada dos trabalhos científicos oriundos do GT2, foi possível identificar uma predominância na utilização da Teoria da Classificação Facetada de Ranganathan seguida da Teoria do Conceito de Dahlberg.

Como metodologia identificou-se a utilização da Análise de Domínio dentre outros métodos referentes a elaboração de Linguística documentária, ontologia aplicada, Análise comparativa, Análise de Conteúdo, Análise Conceitual, Análise Cognitiva. No entanto, no conjunto que totaliza os documentos analisados, percebeu-se por vezes a ausência de declaração explícita da metodologia utilizada nas pesquisas construídas pelos autores.

Cabe notar que as bases epistemológicas que fundamentam as pesquisas variam de acordo com a necessidade dos autores em demonstrar estudos voltados à teorias e ainda a metodologias e modelos. As abordagens são apresentadas hora de maneira mais geral e em outros momentos com um foco mais instrumentalizado e em alguns trabalhos percebeu-se que as bases epistemológicas e teóricas não são declaradas.

Os títulos dos trabalhos são pertinentes ao GT2 quanto a abordagem – Representação e Organização da Informação e do Conhecimento e ainda ao conteúdo tratado na pesquisa dos autores e, de forma geral, esclarecem e refletem a temática descrita nos artigos apresentados.

Quanto a análise da categoria referente aos resultados das pesquisas observou-se que alguns trabalhos, embora apontem resultados, não são de aplicações práticas referentes ou relacionados as bases epistemológicas e teóricas enquanto fundamentadoras de metodologias, modelos ou métodos de pesquisa que possam assegurar a evolução de Sistemas de Organização de Conhecimento (SOC).

Cabe ressaltar que as categorias estabelecidas para a avaliação a que este trabalho se propôs foram úteis e fundamentais para guiar a análise dos trabalhos apresentados no GT em questão, sem que houvesse a necessidade da leitura integral dos textos e guiou os autores para os achados desta pesquisa e que serviu de um referencial importante para trabalhos em andamento e propostas de possíveis investigações futuras para o grupo Difusas.

CONCLUSÕES

As pesquisas em OC no Brasil vêm se consolidando e aprimorando estudos de teorias e metodologias, principalmente no âmbito de análise de conteúdo e ontologias que resultam em modelos de análise, mapeamentos conceituais e modelização de domínios de conhecimento.

A partir das verificações estabelecidas nos processos de discussão deste artigo, pode-se inferir para a carência referente aos modelos de Análise conceitual e formatos de representação do conhecimento que atendam as tecnologias usuais da Web Semântica e ainda que os trabalhos do GT2 analisados demonstraram discussões a respeito dos objetos de estudo propostos, porém, evidenciou-se nos resultados a necessidade de caminhar para aplicações efetivas que atinjam o âmbito social da Ciência da Informação. Percebeu-se que modelos e metodologias propostas estão concentrados no atendimento de uma fatia seletiva do campo científico.

No entanto, como o objetivo do grupo de pesquisa Difusas é o comprometimento com o acesso e uso da informação na sociedade. Acredita-se que pesquisas em Organização Social do Conhecimento que alcance propostas para uma sociedade plural devam contemplar *práxis* voltadas para acessibilidade e diversidade da pessoa humana. É fundamental destacar que não se levou em consideração as implicações políticas e institucionais neste artigo, contudo, não se desconsidera a possibilidade de uma análise destes fatores em um estudo futuro.

Durante o processo, constatou-se a diversidade de teorias utilizadas na fundamentação dos trabalhos que seguem bases epistemológicas e cujos resultados atedem aos pesquisadores de OC. Em geral, o que se pode inferir foi que não se demonstram na maioria dos resultados das pesquisas a efetiva aplicação das teorias estudadas ou apontadas pelos autores e ainda resultados práticos. Além disso nota-se que as teorias ou bases epistemológicas se repetem.

Nesse sentido, a análise enfrentou a existência de duas dimensões epistemológicas que estão interligadas. A primeira relacionada as ideias que estão colocadas de maneira mais claras, onde é possível compreender o discurso dos autores em relação as teorias utilizadas de maneira direta, isso configura os aspectos mais comuns e transversais entre os trabalhos, o que se reflete na prática da escrita através da utilização dos autores mais recorrentes no domínio da OC. A segunda, por sua vez mais

profunda, implica em uma dispersão maior das teorias, gerando uma linha tênue entre os conceitos apropriados pelos autores e suas próprias ideias. Até o momento essa identificação e reflexão carecem de uma análise mais detalhada.

Os resultados indicam que existe a intenção de buscar soluções práticas para a evolução de SOC, principalmente com os avanços da web semântica, no entanto é preciso fomentar discussões sobre ferramentas de inclusão e ainda que atendam aos anseios da sociedade como um todo, principalmente quando as questões sociais, culturais, educacionais vêm sendo modificadas e traçadas respondendo aos interesses políticos. Cabe aos Cientistas da Informação intervir de forma efetiva em seus estudos e práticas para que a representação da informação contribua para o acesso e uso dos itens informacionais, de forma esclarecedora. Os autores deste artigo pertencentes ao grupo de pesquisa Difusas vem buscando colaborar para que a pessoa humana tenha acesso a informação de forma plural, acredita-se que estudos nessa linha devem ser ampliados na CI.

Este texto foi construído e pensado com base nas vozes de outros sujeitos além dos autores e os pesquisadores da CI. Ouvimos os sujeitos que frequentam espaços sociais e culturais como bibliotecas e arquivos públicos e universitários, enquanto ocupantes de cargos de gestão, pois entende-se que para o trabalho de representação alguns elementos são fundamentais e se relacionam: O objeto, signo que o representa e a cognição humana, afinal, interpretá-se e representa-se um objeto também pela experiência que pode ser chamada de conhecimento.

REFERÊNCIAS BIBLIOGRÁFICAS

Alvares, Lillian (Org.). (2012). *Organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações*. São Paulo: B4 Editora.

Bardin, Laurence. (2010). *Análise de conteúdo*. Lisboa: Edições 70.

Brascher, M. & Café, L. (2008) Organização da informação ou organização do conhecimento? *Anais do Encontro Nacional de Pesquisa em Ciência da Informação*, São Paulo, SP, Brasil, 9, Recuperado em 6 de junho de 2017, de <http://skat.ihmc.us/rid=1KR7TM7S9-S3HDKP-5STP/BRASCHERCAF%C3%89>

Campos, Claudinei José Gomes. (2004). Método de análise de conteúdo: ferramenta para a análise de dados qualitativos no campo da saúde. *Revista Brasileira de Enfermagem*, Brasília, 57, (5), 611- 614, Recuperado em 16 outubro, 2016, de <http://www.scielo.br/pdf/reben/v57n5/a19v57n5.pdf>.

Campos, Maria Luiza de Almeida, Souza, Rosali Fernandez & Campos, Maria Luiza Machado. (2003). Organização de unidades de conhecimento em hiperdocumentos: o modelo conceitual como espaço comunicacional para realização da autoria. *Ciência da Informação*, Brasília, 3, (2), 7-16, Recuperado em 15 de junho, 2017, de <http://www.scielo.br/pdf/ci/v32n2/17029.pdf>.

Currás, emilia. (2010). *Ontologias, taxonomias e tesouros em teoria de sistemas e sistemática*. Brasília: Thesaurus.

Dahlberg, I. (2006). Knowledge organization: a new science? *Knowledge organization*, 33, (1).

Foucault, Michel. *As palavras e as coisas: uma arqueologia das ciências humanas*. São Paulo: Martins Fontes, 1999.

Gil, Antonio Carlos. (1991). *Como elaborar projetos de pesquisa*. São Paulo: Atlas.

Hjørland, B. *Knowledge organization systems*. (2008). Available from: Cite: 27 May.

Lima, Gercina Angela de. (2015). *MHTX: modelagem hipertextual para organização de documentos: princípios e aplicação*. Rio de Janeiro: Interciência.

Maculan, Benildes Coura M. Dos. (2014). *Taxonomia facetada e navegacional: um mecanismo de recuperação*. Curitiba: Appris.

Minayo, Maria Cecília de Souza. (1996) . *O desafio do conhecimento: pesquisa qualitativa em saúde*. (4a.ed.). São Paulo: Hucitec-Abrasco.

Novo, Hildenise Ferreira. *Análise conceitual e cognitiva: Modac - um modelo dinâmico para auxiliar a construção de Sistemas de Organização do Conhecimento (SOC)*. Salvador, 2014. Tese (doutorado em Difusão do Conhecimento, UFBA, 2014.



AS CONTRIBUIÇÕES DA TERMINOLOGIA PARA A ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO DE TEXTOS NARRATIVOS DE FICÇÃO

Mariana da Silva Caprioli¹, Larissa de Mello Lima², Walter Moreira³,
João Batista Ernesto de Moraes⁴

¹Unesp - Universidade Estadual Paulista - Faculdade de Filosofia e Ciências PPGCI,
0000-0002-4206-8746, mariana.caprioli@gmail.com

²Unesp - Universidade Estadual Paulista - Faculdade de Filosofia e Ciências PPGCI,
0000-0002-5377-9535 - larissalima.unesp@gmail.com

³Unesp - Universidade Estadual Paulista - Faculdade de Filosofia e Ciências PPGCI,
0000-0001-9454-441X, walter.moreira@marilia.unesp.br

⁴Unesp - Universidade Estadual Paulista - Faculdade de Filosofia e Ciências - PPGCI,
0000-0002-8907-3593 - prof.joao@gmail.com

RESUMO No contexto da Organização e Representação do conhecimento, acredita-se que a terminologia tem potencial para auxiliar o processo de representação de textos narrativos de ficção. Parte-se da problemática de que a representação de textos literários é realizada, muitas vezes, sem considerar as singularidades inerentes aos textos narrativos de ficção resultando em problemas na representação do conteúdo dos mesmos. O objetivo deste trabalho é realizar uma revisão de literatura no cenário brasileiro sobre Terminologia, Organização e Representação do Conhecimento e Textos narrativos de ficção a fim de mostrar como estes temas estão interligados. Para tanto utilizou-se a base de dados BRAPCI e Google Scholar combinando operadores booleanos para encontrar publicações que interligam as temáticas acima citadas a fim de entender como elas estão sendo tratadas na atualidade. Como conclusão foi possível perceber que as três áreas pesquisadas estão diretamente ligadas e que quando relacionadas de modo metodologicamente preciso, podem promover qualitativamente a organização e a representação do conhecimento, com ganhos em termos de eficácia para o usuário no momento da recuperação da informação.

PALAVRAS-CHAVE *Terminologia, textos literários, ORC.*

ABSTRACT In the context of Organization and Representation of Knowledge, it is believed that terminology has the potential to assist the process of representing fictional narrative texts. Start with the problem is that the representation of literary texts is often carried out without considering the singularities inherent in narrative fiction texts resulting in problems in the representation of their content. The objective of this work is to perform a literature review in the Brazilian scenario on Terminology, Organization and Representation of Knowledge and Narrative Fiction Texts in order to show how these themes are interconnected. For this purpose, the BRAPCI and Google Scholar databases were used, combining Boolean operators to find publications that interconnect the above themes in order to understand how they are being treated today. As conclusion it was possible to perceive that the three areas studied are directly linked and that when related in a methodologically precise way, can qualitatively promote the organization and representation of knowledge, with gains in terms of effectiveness for the user at the time of information retrieval.

KEYWORDS *Terminology, Literary texts, ORK.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

As linguagens documentárias (LDs) atuam como instrumentos de intermediação para que a representação da informação possa ser feita, uma vez que uma das etapas da indexação refere-se à atribuição de termos de uma linguagem controlada para representar um documento em determinado sistema de informação.

O papel da Terminologia teórica, seguindo a corrente que interessa a esse estudo enquanto uma discussão situada no âmbito da Ciência da Informação, é fornecer referência para interpretação de termos do modo como são expressos e definidos no domínio das terminologias concretas. Então, a Terminologia serve de apoio às LDs para a construção de instrumentos terminológicos que visem a representação eficaz da informação em unidades de informação.

Sendo assim, a Terminologia pode servir de apoio para a representação de textos literários, considerados de modo geral, mas principalmente os textos narrativos de ficção. Os textos narrativos de ficção requerem um olhar diferenciado e atento nas questões que envolvem a representação, conforme será apresentado posteriormente. Esta pesquisa assenta-se no pressuposto de que a construção de instrumentos, como tesouros, por exemplo, pode colaborar com a representação desses textos que são, normalmente, tratados de modo inadequado em unidades de informação. Acredita-se que este trabalho se justifica na medida em que busca mostrar como a Terminologia extrapola a perspectiva de apoiar as LDs para a construção de vocabulários controlados relacionados aplicados à organização e representação de textos científicos exercendo também um papel fundamental de auxílio no tratamento temático de textos narrativos de ficção.

Para o trabalho, optou-se por uma revisão de literatura em periódicos científicos brasileiros da Ciência da Informação, conforme detalhamento apresentado na seção subsequente. Esta delimitação geográfica visa atender aos interesses imediatos da pesquisa, isto é, compreender o contexto brasileiro de tratamento da temática e, ao mesmo tempo, permitir a identificação de lacunas e de novas categorias de análise entre os temas relacionados para a ampliação da cobertura em estudos posteriores.

METODOLOGIA

Realizou-se uma pesquisa exploratória por meio de uma revisão de literatura na Base de Dados Referenciais de Artigos e Periódicos em Ciência da Informação (BRAPCI)¹ Trata-se de uma base brasileira que tem como principal objetivo subsidiar estudos na área da Ciência da Informação. A BRAPCI aplica-se também à ampliação do espaço de pesquisa, facilitando a visão de conjuntos de produções na área e, ao mesmo tempo, revelando especificidades do domínio científico. Assim, a base foi escolhida para o levantamento dos dados relativos ao presente artigo por sua especificidade temática,

¹ <http://www.brapci.ufpr.br/brapci/index.php/main/about>.

por sua cobertura nacional e pelo nível de confiança que possui entre pesquisadores brasileiros, sendo utilizada em diversos trabalhos que envolvem a pesquisa bibliográfica.

Utilizou-se, também, o *Google Scholar*², ou Google Acadêmico, que se trata de um sistema de busca desenvolvido pela Google, com a intenção de oferecer ferramentas específicas para que pesquisadores busquem literatura acadêmica como artigos, teses, livros etc. Essa ferramenta foi escolhida por ser de ampla cobertura sobre vários temas, com a intenção, então, de obter o maior número possível de material para o levantamento.

Para a composição da estratégia de busca, foram utilizadas as seguintes expressões: “terminologia”; “linguagem documentária”; “organização do conhecimento”; “representação da informação”; “organização e representação do conhecimento”; “textos literários”; “terminologia AND organização e representação do conhecimento”; “organização e representação do conhecimento AND textos literários” e “terminologia AND textos literários”, sendo selecionados, trabalhos de publicação brasileira para se observar como os temas são tratados e trabalhados no país, porém em alguns momentos se encontra a necessidade da utilização de literatura internacional como, por exemplo, se um trabalho indexado em base de dados brasileira tivesse como autor um pesquisador estrangeiro, este mesmo não seria desconsiderado na análise. O levantamento bibliográfico contemplou o período compreendido entre 1990 a 2016.

Foram identificados, em média, 60 textos com a estratégia de busca, valendo ressaltar que quando se trata da combinação entre Representação e Organização do conhecimento e Textos Literários os resultados são poucos e pouco relevantes com um média de 5 textos, porém, separados, o mesmo não acontece, podendo, assim, ser traçado um paralelo interessante. O que não ocorre com os termos “terminologia”; “linguagem documentária”; “organização do conhecimento”; “representação da informação”; “organização e representação do conhecimento”; “textos literários”; “terminologia AND organização e representação do conhecimento”; “organização e representação do conhecimento AND textos literários” e “terminologia AND textos literários”, os quais existindo uma grande extensão de textos a serem recuperados. Neste artigo, observou-se que mais da metade dos textos se encontram relacionados a estas palavras-chaves.

Moraes (2012) em trabalho publicado na 12ª Conferencia Internacional ISKO Índia, identifica e relata a falta de metodologias destinadas à análise e representação de assuntos quando se trata de literatura de ficção. Este “silêncio” é maior quando comparado com que se possui nesta temática em relação à literatura científica, mostrando que o campo da Ciência da Informação, aparentemente, não dá a atenção necessária para a literatura, tratando-a apenas como lazer e diversão. Com tal texto observado, o presente artigo encontra bases para o levantamento de trabalhos que dizem respeito à representação de textos de ficção no contexto brasileiro.

Para a consecução da revisão de literatura pretendida, foram observados os seguintes passos: 1º - identificação de palavras-chave ou descritores e organização das estratégias de busca; 2º - revisão das fontes secundárias; 3º-recolhimento das fontes primárias e 4º - leitura analítica da literatura.

² <https://scholar.google.com.br/>.

TERMINOLOGIA COMO INSTRUMENTO DAS LINGUAGENS DOCUMENTÁRIAS PARA A REPRESENTAÇÃO DO CONHECIMENTO

Pinho, Nascimento e Melo (2015) observam que no âmbito da Ciência da Informação existe uma ambiguidade em relação ao uso dos termos “Representação da Informação” (RI) e “Representação do Conhecimento” (RC), e é importante explicar a diferença entre os termos.

Partindo disso, tem-se que a RC, segundo Dahlberg (2006), se trata de uma estrutura lógica da representação conceitual e que também é o resultado da identificação de conceitos por determinados termos em função da terminologia usada. Já a RI consiste na descrição física e de conteúdo de um objeto informacional. Então, na área da Ciência da Informação, a representação do conhecimento possui dois aspectos distintos: “o resultado da representação de conteúdo pela identificação de conceitos e a representação da estrutura lógica do conhecimento”.

Barité (2001) afirma que a OC tem a intenção de oferecer subsídios teóricos para as questões relacionadas com o tratamento da informação, em particular o tratamento da informação temática. Ainda, para Esteban Navarro (1996), a OC se caracteriza por uma disciplina que se dedica ao estudo de fundamentos teóricos de tratamento e recuperação da informação. Então, esses estudos segundo Martins; Moraes (2012, p. 178-179) “contemplam construção, manutenção, uso e avaliação dos instrumentos lógico-linguísticos mais adequados para controlar os processos de representação, classificação. Ordenação e armazenamento do conteúdo informático dos documentos”.

Hjørland (2008) já afirmava o que foi dito anteriormente, destacando que a OC, diz respeito à natureza e à qualidade de processos de OC., também dos sistemas de OC., os quais são usados para organizar e representar documentos, palavras e conceitos.

Vistos tais aspectos, observa-se que a OC tem a intenção de compreender como a representação do conhecimento se utiliza de instrumentos da Linguagem Documentária para que os documentos sejam representados e recuperados de forma eficiente.

Cintra et. al. (1996), propõe que a Linguagem Documentária é composta por termos descritores que definem a forma de entrada e busca a serem usadas pelo indexador ou usuário em um sistema. Tratam-se de escolhas feitas em um universo lexical amplo, ocorrendo no âmbito de categorias de assunto para intermediar as inconsistências do texto e a representação da linguagem natural para a linguagem documentária, ou especializada. Representa os assuntos e classes de assuntos, mas não textos individuais. Pode não ter serem tratados com clareza e “a remessa de assuntos e classes de assuntos acaba sendo aleatória” (CINTRA, 1996, p.19) em um universo amplo.

Cintra et. al (1996) explica que na Terminologia uma palavra designa determinado objeto com propriedades e características, sendo remetido a certos universos de valores consubstanciados nos discursos de especialidade. A terminologia tem como objetivo principal organizar e harmonizar as noções, ou conjunto deles dos domínios específicos do conhecimento. Por meio de procedimentos sistemáticos seleciona e cria termos para as noções, e os relaciona entre si por meio de definições, como visto anteriormente. E é dessa forma que se obtém repertórios ou listas de termos especializados de um domínio, acompanhados de definições que remetem o termo ao seu referente.

Em momentos mais recentes, a terminologia tem servido para aplicações mais práticas e próximas da realidade e do cotidiano da sociedade (DIAS, 2000). Com a alta expansão da informação, a diversidade

de termos técnicos e científicos vindos de novas tecnologias aumentou, fazendo crescer também a necessidade de comunicação internacional com mais eficiência, maior facilidade e rapidez na recuperação da informação que se encontra espalhada por diversos banco de dados, arquivos e outros meios eletrônicos, acreditando que se faz necessário o desenvolvimento de sistemas mais efetivos e avançados de organização e gestão de informação, baseando-se em metodologias de processamento de dados terminológicos (DIAS, 2000, p.91-92).

Pereira e Bufrem (2005) apresentam um quadro com diversas características para a Análise Documentária, pertinentes em relação ao trabalho intelectual do indexador em um sentido amplo, sendo uma delas a Representação:

PRINCÍPIOS: 7. Representação

CARACTERÍSTICAS: Análises – sínteses – representações. Por meio de regras e instrumentos, tais como: Códigos, Diretrizes e Normas. Ex.: códigos de classificação; vocabulários controlados (thesaurus, CDD, CDU); linguagens documentárias (thesaurus, terminologias, listas de cabeçalhos); política do sistema (PEREIRA; BUFREM, 2005, p.34).

Segundo Maimone e Tálamo (2011), o principal objetivo da Terminologia é observar os discursos especializados, com a intenção de construir dicionários e glossários especializados, ou seja, a “Terminologia é veículo de conhecimento, aspecto importante para a descrição e recuperação da informação” (LARA; TÁLAMO, 2007, p. 2).

As mesmas autoras ainda definem que dicionários terminológicos, vocabulários e glossários se tratam de produtos terminológicos, enquanto os outros produtos documentários ficam por conta das LDs, que se utilizam dos produtos terminológicos para a elaboração e utilização.

Então, até aqui, pode-se observar que a Terminologia contribui significativamente para a organização e representação documentária, uma vez que contribui metodologicamente na identificação das relações *conceito-termo*, *termo-conceito* e *sistemas de conceitos* (MAIMONE; TÁLAMO, 2011).

O trabalho terminológico se evidencia quando se compreende que o termo ou unidade lexical terminológica adquire status de descritor ao se inserir em uma linguagem documentária, assim, podendo mostrar a importância da Terminologia na representação da informação com a construção de tesouros, por exemplo.

Então, a terminologia se faz presente no momento de elaboração de instrumentos nas Linguagens Documentárias, para que a representação do conhecimento seja feita e, assim, corrobora o que se pensa inicialmente nesse trabalho, de que possa ajudar na representação literária, colaborando não apenas para representar de modo eficiente, como também para que a informação seja recuperada e utilizada pelos usuários, sempre se preocupando com suas necessidades.

REPRESENTAÇÃO DE TEXTOS LITERÁRIOS E O APOIO DA TERMINOLOGIA

Quando se trata de representação em unidades de informação, deve-se observar que o documento, objeto da Análise Documental, não se trata apenas de textos científicos, pois existem bibliotecas públicas e

escolares, por exemplo, onde o texto narrativo é o maior foco e objeto de especial atenção, seja na organização par ao atendimento do usuário, ou no desenvolvimento de coleções (MORAES 2011).

Quando se pensa nos textos narrativos de ficção como “matéria prima” para a Análise Documental, a complexidade de representação aumenta consideravelmente, uma vez que o modelo metodológico de identificação de conceitos é elaborado pautado na leitura documental de partes canônicas do texto, como o título, subtítulo, resumo, palavra-chave, etc, e tais aspectos são muitos presentes em textos científicos, não se aplicando, então, a textos narrativos, os quais, por sua própria natureza, assumem outras formas para evidenciar seu conteúdo. Moraes (2011) coloca que o texto narrativo de ficção pode ser entendido com uma obra de arte, a obra literária, que se caracteriza pelo meio verbal, e tem seu objeto artístico composto por palavras que são reveladas e entendida pela capacidade de decodificação do sujeito.

Há uma conclusão feita por Moraes (2011) no que tange ao texto narrativo de ficção explicando que este tipo de texto é aquele que é escrito pensando na seleção de temas que melhor se adéquem e reforcem o contexto sugerido pelo conteúdo semântico da obra, ou em outras palavras, o texto de ficção é uma manifestação artística, onde o artista (autor) se supre de palavras para criar sua obra de arte, “buscando com esta fornecer um novo olhar para as situações humanas, ou mesmo criando uma realidade que pode ser chamada de paralela” (MORAES, 2011, p. 21). Uma diferença dos textos de ficção é que a escolha dos termos empregados durante sua extensão se dá para causar um efeito de nebulosidade para o que seria o sentido real do que está escrito, servindo até mesmo para sugerir múltiplos e imaginários significados.

Para finalizar, Moraes (2011) faz uma colocação reflexiva sobre o assunto, que encerra de maneira poética, ele diz que então, após todas as colocações feitas, pode-se afirmar que a ficção se trata de uma forma de arte construída com palavras, que vai muito além de ser apenas escritos originários da imaginação do escritor.

Com isso, é possível observar que existem incompletudes e desvios na representação de obras literárias, visto que as unidades de informação de modo geral, representam os assuntos a partir de metodologias que se aplicam de modo mais preciso às estruturas mais rígidas, ou mais facilmente identificáveis dos textos científicos.

Diante disso, Anjos (2008) expõe que existem avanços nas discussões da análise de textos narrativos de ficção, mostrando que, realmente, análise e classificação de ficção tem tido uma relação historicamente pobre, mas recentemente isso tem mudado, com a criação de sistemas de recuperação inovadores, como o *Bookhouse* e *Book Forager* (<http://branching-out.net/forager>)³ que tem a intenção de tratar as questões de significado e análise de assunto na ficção. A autora acredita que tais mudanças de atitude em relação à ficção, por bibliotecários e especialistas, no momento da recuperação estão ligadas ao tratamento mais generalista da ficção dentro da academia e da sociedade contemporânea.

Bibliotecários aparentam ter cada vez mais interesse, conforme a informa Anjos (2008), em sistemas de recuperação do conhecimento que facilitem o acesso à literatura de ficção. A facilidade talvez seja influenciada pelos instrumentos sofisticados de recuperação utilizados por livrarias *online* como a

³ O link e a página não estão mais em funcionamento, mas pode-se acessar <<https://web.archive.org/web/20071109184925/http://www.whichbook.net:80/>> para uma noção de como funcionava o sistema de recuperação.

Amazon.com, que investe em informação, mais especificamente a informação bibliográfica levando em consideração a crítica dos leitores para o aperfeiçoamento de seus instrumentos tendo por objetivo fins econômicos. Então “o tratamento da ficção em bibliotecas espelha amplamente o tratamento da ficção na sociedade e influencia a sociedade como um todo por meio da educação dos usuários das bibliotecas”(ANJOS, 2008, p. 166).

Para amenizar os problemas para a representação da informação e do conhecimento de textos científicos, Lara (1993) propôs algumas referências que contribuem para identificação das informações brutas no momento da leitura técnica desse tipo de documento. Acredita-se que estas referências também se aplicam a textos que narrativos de ficção, que podem ser aliadas às outras estratégias, sendo elas:

- a) A posse de um quadro de referência que permita a identificação tipológica do texto;
- b) O conhecimento aprofundado dos objetivos institucionais e dos perfis de usuários;
- c) As terminologias de área que constituem referências de enciclopédias necessárias à interpretação textual e discursiva (LARA, 1993, p. 55).

Sendo assim, observa-se que no momento da representação, a terminologia enquanto base para a construção de tesouros, enciclopédias ou listas de cabeçalhos de assunto, contribuem para uma melhor representação, o que corrobora com todo o material levantado e mostra que a terminologia é sim, útil para colaborar com a representação da informação de obras narrativas de ficção.

Um exemplo de terminologia criada para o auxílio de representação da informação literária é o “Tesouro sobre Literatura”, publicado em 1985, editado pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT). Sua criação teve início no Projeto CAU (“Cabeçalho de Assunto Unificado”, da BN-1983), em que o IBICT se propôs a patrocinar a pesquisa e editá-la, como dito anteriormente, de maneira experimental, para posterior lançamento à comunidade. Em sua descrição, temos:

Para a elaboração do presente trabalho, desde o início rejeitou-se a mera listagem alfabética, adotando-se o formato de um tesouro terminológico. Na verdade, o tesouro, que tem por base as conceituações, patenteia as relações vigentes entre os termos - sinônimas, hierárquicas e outras - e, por isto mesmo, permite exercer maior controle sobre a linguagem e assegura uma prática consistente. A peculiar elaboração do Tesouro tornou-o apto a ser usado não só em sistemas pós-coordenados de indexação como também em sistemas pré-coordenados, em virtude da inclusão de uma Ordem de Citação, com vistas ao interesse específico das bibliotecas de caráter geral que empregam cabeçalhos de assunto (IBICT, 1985, p. 6).

A partir da citação acima é possível afirmar que o Tesouro foi criado com a intenção de servir de base para representação a quaisquer bibliotecas que dele precisassem, confirmando a hipótese, mais uma vez, que a Terminologia pode colaborar pra a representação da informação de obras literárias, focando, aqui, em textos narrativos de ficção.

CONSIDERAÇÕES FINAIS

Para concluir, é preciso pontuar que a intenção do presente estudo foi levantar um recorte da literatura referente à “Linguagens Documentárias”, “Organização e Representação do Conhecimento”, “Terminologia” e “Textos narrativos de ficção”, para que se pudesse observar como os assuntos vem sendo tratados em estudos. O que entende-se que ocorreu no presente artigo, pois ao concluir o que foi

proposto, observou-se que futuras pesquisas e investimentos em instrumentos de representação literárias pautadas na terminologia podem colaborar para muitas bibliotecas no momento de representar suas obras.

Até o momento, viu-se então, que os aspectos pretendidos na introdução, percorreu o universo das “Linguagens Documentárias” que, conforme já explicado ao longo do texto, se tratam de ferramentas essenciais para a representação documentária em unidades de informação pois criam pontes entre os documentos e os usuários, ou seja, são linguagens construídas a fim de facilitar o processo de busca, seleção e apropriação do documento.

A Organização e Representação do Conhecimento dos textos narrativos de ficção só podem ser realizadas de maneira profícua levando-se em consideração o papel da terminologia enquanto recurso necessário e indispensável à construção e manutenção de LDs, contemplando o aspecto representacional tanto em relação à tarefa do indexador quanto à que realiza o usuário da informação.

Ou seja, foi possível observar que esses três domínios estão intimamente ligados, servindo, quase que exclusivamente, para que o conhecimento seja representado da melhor maneira possível para a recuperação dos usuários. E, com isso, foi possível observar também, que a premissa da terminologia ter poder efetivo de colaborar com a representação de obras narrativas de ficção, se torna real.

Então, por meio desta breve explanação, pode-se concluir que a terminologia possui bases para representação de obras literárias, incluindo os textos narrativos de ficção, uma vez que glossário, tesaurus e outros instrumentos são construídos para que termos e conceitos sejam retirados das obras e, assim, se possa representa-las da melhor maneira possível.

REFERÊNCIAS BIBLIOGRÁFICAS

Anjos, L. (2008). *Sistemas de classificação do conhecimento na Filosofia e na Biblioteconomia: uma visão histórico-conceitual crítica com enfoque nos conceitos de classe, de categoria e de faceta*. 2008. (Tese de Doutorado em Ciência da Informação). Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo.

Barité, M. 2001. *Organización del conocimiento: um nuevo marco teórico-conceptual en Bibliotecología y Documentación*. Em Educação, universidade e pesquisa. Marília: Unesp-Marília-Publicações. São Paulo: FAPESP.

Cintra, A. M. M, Tálamo, M.F.G.M. & Lara, M.L.G., Kobashi, N.Y. (1996). Linguagens documentárias e terminologia. In: ALVES, I. M. (eds.) *A constituição da normalização terminológica no Brasil*. (p.17-22) São Paulo, Sp: FFLCH/CITRAT.

Dahlberg, I. (2006) Knowledge organization: a new science? *Knowledge organization*. 33 (1).

Dias, C. A. Terminologia: conceitos e aplicações. (2000) *Ciência da Informação*, 29(1), p.90-92. Recuperado de: <<http://www.scielo.br/pdf/ci/v29n1/v29n1a9>>.

Hjorland, B. (2008) What is knowledge organization? *Knowledge Organization*, v.35(2) , p.86-101.

Instituto Brasileiro de Informação e Tecnologia. Tesaurus sobre literatura / IBICT.1985. Brasília.

Lancaster, F.W. (2004). *Indexação e resumos*. Brasília, DF: Briquet de Lemos.

Lara, M. L. G. (1993). *A representação documentária: em jogo a significação*. (Dissertação de mestrado em Ciências da Comunicação). Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo.

Lara, M. L. G. de. & Tálamo, M. F. G. M. (2007). Uma experiência na interface Linguística Documentária e Terminologia. *DataGramZero*, 8(5) Recuperado de: <http://basessibi.c3sl.ufpr.br/brapci/index.php/article/view/0000007759/d50ccdddef59b7aa3edfacdffce8228ee>>

Maimone, G. D. & Tálamo, M. F.G. M. (2011). Linguística e terminologia: contribuições para a elaboração de tesouros em ciência da informação. *DataGramZero*, 12(2). Recuperado de: <http://basessibi.c3sl.ufpr.br/brapci/v/a/10072>>. Acesso em: 13 fev. 2017.

Martins, G. K, & Moraes, J. B. E. 2012. *Aspectos sociais na representação da informação: concepção integradora e democratizada a partir do âmbito da organização do conhecimento em Ciência da informação e contemporaneidade*. Tessituras e Olhares, Fortaleza: edições UFC.

Moraes, J. B. E. (2011) *A questão do aboutness no texto narrativo de ficção: perspectivas metodológicas para a Ciência da Informação*. (Tese de Livre-docência em Linguística e Documentação) - Departamento de Ciência da Informação da Faculdade de Filosofia e Ciências- Unesp, Marília.

Moraes, J. B. E. (2012) *Aboutness in Fiction: Methodological Perspectives for Knowledge Organization*. 12th International ISKO Conference: Categories, Relations and Contexts in Knowledge Organization, 6-9 de agosto 2012, Índia. ISKO (Capítulo Índia)

Pereira, E. C. & Bufrem, L. S. (2005). Princípios de Organização e Representação de conceitos em Linguagens Documentárias. *Enc. Bibli: Revista eletrônica de Bibl. Ci. Inform.*,10(20), 21-37. Recuperado de: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2005v10n20p21/302>.

Pinho, F. A., Nascimento, B. L. C. & Melo, W. L. (2015) As dimensões ôntica, epistêmia e documental na representação da informação e do conhecimento. *ACB: Biblioteconomia em Santa Catarina*, 20(1),112-123, recuperado de: https://revista.acbsc.org.br/racb/article/view/995/pdf_114. Acesso em: 15 jan. 2017.

Vargas, D. F. &, Van Der Lann, R. H. (2011) A contribuição da Terminologia na construção de Linguagens Documentárias como os Tesouros. *Biblios*. 25(1), .21-34. Recuperado de: <https://www.seer.furg.br/biblos/article/view/1988>.



CENÁRIO DA REPRESENTAÇÃO DA INFORMAÇÃO NA ORGANIZAÇÃO DO CONHECIMENTO: VOCABULÁRIOS CONTROLADOS DE ACERVOS ARQUIVÍSTICOS NO RIO DE JANEIRO

Rosale de Mattos Souza¹, Antonio Victor Rodrigues Botão²,
Julia de Vasconcellos Machado da Silva³

¹Universidade Federal do Rio de Janeiro (UFRJ), ECO, IBICT. Universidade Federal do Estado do Rio de Janeiro (UNIRIO), 0000-0002-5302-5629, rosaledemattossouza@gmail.com

²Universidade Federal do Rio de Janeiro (UFRJ), ECO, IBICT, 0000-0003-3377-0397, antoniobotao@gmail.com

³Universidade Federal do Estado do Rio de Janeiro (UNIRIO), 0000-0001-71482586, vasconcellosmachado@gmail.com

RESUMO Este estudo visa traçar o cenário da utilização dos vocabulários controlados como instrumentos de pesquisa e trabalho no âmbito da representação e recuperação da informação nas instituições que possuem acervos arquivísticos, a fim de ressaltar a necessidade de um maior destaque para o tema da representação da informação arquivística no meio acadêmico, ampliando a visibilidade para o estudo e análise dos acervos em questão e dos seus conteúdos, relevantes para a sociedade da informação. Investiga o atual cenário da utilização de vocabulários controlados nos acervos arquivísticos do Rio de Janeiro, as mudanças teórico-metodológicas advindas dos impactos das tecnologias da informação na análise, representação e recuperação de conteúdos, seus usos e necessidades informacionais no mundo contemporâneo. A representação da informação está associada juntamente com a classificação e a recuperação da informação à Organização do Conhecimento na Ciência da Informação. O conceito de representação da informação corresponde em obras de referência da área de documentação e informação com o estabelecimento de conceitos e notações simbólicas em Sistemas de Informação analógicos e digitais. A abordagem da representação da informação arquivística ainda apresenta-se como uma lacuna, pois existem correntes de pensamento desta área oriundas da historiografia do século XIX, que privilegiam o termo descrição em detrimento do termo ou atividade de representação. A Representação da Informação está mais associada à noção pós-moderna da Arquivística, de processo dinâmico, fluido, independente das idades documentais arquivísticas que balizam a gestão dos documentos e informações, já que a área arquivística consiste em um campo em construção, que vem se consolidando ao longo do tempo; um ambiente de constantes mudanças, devido à sua interdisciplinaridade com outras áreas, tais como a Ciência da Informação, o que justifica sua natureza mutável, pela forma como se relaciona com outras áreas, como Administração, Biblioteconomia e Informática, e não se mantém meramente como um simples campo autônomo auxiliar da História. A metodologia proposta para o desenvolvimento deste estudo consiste em um levantamento teórico e empírico do campo da Organização do Conhecimento, abrangendo além da área arquivística, o material científico relevante a respeito do tema da representação da informação contido em publicações nas áreas interdisciplinares mencionadas anteriormente, como também será utilizada a aplicação de questionários em instituições arquivísticas, museológicas, biblioteconômicas e de centros de documentação acerca da utilização de vocabulários controlados utilizados por essas instituições em suas práticas cotidianas no que tange ao tratamento da informação em acervos arquivísticos. Espera-se como resultado desta pesquisa, a ratificação da importância da absorção e adoção do conceito de representação

da informação arquivística - por meio da utilização de vocabulários controlados - para a área em questão, como também, provar que a teoria arquivística não se mantém estagnada através dos anos, que sua natureza flexível e adaptável às novas tecnologias e necessidades informacionais dos usuários, consiste na condição de sua própria manutenção e existência como campo científico e interdisciplinar à Ciência da Informação.

PALAVRAS-CHAVE *representação da informação, organização do conhecimento, vocabulário controlado, descrição arquivística.*

ABSTRACT This study aims to trace the scenario of the use of controlled vocabularies as tools of research and work in the scope of representation and retrieval of information in institutions that have archival collections, in order to highlight the need for greater emphasis on the topic of representation of archival information in the academic environment, increasing the visibility for the study and analysis of the collections in question and their contents, relevant to the information society. It investigates the current scenario of the use of controlled vocabularies in archival collections in Rio de Janeiro, the theoretical-methodological changes arising from the impacts of information technologies on the analysis, representation and retrieval of contents, their uses and information needs in the contemporary world. The representation of information is associated with the classification and retrieval of information to the Knowledge Organization in Information Science. The concept of information representation corresponds in works of reference of the area of documentation and information with the establishment of concepts and symbolic notations in Analog and digital Information Systems. The approach of the representation of archival information still presents itself as a gap, since there are currents of thought of this area originating from nineteenth-century historiography, which privilege the term description to the detriment of the term or activity of representation. Information Representation is more associated with the postmodern notion of Archival, dynamic, fluid, independent of the archival documentary ages that guide the management of documents and information, since the archival area consists of a field under construction, which comes Consolidating over time; An environment of constant change, due to its interdisciplinarity with other areas, such as Information Science, which justifies its changing nature, the way it relates to other areas, such as Administration, Librarianship and Informatics, and does not merely maintain A simple autonomous auxiliary field of History. The methodology proposed for the development of this study consists of a theoretical and empirical survey of the Knowledge Organization field, including addition of archival area, the relevant scientific material about the representation of information of the subject contained in publications in interdisciplinary areas mentioned above, as The use of questionnaires in archival, museological, library and documentation centers will also be applied to the use of controlled vocabularies used by these institutions in their daily practices regarding the processing of information in archival collections. As a result of this research, we hope to ratify the importance of absorbing and adopting the concept of representation of archival information - through the use of controlled vocabularies - for the area in question, as well as proving that the archival theory does not remain stagnant Over the years, that its flexible and adaptable nature to the new technologies and informational needs of users, consists in the condition of its own maintenance and existence as a scientific and interdisciplinary field for Information Science.

KEYWORDS *information representation, knowledge organization, controlled vocabulary, archival description.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A partir do projeto de pesquisa “Linguagens Documentárias, Semântica e Representação da Informação Arquivística”, viemos trabalhando desde 2014, na área de pesquisa, de uma universidade pública federal, no Rio de Janeiro. Desta forma, objetivamos o levantamento da situação de agências de informação e instituições arquivísticas que contém em seus programas de trabalho a existência de vocabulários controlados, como ferramenta que possa auxiliar na elaboração de Planos de Classificação, Tabelas de Temporalidade de Documentos, e como instrumento de controle de termos utilizados em sistemas de informação arquivísticos. Destaca-se, que na Arquivologia existem lacunas quanto à utilização da Teoria da Classificação ligada à Organização do Conhecimento, da Teoria do Conceito, da Representação da Informação, nas investigações relacionadas com vocabulário controlado e políticas de indexação de assuntos.

Pombo (1998) trata da relação da classificação como objeto antigo de análise por diversos teóricos e áreas do conhecimento, tais como, da classificação filosófica à classificação científica, em Aristóteles na Filosofia, Piaget na Educação, Augusto Comte e tantos outros autores e pensadores. Destaca também a Árvore de Porfírio, que foi inspirada no pensamento Aristotélico e estabeleceu uma hierarquia de termos e conceitos. Pombo ainda dá como exemplo que na literatura também encontra-se no escritor Jorge Luis Borges inspiração para Michel Foucault, na obra “As Palavras e as Coisas” para classificar o mundo.

Um vocabulário controlado trata-se de uma linguagem artificial, que resulta de escolhas de termos do código linguístico. Smit e Kobashi, (2003, p.14) ressaltam a importância de uma “classificação rigorosa”, ou seja, um plano de classificação que possa ser considerado adequado em relação à instituição. É importante analisar se as atividades-fim das instituições. Assim, para que se possa fazer a representação do universo documental, é necessária a incorporação de uma linguagem mais especializada. A ordenação, hierarquização, estruturação ou categorização de termos geram a arquitetura de funções e atividades da instituição, atribuindo sentido ao todo. Um vocabulário controlado deve hierarquizar as funções e atividades, contemplar o controle de vocabulário – principalmente no que se refere à sinonímia, homonímia, e quase sinonímia, relacionando os termos presentes em categorias diferentes, além de conceituar os termos e as condições de uso.

Conforme (Sousa, 2002, p. 15) a função arquivística classificação de documentos, considerada matricial para o *fazer arquivístico*, é a explicitação de funções ou estruturas dentro de um contexto de produção documental. É a representação do mapeamento deste contexto produtivo e que subsidia as demais etapas da gestão de documentos. Classifica-se o item documental pensando no todo e nas relações existentes, no papel desempenhado pela informação dentro do organismo. Ao contrário do livro para a biblioteca – unidade de informação que independe das outras unidades, o documento, ou ainda, a informação arquivística, pensada em nível de unidade de informação, só pode ser significada se relacionada com outras unidades informacionais que representam uma atividade, subfunção ou função específica. À medida que vão sendo acumuladas, as informações estabelecem relações entre si. Elas estão unidas por um elo criado no momento em que são produzidas e recebidas, denominada de organicidade, determinado pela razão de sua elaboração e que é necessário à própria existência e à capacidade de cumprir seu objetivo. Elas são um conjunto indivisível de relações intelectuais. (Sousa, 2002, p. 15)

Segundo Sousa (2002, p. 15), não existe instrumentalização que não seja sustentada por bases teóricas e/ou metodológicas. Para o autor, deve-se estar atento para que a classificação não seja um processo

intuitivo, mas levada como uma atividade técnica que demanda procedimentos metodológicos e científicos, propondo que deve-se perceber uma triangulação interdisciplinar entre os princípios arquivísticos, a Teoria da Classificação, e a metodologia da coleta de dados.

Na classificação arquivística é preciso levar em consideração a missão, as funções, as atividades, as tipologias, espécies e tipos documentais, presentes em todas as idades documentais arquivísticas, desde o arquivo corrente até ao arquivo permanente, facilitando a elaboração de planos de classificação e de tabelas de temporalidade de documentos, e, conseqüentemente a Gestão de Documentos. O aspecto objetivo, dado tanto pela Diplomática e pelas espécies documentais como pelas características orgânicas documentais, facilita na determinação conceitual dos termos, de suas séries documentais, enquanto que a representação temática que leva à subjetividade pode ser observada com maior relevância em arquivos permanentes.

Com relação aos princípios hierárquicos e classificatórios pertinentes a um arquivo, deve-se levar em consideração os seguintes princípios: Categorias de Funções (meio e fim); Áreas Funcionais; Função; Subfunção; Atividade; Subatividade; Tarefa ou ato.

Portanto, faz-se importante o diálogo da Arquivologia com outras áreas como a Biblioteconomia, a Ciência da Informação e a Comunicação para que se desenvolva a teoria do conceito, os princípios e o fazer arquivísticos na sociedade do conhecimento. Espera-se como resultado deste estudo, a ratificação da importância da absorção e adoção do conceito de representação da informação arquivística e do corpo teórico da Organização do conhecimento, por meio da utilização de vocabulários controlados. E provar que a teoria arquivística não se mantém estanque, que sua natureza deve ser adaptável às novas tecnologias e necessidades informacionais dos usuários, consiste na condição de sua própria manutenção e existência como campo científico e interdisciplinar à Ciência da Informação.

OS VOCABULÁRIOS CONTROLADOS COMO INSTRUMENTOS AUXILIARES DE PESQUISA ASSOCIADOS À DESCRIÇÃO NA ARQUIVÍSTICA E À REPRESENTAÇÃO DA INFORMAÇÃO

As normas de descrição na Arquivística, tais como, a Norma Internacional de Descrição Arquivística - ISAD (G) e a Norma Nacional de Descrição Arquivística - NOBRADE, entre outras normas de descrição, inauguraram um novo momento na forma de entendimento de descrição e recuperação da informação na Arquivística, levando-se em consideração as tecnologias e as novas necessidades dos sistemas de informação. Não podemos deixar de refletir sobre as características principais destas normas e da descrição multinível, tais como: descrição do geral para o particular, informação relevante para o nível de descrição; relação entre descrições, não repetição da informação.

Existem sete áreas de informação na Norma Internacional de Descrição Arquivística - ISAD (G), que são: 1. Área de Identificação; 2. Área de Contextualização; 3. Área de Conteúdo e Estrutura; 4. Área de Condições de Acesso e Uso; 5. Área de fontes relacionadas; 6. Área de Notas; 7. Área de Controle da Descrição; e oito áreas de informação na Norma Nacional de Descrição Arquivística - NOBRADE, que deriva da norma anteriormente citada e na qual foi incluída a Área de n. 08, de Pontos de Acesso e indexação de assuntos, a qual não iremos detalhar neste artigo, deixando para outro momento. Houve uma preocupação neste trabalho em particular com esta área de informação, pois os arquivistas carecem de maior conhecimento e treinamento no que diz respeito à representação da informação, em particular,

quanto à indexação, tratamento e disseminação da informação, pois não há um entendimento claro de Organização do Conhecimento, subárea da Ciência da Informação, associada à Teoria Arquivística. No Glossário da NOBRADE há apenas uma conceituação do que é indexação coordenada e indexação pré-coordenada, porém, não há uma maior explanação sobre indexação e políticas de indexação naquela norma de descrição. Inclusive, a expressão cabeçalho de assunto não faz parte da terminologia arquivística, necessitando de uma interdisciplinaridade no conhecimento do processo documental, conforme poderemos ver a seguir: “Indexação pós-coordenada: Indexação por termos que devem ser combinados no momento da busca para filtragem da informação desejada, em geral chamados de descritores” e “Indexação pré-coordenada: Indexação por termos combinados previamente, em geral identificados com cabeçalhos de assunto” (Conselho Nacional de Arquivos, 2005).

Segundo CAMPOS (1999) faz uma análise da representação descritiva e temática da informação na Arquivística:

Podemos citar como preocupação desta natureza a própria ISAAR – Norma Internacional de Registro de Autoridade Arquivística para entidades coletivas, pessoas e famílias. Além disto, a própria ISAD(G) recorre às normas ISO para descrição documentária para padronização de fontes. Entretanto, quanto aos pontos de acessos relacionados à indexação temática, esta mesma norma somente se preocupa em indicar que se deva controlar o vocabulário, mas não aborda aspectos metodológicos relativos aos procedimentos de análise e mesmo de elaboração de vocabulários controlados. (Campos, 1999, p. 17-31)

É preciso haver estudos interdisciplinares mais consistentes a respeito do tema, como, por exemplo, a representação descritiva e a representação temática da informação. Além disto, destacamos que os vocabulários controlados são ferramentas e podem ser utilizadas como metodologia para auxiliar nos Planos de Classificação, Tabelas de Temporalidade de Documentos e na elaboração de instrumentos de pesquisa.

Nesse sentido, chegamos ao Vocabulário Controlado: instrumento que nomeia os termos correspondentes à linguagem documentária utilizada nos arquivos de determinado acervo, visando organizar e recuperar com consistência e precisão os documentos e as informações arquivísticas. Para tanto, procura padronizar toda linguagem a ser utilizada pela instituição, através da nomeação de um termo único para determinado fim, evitando assim múltiplas entradas, garantindo a recuperação completa das informações.

Imaginemos que os pedidos de carros sejam nomeados de cinco formas diferentes: pedido de transporte; requerimento de veículo; requisição de veículo; solicitação de transporte e solicitação de veículo. [...] Se, por exemplo, ocorrerem problemas no agendamento de um carro, para localizar o respectivo documento e verificar o que aconteceu, será necessário lembrar como o pedido foi nomeado no momento da solicitação. [...] Com o objetivo de evitar esses dissabores, a solução mais racional passa pelo controle de vocabulário: na entrada do sistema, com a adoção de uma única forma de designação para nomear documentos que são gerados pela mesma atividade; e na saída do sistema (na busca), informando como cada atividade é nomeada pelo sistema (Smit & Kobashi, 2003, pp. 16-17).

A partir das considerações anteriores, constata-se que o vocabulário controlado é um método utilizável tanto na organização como na recuperação dos documentos, servindo para todos os processos arquivísticos rotineiros, uma vez que consiste em um sistema de referência comum e compartilhado pelos produtores e pesquisadores da informação arquivística de uma mesma entidade custodiadora,

possibilitando a otimização de Planos de Classificação e de Tabelas de Temporalidade de e Destinação de Documentos.

Segundo o manual publicado pelo Arquivo do Estado e Imprensa de São Paulo, “Como Elaborar Vocabulário Controlado Para Aplicação Em Arquivos”, de autoria de Smit e Kobashi (2003) existem dois diferentes procedimentos na construção e utilização dos vocabulários controlados: o procedimento micro e o procedimento macro.

O primeiro procura introduzir o controle nas expressões ou termos e entre eles, englobando aspectos referentes à linguística (como o Controle de sinonímia, homonímia e quase-sinonímia e a normalização gramatical), enquanto o segundo atenta para a ordenação dos termos, buscando evitar que o instrumento se torne uma longa lista de expressões controladas para as quais não se sabe ao certo o momento e a finalidade de utilização, e o faz através do estabelecimento de uma hierarquização, ordenação, estruturação ou categorização de termos.

Embora apareça com maior incidência, em se tratando do processo de recuperação da informação, conclui-se que a utilização de vocabulários controlados facilita o cotidiano dos arquivos em todas as idades documentais, quando há a preocupação com os termos adequados ao tipo documental já em sua produção (arquivo corrente) e/ou quando revela as entradas apropriadas no momento da busca para o acesso (arquivo permanente).

METODOLOGIA

O panorâma da utilização de vocabulários controlados no Rio de Janeiro, no Brasil está consolidado a partir da aplicação de questionários às agências de informação que têm acervos arquivísticos, com relação a existência ou não da presença do vocabulário controlado como ferramenta de trabalho e de busca das informações; possibilitando futuras políticas e Manuais de Indexação de termos. Diante disso, fizemos um levantamento das Instituições que possuem acervos arquivísticos na cidade do Rio de Janeiro, a fim de construirmos quantitativamente e qualitativamente o cenário atual do tratamento da informação arquivística, considerando sua representação. Foram aplicados questionários em (26) instituições, entre elas oito (8) Centros de Documentação, dez (10) Arquivos, três (3) Bibliotecas e cinco (5) Museus, conforme a figura 1.

Aplicações do Questionário



Figura 1. Gráfico expositivo de aplicação de questionário em instituições com acervos arquivísticos

Referência: Elaborado pelos autores

Entretanto, apenas dezesseis instituições nos retornaram com questionários respondidos ou direcionamentos para questões apresentadas, de acordo com a figura 2.

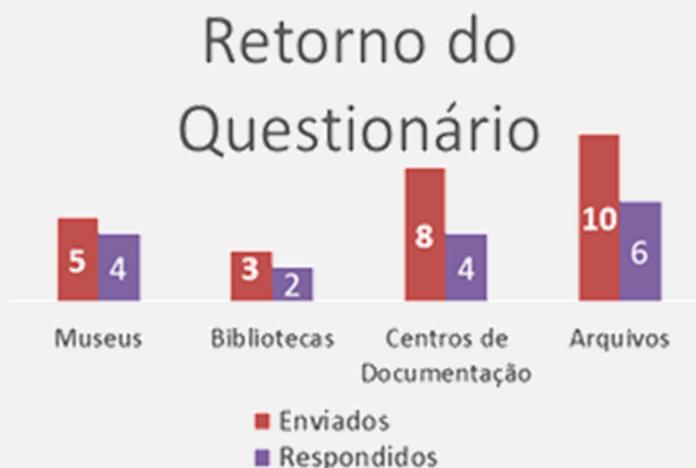


Figura 2. Estatística de retorno dos questionários enviados

Referência: Elaborado pelos autores

A partir dos resultados, iniciamos o reconhecimento das instituições que possuem e utilizam vocabulários controlados como instrumento de trabalho e pesquisa, apresentadas na figura 3.



Figura 3. Instituições que utilizam vocabulário controlado como ferramenta de trabalho e pesquisa

Referência: Elaborado pelos autores

A partir dos subsídios anteriores com relação à utilização de vocabulários controlados, esboça-se por meio das figuras 4 (Sistema eficiente X Vocabulário Controlado), 5 (Utilização de Representação da Informação em Centros de Documentação, Bibliotecas e Museus) e 5.1 (Arquivos que trabalham com Representação da Informação), um panorama geral do uso da Representação da Informação e do Vocabulário Controlado em instituições que abarcam acervos arquivísticos, demonstrados pela figura 6.

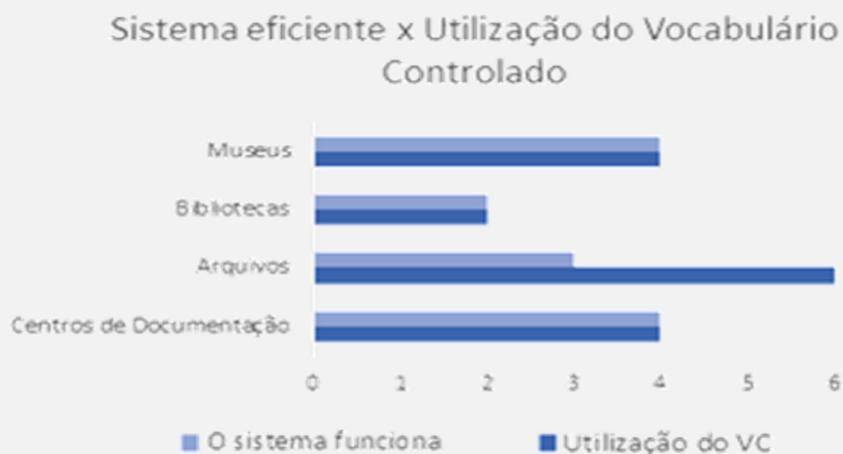


Figura 4. Sistema eficiente X Vocabulário Controlado

Referência: Elaborado pelos autores

A Figura 4 demonstra que em arquivos, a eficiência de um sistema de informação, não está atrelada somente à aplicação de vocabulários controlados, o que denota a utilização de outros instrumentos de controle da eficácia do processamento da informação. Já a partir da figura seguinte, temos outro tipo de observação:

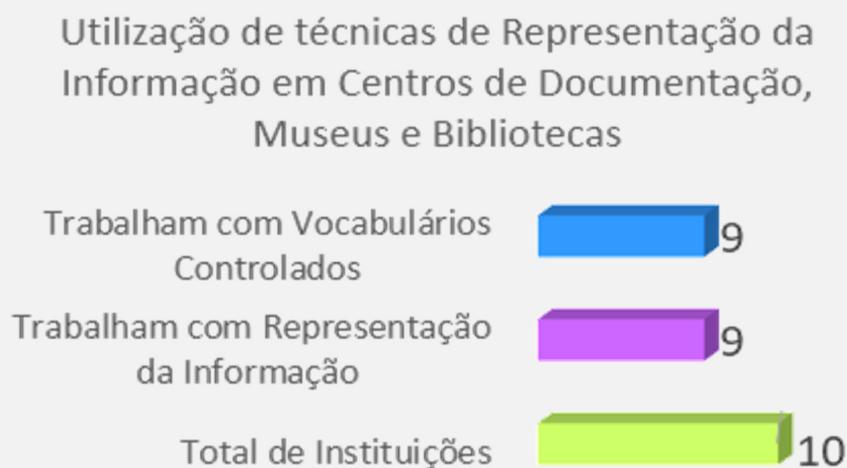


Figura 5. Utilização de Representação da Informação em Centros de Documentação, Bibliotecas e Museus

Referência: Elaborado pelos autores

Constatamos a partir da Figura 5, que para Centros de Documentação, Museus e Bibliotecas, pelo menos 50% das entidades questionadas trabalham com essas ferramentas, enquanto apenas um Centro de Documentação consultado tem essa preocupação de forma mais avançada, o da Casa de Oswaldo Cruz (COC), da Fundação Oswaldo Cruz (FIOCRUZ), que tem um grupo de trabalho interdisciplinar de uma Política de Indexação, tendo promovido a elaboração de um Manual de Indexação de termos.

Atribuímos esse resultado à dois fatores estudados nessas áreas: o tratamento focado na informação/conteúdo do acervo, que exige maior atenção para a questão da representação da informação; e a maior preocupação com o usuário. É evidente que a consideração de ambos está constantemente presente em discussões da Biblioteconomia e Museologia que, por sua vez, afetam os Centros de

Documentação, enquanto para Arquivos, ainda são consideradas questões muito recentes e pouco estudados, o que se ratifica na representação do gráfico na figura 5.1.

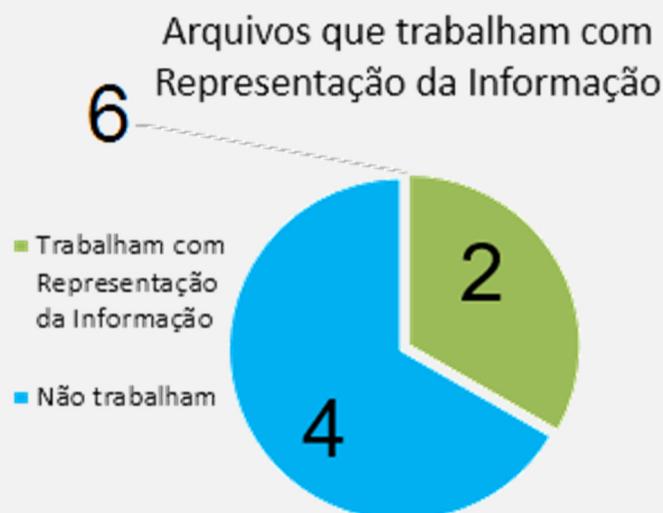


Figura 5.1. Arquivos que trabalham com Representação da Informação

Referência: Elaborado pelos autores

Nota-se na figura 5.1, que há um maior número de ocorrências de instituições arquivísticas que não trabalham com o conceito e técnicas de Representação da Informação, ferramenta esta, utilizada em larga escala nas demais instituições de documentação e informação, como bibliotecas, museus e centros de documentação e informação (figura 5).

A discussão se torna ainda mais relevante quando notamos que a inversão apenas não é unânime, pois a única biblioteca que não utiliza a ferramenta possui majoritariamente acervo arquivístico e que, se não fosse por isso, todas as demais instituições, que não os arquivos, utilizariam o instrumento, dando ainda mais força ao argumento anterior. Observemos tais resultados do panorama traçado na Figura 6:

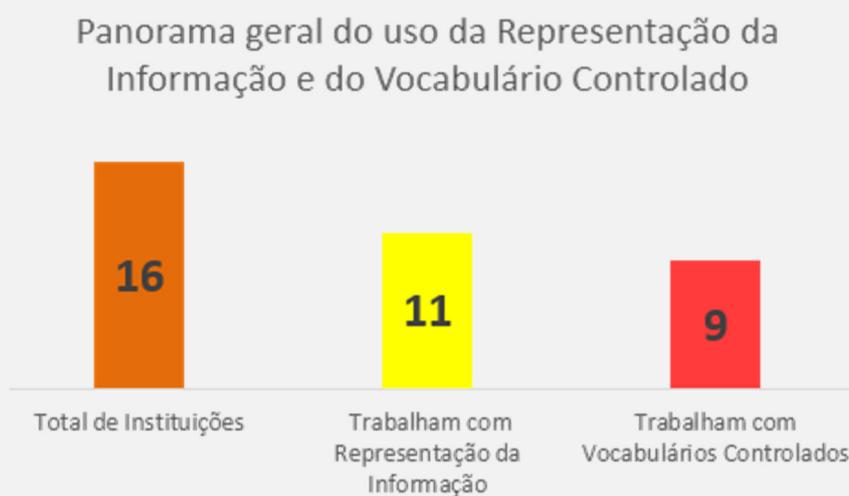


Figura 6. Panorama geral do uso da Representação da Informação e do Vocabulário Controlado

Referência: Elaborado pelos autores

Embora os resultados anteriores ilustrem nitidamente a necessidade do investimento nessa questão, é importante salientar que quantidade não é qualidade. Por isso trouxemos também o resultado das práticas do uso dessa ferramenta. É justo argumentar que, mesmo que os arquivos não a utilizem, seus sistemas de recuperação da informação funcionam, ainda que não em sua totalidade. Entretanto, a constatação não invalida a discussão, principalmente quando percebemos que o retorno da eficiência do sistema não é ainda absoluto nas entidades que trabalham exclusivamente com o vocabulário controlado (como os Museus e os Centros de Documentação).

RESULTADOS

Ratifica-se a importância dos vocabulários controlados como ferramentas da representação da informação arquivística no contexto da produção documental, peculiar às instituições produtoras de documentos arquivísticos, levando-se em conta o contexto de produção, que peculiariza os arquivos em relação às outras unidades de informação, como bibliotecas, museus e centros de documentação;

Destacar a adoção do termo representação da informação arquivística na área em questão, já que, mesmo em detrimento de algumas correntes arquivísticas eminentemente históricas, há que se adotar tal referencial teórico e prático, coadunando com o perfil mutante e variável da área arquivística frente às novas necessidades informacionais que desencadeiam emergentes comportamentos por parte dos usuários; caracterizados pela mudança de paradigma na área arquivística, que migrou do paradigma documental para o informacional, este último influenciado pelas tecnologias de informação e comunicação e seu inegável impacto na chamada Sociedade da Informação e do Conhecimento.

Avançar na área de Arquivologia com as relações interdisciplinares com a Biblioteconomia, a Ciência da Informação e a Organização do Conhecimento, como subárea da CI, ampliando as investigações teóricas e práticas do uso de linguagens documentárias e a representação temática da informação em sistemas de informação em instituições que contenham acervos arquivísticos.

CONCLUSÕES

Esta pesquisa pretende contribuir e consolidar as linguagens artificiais na arquivística como ferramentas representacionais, com vistas à construção de melhores instrumentos de classificação arquivísticos, descrições e sistemas de recuperação da informação mais eficientes, além da excelência no atendimento aos usuários.

Considera-se também como objetivo, o estímulo a reflexão sobre o conceito de representação da informação no campo arquivístico, já que tal conceito permeia as demais áreas da documentação-informação com perspectivas comuns ao tratamento informacional, visando a otimização da construção de sistemas de recuperação de informações nas organizações, tanto públicas como privadas.

O impacto da representação da informação arquivística, que é notória apesar de algumas contestações conceituais na própria área arquivística, podem se refletir na otimização de melhores e mais eficientes sistemas de Recuperação da Informação – SRI e de Gestão da Informação, promovendo maior atendimento aos usuários, conforme seu perfil de pesquisas. A aceitação e aplicação do conceito de

representação em arquivos traz à tona a discussão sobre tal atividade para além da qual é conhecida como descrição arquivística.

Há que se estimular o pensar e o fazer arquivísticos com uma maior produção intelectual sobre reflexões que envolvam trabalhos em congressos, artigos e livros sobre representação da informação e suas contribuições para o campo arquivístico. O arquivista deve estar sempre em constante adaptação às novas demandas de mercado com relação à sua conduta profissional, teórica e prática, por conta da época de mudanças que vivenciamos, de informações fluidas e mutáveis, da modernização dos aparatos tecnológicos, da relação cada vez mais dinâmica entre profissionais da informação e usuários, seja ele um ator inserido em bibliotecas, arquivos, centros de documentação ou museus, dos quais indiscutivelmente temos que lançar mão para justificar a existência da área Arquivística e da Ciência da Informação.

REFERÊNCIAS BIBLIOGRÁFICAS

CAMARGO, A.M. A., & BELLOTO, H. L. (2010). *Dicionário de Terminologia Arquivística: (versão bolso)*. São Paulo: FEUSP/FAPESP.

CAMPOS, M. L. A. (2001). *Linguagem Documentária: Teorias que fundamentam sua elaboração*. Niterói: EDUFF.

CAMPOS, M. L. A. (2006). Indexação e descrição em arquivos: a questão da representação e recuperação de informações. *Arquivo & administração*, 5(1), 17-31.

CONSELHO NACIONAL DE ARQUIVOS (BRASIL). (2005). *Norma Brasileira de Descrição Arquivística*. Rio de Janeiro: Arquivo Nacional.

CONSELHO INTERNACIONAL DE ARQUIVOS. (2000). *Norma Internacional de Descrição Arquivística - ISAD (G)*. Rio de Janeiro: Arquivo Nacional.

CONSELHO NACIONAL DE ARQUIVOS (Brasil). Câmara Técnica de Documentos Eletrônicos. (2006). *Modelo de Requisitos para Sistemas informatizados de Gestão Arquivística de Documentos- E-ARQ Brasil*. Rio de Janeiro: Arquivo Nacional.

COOK, T. (2013). Moda absurda ou Renascimento Profissional: Pós-modernismo e a Prática de Arquivo. *Informação Arquivística*, 2(1), 158-187.

DAHLBERG, I. (1978). Teoria do Conceito. *Ciência da Informação*. 7(2), 101-107.

DUCHEIN, M. (1986). O Respeito aos Fundos em Arquivística: Princípios Teóricos e Problemas Práticos. *Arquivo & administração*. 1, 10-14.

LANCASTER, F. W. (2004). *Indexação e Resumos: Teoria e Prática*. Brasília: Briquet de Lemos.

POMBO, O. (1988) Da classificação dos seres à classificação dos saberes. Leituras. *Revista da Biblioteca Nacional de Lisboa*, 2, 19-33.

SMIT, J. W. & KOBASHI, N. Y. (2003). *Como elaborar vocabulário controlado para aplicação em arquivos*. São Paulo: Arquivo do Estado e Imprensa Oficial do Estado.

SOUSA, R. T. (2003) Os princípios arquivísticos e o conceito de classificação. In: Rodrigues, MEDLEG, Georgete. LOPES, Ilza Leite (Orgs.). *Organização e Representação do Conhecimento na perspectiva da Ciência da Informação*. Brasília: CID/Thesaurus.

SOUSA, R. T. (2002). *As bases do processo classificatório em Arquivística: um debate metodológico*. São Paulo: Associação dos Arquivistas de São Paulo.

SOUSA, R. T. (2007). A análise do desenvolvimento da classificação em arquivística. In V. B. Santos & C. H. Inarelli. *Arquivística: temas contemporâneos*. Brasília: SENAC.



POSSIBILIDADES E VIABILIDADE DA DESCRIÇÃO ARQUIVÍSTICA DESDE A PRODUÇÃO DA INFORMAÇÃO DIANTE DOS DESAFIOS DE SUA RECUPERAÇÃO

Angelica Alves da Cunha Marques¹, Thaianie Honda Cotts²

¹Universidade de Brasília (UnB), 0000-0003-4642-5912, angelicacunha@unb.br

²Universidade de Lisboa (Ulisboa), 0000-0001-9949-3834, thondacotts@gmail.com

RESUMO Esta comunicação objetiva compreender a relevância da descrição arquivística desde a criação dos documentos, considerando a transversalidade das informações produzidas, organizadas, recuperadas, acessadas e utilizadas nos arquivos. Trata-se de uma pesquisa bibliográfica na literatura da Arquivologia e da Ciência da Informação. Os resultados demonstram que, embora os fundamentos epistemológicos da Arquivologia ainda ressoem como basilares, há que se pensar na sua viabilidade, de maneira a se reformular os processos de criação, produção, organização, representação, recuperação, acesso e utilização dos documentos de arquivo. Amparados por instrumentos conceituais, metodológicos e produtos advindos do trabalho do arquivista, deve-se redimensionar esses fundamentos mediante os desafios contemporâneos, ampliados e aprofundados quanto à organização e à recuperação dessas informações diante das leis de acesso à informação.

PALAVRAS-CHAVE *Descrição arquivística, representação da informação, recuperação da informação, fundamentos arquivísticos.*

ABSTRACT This communication aims to understand the relevance of the archival description from the creation of the documents, considering the transversality of the information produced, organized, retrieved, accessed and used in the archives. Within this article is a bibliographical research of the Archival Science and Information Science's Literature. The results show that, although the epistemological foundations of Archival Science still resonate as basilar, one has to think about its viability, to reformulate the processes of creation, production, organization, representation, retrieval, access and use of archival documents. Supported by conceptual, methodological and product tools derived from the archivist's work, these fundamentals must be re-dimensioned through contemporary, expanded and in-depth challenges regarding the organization and retrieval of this information under the laws of access to information.

KEYWORDS *Archival description, information representation, information retrieval, archives fundamentals.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Se informação é o que é informativo para determinada pessoa (Capurro; Hjørland, 2007), os diversos conceitos de informação que perfazem o objeto de estudo da Ciência da Informação em torno da compreensão da sua criação, da sua organização, da sua recuperação, do seu acesso e dos seus usos são,

inexoravelmente, perpassados por contextos e necessidades específicos, mesmo que reflitam demandas sociais.

Semelhantemente, ainda que em abordagens mais contingenciais, o objeto de estudo, os métodos e as técnicas arquivísticas também impescindem da compreensão dos contextos e das necessidades que conduzem a criação, a produção, a organização, a recuperação, o acesso e os usos dos documentos de arquivo, independentemente da fase nas quais os conjuntos documentais se encontrem – corrente, intermediária ou permanente.

A descrição arquivística, um dos métodos de representação da informação orgânica registrada, ou seja, dos documentos de arquivo e dos contextos da sua produção, acumulação e usos (Esteban Navarro, 1995), é apontada em grande parte da literatura da área como uma atividade típica dos arquivos permanentes. Para Bellotto (2006, p. 179), por exemplo, “Não se trata mais da utilização do documento pelo produtor, do seu valor primário, ligado à própria razão de ser do ato escrito consignado no documento; a descrição destina-se àquele cuja tarefa é explorar o que restou, após ter cumprido a finalidade administrativa ou jurídica do ato”. Desse modo, prossegue a autora, “o processo de descrição consiste na elaboração de instrumentos de pesquisa que possibilitem a identificação, o rastreamento, a localização e a utilização de dados”.

Assim, por muito tempo, os documentos correntes, nutridos pelo seu valor primário, não eram objeto da descrição; os documentos que apresentavam predominantemente valor secundário eram, por sua vez, considerados em seu potencial informacional na perspectiva história. O processo de descrição limitava-se aos arquivos permanentes e objetivava a elaboração de instrumentos de pesquisa alinhados ao papel social do arquivo, com a difusão e promoção do acesso informacional para usuários diversos. Para Schellenberg:

Ao realizar esse trabalho, o arquivista se inteira da procedência, do conteúdo, do arranjo e do valor dos papéis. Esses dados são por ele registrados em instrumentos de busca que servem a um duplo propósito: a) tornar os papéis conhecidos às pessoas que possam vir a se interessar pelos mesmos; e b) facilitar ao arquivista a pesquisa. (Schellenberg, 2006, p. 313).

Nesse sentido, a descrição não se limita à elaboração de instrumentos de pesquisa (Oliveira, 2012), tampouco às atividades relacionadas aos arquivos permanentes (Rodrigues, 2003). Considerando essa perspectiva mais abrangente da descrição, esta comunicação busca compreender a relevância da descrição arquivística desde a criação dos documentos, considerando a transversalidade das informações produzidas, organizadas, recuperadas, acessadas e utilizadas nos arquivos, bem como dos desafios decorrentes das leis de acesso à informação, ou seja, o dever do Estado de preservação dos documentos e o direito de acesso, por parte dos cidadãos.

A pesquisa que resulta nesta comunicação caracteriza-se como qualitativa, exploratória, descritiva e bibliográfica mediante breve pesquisa em algumas bases de dados e repositórios digitais (Repositório Científico de Acesso Aberto de Portugal (RCAAP), *Scientific Electronic Library Online* (SciELO.br) e B-on: Biblioteca do conhecimento on-line). Assim, argumentos para a descrição de documentos de valor primário foram buscados a partir da literatura da Ciência da Informação, numa perspectiva mais ampla, e da Arquivologia, mais especificamente, tendo em vista o conhecimento e a compreensão dos métodos e modelos já propostos, no âmbito dos elementos necessários ao alcance da gestão estratégica da informação e da gestão de documentos.

A partir dessa breve revisão de literatura, observamos que, apesar de ressoarem como basilares, os fundamentos arquivísticos se deparam com desafios contemporâneos – novos modelos de organizações, transformação do perfil dos usuários, novos suportes documentais, leis de acesso à informação, novas tecnologias de informação e de comunicação. Nessa nova perspectiva, surge a necessidade do seu redimensionamento, tal como a reformulação dos processos de criação, produção, organização, representação, recuperação, acesso e utilização dos documentos de arquivo, de forma que sejam ampliados e aprofundados quanto à organização e à recuperação dessas informações.

Os resultados da pesquisa bibliográfica ainda são parciais, dada a variedade da literatura nesse sentido, mas já demonstram essa nova perspectiva e a necessidade de reavaliação da viabilidade de aplicação dos fundamentos da Arquivologia tais como eles foram criados, amparados por instrumentos conceituais, metodológicos e produtos advindos do trabalho do arquivista, como manuais de gestão de documentos, planos de classificação, tabelas de temporalidade e instrumentos de pesquisa.

A DESCRIÇÃO ARQUIVÍSTICA NA LITERATURA

A descrição arquivística é um tema recorrente nas discussões da área. Seu destaque aumentou consoante o aumento da necessidade de explicitar o contexto de produção dos conjuntos documentais, pois cabe a ela viabilizar o entendimento das informações registradas (documentos) com o seu contexto de produção e acumulação. Inicialmente uma ferramenta de controle dos acervos permanentes sob a custódia das organizações, passa também a exercer o papel de facilitadora da recuperação e do acesso aos documentos. Carucci evidencia que:

[...] o documento interessa por seu conteúdo, pelas informações que transmite. Todavia, as notícias que ali são representadas ou descritas requerem, de quem as adequa às capacidades técnicas, que sejam traduzidas em cânones de representação, os quais, por sua vez, podem constituir objeto de análise, sendo, esses testemunhos diretos da atividade de documentar. (Carucci apud Lopes, 2000, p. 83).

Os potenciais usos dos documentos de arquivo são também apresentados por Duranti, que afirma que:

Os registros, além das necessidades do direito e da história, servem à “transparência das ações”, um novo e atraente nome para o que mais tradicionalmente constitui a obrigação de prestar contas (*accountability*) tanto do ponto de vista administrativo quanto histórico. (Duranti, 1994, p. 55)

Essa autora também lembra a naturalidade da acumulação dos documentos de arquivo, ao afirmar que:

Os registros arquivísticos não são coletados artificialmente, mas acumulados naturalmente nas administrações, em função dos seus objetivos práticos; os registros arquivísticos se acumulam de maneira contínua e progressiva, como sedimentos de estratificações geológicas, e isto os dota de um elemento de coesão espontânea, embora estruturada (organicidade). (Duranti, 1994, p. 3)

Considerando então o inter-relacionamento entre os documentos que compõem um fundo e o contexto de sua produção e acumulação, há que se pensar em maneiras que permitam a recuperação dos documentos de forma que esse contexto seja preservado. Esse contexto, relacionado ao conteúdo semântico e ao processo de criação dos documentos, engloba a informação registrada em cada documento, bem como no seu conjunto (fundo), refletindo a forma, a estrutura organizacional e os processos administrativos que os originam. E é a preservação desse contexto que viabiliza o registro e

a compreensão dos laços que unem a informação registrada com o processo que a gerou. Segundo Sousa (2014):

À medida que os documentos vão sendo acumulados, estabelecem relações entre si. Eles estão unidos por um elo criado no momento em que são produzidos e recebidos, determinado pela razão de sua elaboração e que é necessário à própria existência e a capacidade de cumprir seu objetivo. Eles são um conjunto indivisível de relações intelectuais, onde o “todo é maior que a soma de suas partes”. (Sousa, 2014, p. 7).

A descrição arquivística de documentos é um processo subjetivo e intelectual. O profissional responsável por essa atividade deve dominar as capacidades de interpretação textual, do idioma em uso e suas regras gramaticais, além de conhecer a história, legislação e hierarquia da instituição. A exemplo de Schellenberg (2006), Heredia Herrera afirma que:

A descrição é o meio utilizado pelo arquivista para obter a informação contida nos documentos e oferecê-la aos interessados nela. A descrição persegue dois objetivos: dar informação aos demais e facilitar o controle pelo arquivista. (Heredia Herrera, 1988, p. 300, tradução nossa).

Para Cruz Mundet (2001, p. 256, tradução nossa), o objetivo da descrição é tornar os fundos documentais eficazmente acessíveis. Ele afirma que “a descrição dos documentos constitui a parte culminante do trabalho arquivístico e coincide exatamente com sua finalidade própria com a da própria documentação: informar”.

Tradicionalmente, a descrição dos documentos de arquivo era realizada a partir da memória do descritor, e construída de maneira breve e objetiva. Nas organizações tradicionais, a memória da instituição era mantida pelos funcionários que nelas trabalhavam há anos e que conheciam bem a sua história. Os registros eram comumente feitos com base em numeração de caixa e de documentos, listagens, materiais de acondicionamento, ou mesmo listas de assuntos ou palavras-chave (que eram considerados descritores). Esse processo de descrição acarretava perda ou comprometimento do significado informacional.

Em 1898, a Associação dos Arquivistas Holandeses publica um manual que apresenta um padrão para a realização da descrição de documentos, com normas para as práticas dos profissionais de arquivos (Mueller, Feith e Fruin, 1960). O capítulo dedicado à descrição aborda os procedimentos a serem seguidos para a elaboração dos diversos instrumentos de pesquisa, sem chegar a definir claramente o que era entendido como “descrição”. Discutem-se os tipos de instrumentos de pesquisa, seu grau de detalhamento e a importância de serem seguidas regras, tais como o tratamento do acervo como um todo antes de se descrever um fundo ou uma série em particular. A intenção do texto parece ser a de estabelecer um procedimento propriamente arquivístico que contemplasse o conjunto dos documentos do arquivo, diferente das práticas dos historiadores (especialmente aqueles de orientação positivista), que se dedicavam em maior profundidade aos documentos considerados mais “importantes” (Hagen, 1998).

Ainda segundo essa autora, as primeiras tentativas de padronização na área foram os dicionários de termos técnicos. A descrição passou, assim, a ser objeto de esforços de padronização basicamente em função do impacto de novas tecnologias e, para se beneficiar desses recursos, a comunidade arquivística teve de desenvolver a comunicação do conhecimento, envolvendo conceitos utilizados por outras disciplinas (Hagen, 1998).

Mesmo assim, vários autores da Arquivologia reconhecem a descrição como uma função típica dos arquivos permanentes, como é o caso de Bellotto (2006). Apesar de definir o processo de descrição como ininterrupto (visto que a incorporação ou produção de novos documentos não finda), ela afirma que:

Ela [a descrição] não cabe nos arquivos correntes, onde seu correspondente é o estabelecimento dos códigos do plano de classificação – que acabam por servir de referência para a recuperação da informação –, assim como de outras categorias de controle de vocabulário e indexação que se usem para o mesmo fim. Tampouco a descrição faz sentido no âmbito dos arquivos intermediários, onde a frequência de utilização secundária é quase nula. (Bellotto, 2006, p. 173)

Essa limitação da descrição aos documentos de guarda permanente pode ser justificada pela complexidade do trabalho arquivístico quando da elaboração dos instrumentos de pesquisa, criados especificamente para tornar possível a busca por informações.

Há, também, perspectivas mais amplas da descrição, a exemplo do que é apresentado nas normas internacionais de descrição, que preconizam que o arquivista tem o trabalho subjetivo de preservar a história e o contexto de produção, acumulação e organização dos documentos. De acordo com a Norma Geral Internacional de Descrição Arquivística (ISAD(G)):

O objetivo da descrição arquivística é identificar e explicar o contexto e o conteúdo de documentos de arquivo a fim de promover o acesso aos mesmos. Isto é alcançado pela criação de representações precisas e adequadas e pela organização dessas representações de acordo com modelos predeterminados. **Processos relacionados à descrição podem começar na ou antes da produção dos documentos e continuam durante sua vida.** Esses processos permitem instituir controles intelectuais necessários para tornar confiáveis, autênticas, significativas e acessíveis descrições que serão mantidas ao longo do tempo. (Conselho Internacional de Arquivos. ISAD-G, 2000, p. 11, grifos nossos).

Portanto, essa norma prevê a descrição arquivística independentemente da fase documental e reconhece a descrição no âmbito das diversas intervenções do arquivista.

De uma forma geral, a descrição ganhou consistência com a definição de normas internacionais, que buscam a padronização do processo. Cabe a cada país discuti-las, adaptá-las e colocá-las em uso. A aplicação dessas normas pode significar o intercâmbio efetivo de informações, mediante regras internacionalmente definidas que buscam conferir credibilidade e fiabilidade ao trabalho. Nessa perspectiva, a descrição possibilita acesso às informações de domínio público, sejam elas dotadas de valor administrativo, informacional, probatório e/ou histórico.

As informações contextuais dos documentos arquivísticos assumiram um papel de maior destaque na descrição com a publicação da ISAD(G), pois os dados contextuais ganharam importância para a interpretação e compreensão do conteúdo descrito. Os instrumentos tradicionais do processo de descrição arquivística tiveram seus papéis redimensionados, a descrição ultrapassou formatos e suportes tradicionais, que passaram a ser utilizados de maneira mais conectada e integrada. O processo de representação da informação passou de intuitivo e empírico a um processo baseado no objeto (arquivo em todas as suas dimensões) e em normas sedimentadas na teoria arquivística.

As normas de descrição arquivística são fundamentalmente elaboradas para assegurar que as descrições dos documentos de arquivo sejam consistentes, apresentem conteúdo relevante e explícito. Mas para alcançar esse nível de resultados, as normas devem ser adotadas e implementadas por todas as

organizações. Deve-se encontrar um caminho que respeite a história e a tradicional descrição arquivística, mas que viabilize o incremento dos procedimentos, de forma que os documentos sejam entendidos, potencialmente representados com base em sua proveniência e que a descrição seja completa.

Lopes (2009) entende a descrição como um processo que se inicia na classificação e tem continuidade no processo de avaliação. Assim, as funções de classificar, avaliar e descrever documentos de arquivo devem ser consideradas complementares. Na prática, trata-se de identificar um conjunto documental, classificá-lo – e, dessa forma, criar o primeiro nível avaliativo e descritivo – e elaborar uma primeira representação do conteúdo avaliado. Enquanto a classificação estabelece as relações lógicas dos conjuntos documentais, a descrição é construída mediante textos explicativos, mas sintéticos e espontâneos. O processo de descrição deve ser aplicado na classificação dos documentos e continuar durante todo o tratamento documental.

Desse modo, Lopes (2009) defende que todas as operações intelectuais possuem natureza descritiva. Daí a dificuldade em dissociar os processos de classificação, avaliação e descrição dos documentos. A descrição de documentos está relacionada a atividades de diferentes níveis descritivos: nos planos de classificação, quando da definição dos fundos e suas subdivisões; nas tabelas de temporalidade e destinação de documentos, na definição do que dever ser mantido sob custódia, por quanto tempo guardar, se é permanente ou descartável; nos instrumentos de pesquisa; na difusão dos conteúdos informacionais, como em sítios eletrônicos de instituições arquivísticas.

Semelhantemente, Sousa defende a relação entre as funções arquivísticas, destacando o papel da classificação como o ponto de partida desses processos:

Para realizar a descrição de documentos de arquivo precisamos partir dos conjuntos documentais definidos pelo processo classificatório e representados nas unidades de classificação. Todos os instrumentos de descrição ou de pesquisa partem das unidades de classificação (fundo, série, subsérie) para descrever. (Sousa, 2014, p. 4).

Haworth (2001) aborda a necessidade de descrições arquivísticas que retomem o contexto administrativo dos documentos, além de uma definição mais aprofundada sobre o seu conteúdo, apresentando elementos da estrutura organizacional, do contexto e do conteúdo dos conjuntos documentais. Com vistas às relações intelectuais que são estabelecidas entre esses elementos, destaca a necessidade de abordar o processo de descrição como uma forma contínua e fluida de representação da informação. A descrição deve se iniciar na criação do documento, acontecer durante seu trâmite e, também, depois da definição da sua destinação final.

A descrição arquivística, segundo MacNeil (2005), visa à promoção do acesso a documentos arquivísticos; ao entendimento do acervo a partir do registro do seu contexto, da estrutura e do conteúdo; e ao registro da cadeia de custódia dos documentos, seu arranjo, contexto de produção e tramitação para que seja possível presumir a autenticidade. Nessa perspectiva, a descrição supera o viés de instrumento de pesquisa para arquivos permanentes e passa a ser um complexo processo que pode auxiliar na garantia de autenticidade dos documentos de arquivo, assim como na preservação da sua integridade. Essa perspectiva do processo de descrição aplicado desde a produção dos documentos resguarda a autenticidade dos documentos e salienta aos usuários uma imagem confiável do arquivo e de quem possui sua custódia. Contudo, apesar da busca pela descrição completa, deve-se ter em mente que não

é possível representar todo o conteúdo do acervo, dada a complexidade do processo e os diversos recursos nele envolvidos.

Tradicionalmente, o arquivista elabora e se apoia em instrumentos de pesquisa que representam o conteúdo dos documentos (guia, inventário, catálogo, catálogo seletivo, índice, edição de fontes). Esses instrumentos são um produto da descrição de documentos, mas não são suficientes para representar todos os significados do conjunto documental. Por isso, é tão necessário ao arquivista conhecer a estrutura do produtor documental, como mediador das informações orgânicas registradas. Lopes (2009) afirma que “é preciso saber ler, interpretar e escrever, estar em sintonia com os problemas do mundo contemporâneo, em especial com os do seu país, conhecer métodos e técnicas de pesquisa e, sobretudo, estar pronto para aprender a aprender”. Nesse aspecto, a explicitação do contexto ganha cada vez mais força no processo de descrição dos documentos.

Os instrumentos de descrição arquivística devem, assim, retomar o contexto de produção e acumulação dos documentos, permitindo ao arquivista conhecer o conteúdo do fundo, além de orientar os usuários e facilitar a busca pela informação. Conforme Lopes (2009), a documentação organizacional deve ser produzida e descrita de forma normalizada e padronizada, conforme previsto nas normas internacionais e nacionais de descrição. Isso significa que ela deverá ser consistente, adequada e autoexplicativa; além de facilitar a recuperação dos documentos, de tornar a informação um ativo organizacional e proporcionar a sua troca.

CONSIDERAÇÕES FINAIS

Relacionada às demais funções arquivísticas no seu objetivo comum de possibilitar e facilitar o acesso aos documentos/informações arquivísticas, a descrição cumpre essencial papel desde a criação dos documentos até a elaboração dos instrumentos de pesquisa, ao representar as informações contidas nesses documentos e o contexto funcional da sua produção e acumulação. Ao reconstruir o contexto arquivístico dos documentos pertencentes ao fundo durante a descrição, o arquivista deve evidenciar os fatos, os processos de produção e preservação, bem como as tomadas de decisões que participaram da origem dos documentos que estão sob sua análise. Por isso, ter disponível uma definição de elementos que torne possível a descrição de documentos em qualquer fase do ciclo documental corrobora para assegurar a integridade e fidedignidade dos conjuntos documentais.

Se, por um lado, o Estado tem o dever de preservar os seus arquivos como patrimônio histórico-cultural, por outro lado, o cidadão tem o direito de a eles recorrer para satisfazer as suas necessidades de informação. O arquivista, como mediador desses processos – planejamento, organização, preservação, promoção do acesso e usos dos documentos – deve pautar a sua atuação nos princípios científicos, nos métodos e nas técnicas arquivísticas, alinhados às demandas contemporâneas de informação, perpassadas por desafios diversos. Desse modo, deve estudar e aplicar de forma crítica os fundamentos arquivísticos, explorando as potencialidades e a viabilidade da descrição de documentos arquivísticos desde a sua produção, superando as limitações da sua realização somente nos arquivos permanentes. Afinal, há documentos que já são produzidos com valor histórico e há documentos históricos que são reivindicados para demandas administrativas, ultrapassando as fronteiras das três idades documentais anunciadas em parte da literatura da área.

Compreendemos que as principais funções da descrição arquivística relacionam-se (mas não se restringem) à elaboração de mecanismos de acesso a partir da explicitação das relações entre os documentos e as funções que o geraram. Analisar essas funções a fundo evidencia a pertinência da descrição de documentos nas fases corrente, intermediária e permanente. Afinal, os objetivos de controle e acesso estão presentes em todas as fases documentais e a sua aplicação deve estar aliada ao emprego do plano de classificação de documentos e ao controle da sua temporalidade, a fim de que se definam as relações entre os documentos e se preserve as suas inter-relações.

Nesse sentido, a descrição do contexto de produção/acumulação de documentos, bem como da missão, das funções, das atividades e da estrutura institucional (considerando as suas mudanças ao longo do tempo) no âmbito do planejamento da sua criação é vislumbrada como um processo necessário para a economia de recursos (temporais, materiais e humanos), além de possibilitar a celeridade no processo de recuperação da informação já na sua gênese. Dessa forma, os vários desafios para a gestão estratégica da informação (como a grande quantidade de informações produzidas, acumuladas, armazenadas e distribuídas; e a alta velocidade de mutação das formas e dos conteúdos) podem ser superados se a organização investir insumos no planejamento e na implementação de políticas que visem à gestão estratégica da informação.

A descrição padronizada dos metadados, de acordo com tesauros, pode também otimizar a indexação de documentos e, conseqüentemente, agilizar o seu trâmite e a sua recuperação. E até mesmo nas atividades mais simples, como na elaboração de espelhos e guias-fora e no endereçamento de caixas e pastas, a descrição pode ser um processo relevante para a organização padronizada dos documentos com valor predominantemente administrativo, de forma a resguardar os seus vínculos com o contexto funcional de produção/acumulação e com os outros documentos que estejam nesse contexto.

Considerando essas leituras que redimensionam o processo de descrição de forma a contemplá-lo em todas as fases do ciclo documental, observamos que a comunidade arquivística parece conceber a descrição como um procedimento mais econômico, visto que ele passa a ser implementado desde a origem dos documentos, mas de forma que o acesso, a utilização e o entendimento dos conteúdos representados se tornem cada vez maiores, com custos mais dosados.

Os resultados da pesquisa bibliográfica em curso demonstram que, embora os fundamentos epistemológicos da Arquivologia ainda ressoem como basilares, há que se pensar na sua viabilidade, de maneira a se reformular os processos de criação, produção, organização, representação, recuperação, acesso e utilização dos documentos de arquivo. Amparados por instrumentos conceituais, metodológicos e produtos advindos do trabalho do arquivista, como manuais de gestão de documentos, planos de classificação, tabelas de temporalidade e instrumentos de pesquisa, deve-se redimensionar os fundamentos arquivísticos mediante os desafios contemporâneos, ampliados e aprofundados quanto à organização e à recuperação dessas informações diante das leis de acesso à informação.

Portanto, há avanços no estudo da descrição como um amplo processo, até mesmo já normatizado internacionalmente, que merecem ser estudados contingencialmente, para atender demandas particulares de informações, de modo a explorar a potencialidade informativa dos documentos de arquivo, independentemente da fase ou dos valores que majoritariamente os perpassem.

REFERÊNCIAS BIBLIOGRÁFICAS

- Bellotto, H. L. (2006). Arquivos permanentes: tratamento documental. 4. ed. Rio de Janeiro: Ed. FGV.
- Capurro, R., Hjørland, B. (2007). O conceito de informação: Perspectivas em Ciência da Informação, 12, (1), p. 148-207.
- Conselho Internacional de Arquivos. (2000). Comitê de Normas de Descrição. ISAD(G): Norma geral internacional de descrição arquivística. 2. ed. Rio de Janeiro: Arquivo Nacional.
- Cruz Mundet, J. R. (2001). Manual de arquivística. Madrid: Fundación Germán Sánchez Ruipérez. Biblioteca del Libro.
- Duranti, L. (1994) Registros documentais contemporâneos. Estudos Históricos, Rio de Janeiro, 7, (13), p. 49-64.
- Esteban Navarro, M. A. (1995). La representación y la organización del conocimiento en los archivos: los lenguajes documentales ante los procesos de clasificación, ordenación y descripción. In: GARCÍA MARCO, F. J. (ed.). Organización del conocimiento en sistemas de información y documentación. Actas del I Encuentro de ISKO. España, Madrid, 4-5 noviembre de 1993. Librería General: Zaragoza, Espanha, p. 65-90.
- Hagen, A. M. (1998) Algumas considerações a partir do processo de padronização da descrição arquivística. Ciência da Informação, Brasília, 27(3), p. 293-299.
- Haworth, K. M. (2001) Archival description: content and context: in search of structure. In: PITTI, D. V.; DUFF, W. M. (Orgs.). Encoded Archival Description on the Internet. New York: The Haworth Information Press.
- Heredia Herrera, Antonia. (1988). Arquivística general: teoría y práctica. Sevilla: Diputación.
- Lopes, L. C. (2000) A nova Arquivística na modernização administrativa. Rio de Janeiro: [s.n.].
- Lopes, L. C. (2009) A nova arquivística na modernização administrativa. Brasília: Projecto Editorial.
- MacNeil, H. (2005) Picking our text: archival description, authenticity, and the archivist as editor. The American Archivist. Chicago: The American Archivist, v. 68, p.264-278, Fall/Winter.
- Mueller, S.; Feith, A.; Fruin, R. (1960). Manual de Arranjo e Descrição de Arquivos. Rio de Janeiro: Ministério da Justiça e Negócios Interiores; Arquivo Nacional.
- Oliveira, L. M. V. de. (2012). Descrição e pesquisa: reflexões em torno dos arquivos pessoais. Rio de Janeiro: Móbile Editorial, p. 41-70.
- Rodrigues, G. M. A representação da informação em Arquivística: uma abordagem a partir da perspectiva da Norma Internacional de Descrição Arquivística. In RODRIGUES, G. M., & LOPES, I. L. (orgs.) (2003). Organização e representação do conhecimento na perspectiva da ciência da informação. Brasília: Thesaurus.
- Schellenberg, T. R. (2006) Arquivos Modernos: princípios e técnicas. 6ª ed. Rio de Janeiro: Editora FGV.

Sousa, R.T.B. (2014). Alguns apontamentos sobre a classificação de documentos de arquivo. *Brazilian Journal of Information Science: new research trends*, 8, (1/2), p. 1-24.

**FORMAÇÃO E ATUAÇÃO PROFISSIONAL
EM ORGANIZAÇÃO DO CONHECIMENTO**



ATUAÇÃO DO PROFISSIONAL DA INFORMAÇÃO NA ORGANIZAÇÃO DO CONHECIMENTO NA WEB DE DADOS

Francisco Carlos Paletta¹, Armando Manuel Barreiros Malheiro da Silva²

¹Universidade de São Paulo, 0000-0002-4112-5198, fcpaletta@usp.br

²Universidade do Porto, 0000-0003-0491-3758, malheiro@letras.up.pt

RESUMO A inclusão digital está também relacionada ao desenvolvimento nos indivíduos de habilidades e competências para acesso, uso e disseminação eficiente da informação. Além disso, a evolução contínua tecnológica, tal como a emergência da Web semântica, leva a novos perfis e demandas do profissional da informação. A organização e gestão da informação com grandes coleções de materiais necessitam reavaliar constantemente a identidade e propósitos principais de seus usuários. Bibliotecas estão desenvolvendo continuamente processos e serviços de forma a explorar os benefícios da Web Semântica. Neste trabalho, discutimos as principais tendências nesse campo, em especial aquelas que influenciam o contexto da Web de Dados.

PALAVRAS-CHAVE *Organização da Informação, Gestão do Conhecimento, Web de Dados, Era Digital, Profissional da Informação.*

ABSTRACT Digital inclusion is also related to the development of individuals in skills and competencies to access, use and efficiently dissemination of information. Besides, continuous technological developments such as the emergence of the semantic Web, leads to new professional profiles and demands. The design and management of information systems with large collections of materials need to constantly re-evaluate the identity of its users and their main purposes. Libraries are continually developing processes and services to exploit the benefits of the Semantic Web, for instance. In this paper, we discuss the main trends in this field, especially those that influence the context of the Web of Data.

KEYWORDS *Information Organization, Knowledge Management, Web of Data, Digital Age, Information Professional.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A Web é uma rede com conteúdos interligados através de documentos de hipertexto. Seu mapeamento é possível por processos de análise e coleta sucessiva de páginas de conteúdo a partir de um conjunto de localizações de documentos previamente conhecido. Tais buscas são feitas de forma automática por programas de computador normalmente denominados *crawlers*, coletores ou batedores. Mesmo neste mapeamento automático, o conjunto inicialmente assumido de localizações a partir do qual a pesquisa é feita influencia o resultado obtido. Além disso, nem toda a Web está interligada: existem muitos subconjuntos de documentos interligados entre si, de tamanhos variados, sem ligação com o restante da

rede (ou seja, “ilhas” de informação). O primeiro princípio da Web, proposto pelo Consórcio World Wide Web W3C Brasil afirma (W3C, 2011):

“O principal valor da Web é o social. Mais do que tecnológico, este é um ambiente de comunicação humana, de transações comerciais, de oportunidades para compartilhar conhecimentos e, para ser um ambiente universal, deve estar disponível para todas as pessoas, independentemente dos equipamentos e softwares que utilizem, principalmente da cultura em que se inserem, da localização geográfica, das habilidades físicas ou mentais, das condições socioeconômicas ou de instrução”.

A universalidade da Web só pode ser garantida e aprofundada com um modelo de governança democrático e pluralista que tenha foco no acesso por todos e na sua própria evolução tecnológica (CGI.br, 2011).

A temática da inclusão digital vem sendo tratada desde os anos 90 como a necessidade de permitir o acesso a computadores e ferramentas de TICs (Tecnologia da Informação e Comunicação). No entanto, novas proposições têm tratado da inclusão não apenas no acesso às ferramentas digitais, mas também nos aspectos relativos ao acesso e uso dos conteúdos distribuídos. Promover a inclusão digital e por consequência a inclusão social não significa apenas promover universalidade das ferramentas de TICs, mas possibilitar seu uso de forma crítica, estimulando o aperfeiçoamento das potencialidades informativas e cognitivas e atividades cidadãs. A preocupação deixa de ser o aprendizado de ferramentas básicas de navegação na WEB e desloca-se para as diferentes formas de coleta e de produção do conhecimento. Tornam-se centrais questões de inclusão social através da inclusão digital e das práticas sociais e educacionais vigentes (Passarelli 2012).

A expressão Competência Informacional, ou *Information Literacy*, tem suas origens no estudo da sociedade da informação, e está ligada à necessidade de se desenvolver nos indivíduos aptidões sobre habilidades e competências relacionadas ao acesso, uso e disseminação da informação, objetivando fazer uso desta de forma ética e eficiente, para que o ser humano através de seu intelecto e processo cognitivo possa produzir novo conhecimento (Campello 2003).

Literacia informacional tem constituído, nos últimos anos, novo campo de pesquisa que interessa à educação, à ciência da informação e às ciências cognitivas. O desafio está em inicialmente aprender a utilização básica dos recursos tecnológicos - literacia digital - e a seguir apropriar-se dos mesmos para gerar novos conhecimentos - literacia informacional. No contexto da literacia informacional, os desafios são múltiplos: a dificuldade das bibliotecas universitárias, com suas coleções caríssimas, em concorrer com acervos digitais; as bibliotecas virtuais e os motores de busca oferecendo informações de forma imediata e a custo zero; a desconstrução do conceito de autor individual e a emergência dos coletivos digitais. Os novos contornos da sociedade em rede têm constituído objeto de estudo e pesquisa para profissionais das mais diversas áreas de atuação.

Esta pesquisa de caráter qualitativo e exploratório tem como objetivo proporcionar uma reflexão sobre a atuação do profissional da informação na organização do conhecimento na Web de Dados, com vistas a torná-lo mais explícito ou a construir hipóteses.

O PROFISSIONAL DA INFORMAÇÃO E A ERA DIGITAL

No contexto mundial definido pela globalização e pela constante mudança tecnológica, o conhecimento tornou-se a principal riqueza das nações, pode também constituir o principal fator de desigualdade nas sociedades atuais. A Sociedade da Informação é a pedra angular das Sociedades do Conhecimento. O conceito de “sociedade da informação”, relacionado à ideia da inovação tecnológica, caracteriza-se pelo rápido crescimento da informação disponibilizada e as mudanças ocasionadas pela tecnologia usada para gerar, disseminar, acessar e usar a informação (Melo 2007).

Por outro lado, o conceito de “sociedade do conhecimento” inclui uma dimensão de transformação social, cultural, econômica, política e institucional, assim como uma perspectiva mais pluralista e de desenvolvimento, expressando a complexidade e o dinamismo das mudanças que estão ocorrendo. O conhecimento não só é importante para o crescimento econômico, mas também para fortalecer e desenvolver todos os setores da sociedade. De acordo com Unesco (2017):

“O conhecimento e a informação têm impacto significativo na vida das pessoas. O compartilhamento de conhecimento e informação, particularmente por meio das tecnologias de informação e comunicação (TIC), tem o poder de transformar economias e sociedades... as sociedades do conhecimento devem ser construídas sobre quatro pilares: respeito às diversidades cultural e linguística; liberdade de expressão; acesso universal à informação e ao conhecimento; e educação de qualidade para todos”.

A noção de fácil acesso à informação propiciado por avanços da tecnologia de redes de computadores e de telecomunicações criou uma noção errônea do imperativo tecnológico como resposta às deficiências comunicacionais e educacionais da humanidade. O conhecimento e o uso deste ferramental tecnológico são essenciais nos dias de hoje, porém é preciso considerar que por si só a tecnologia não leva à comunicação e à educação. A competência informacional está fortemente relacionada ao processo de interiorização de conhecimentos, habilidades e valores ligados à informação e ao aprendizado. Em um contexto mais prático, representa um conjunto de habilidades necessárias para localizar, interpretar, analisar, sintetizar, avaliar e comunicar a informação em diferentes ferramentas e suportes (Dudziak 2001; 2003).

Indivíduos têm seus estoques de informação e estes são usados para compreender as informações externas e as diferentes situações em que se encontram em dado momento. O comportamento de busca e uso de informação são modelados, do ponto de vista epistemológico, pelas características cognitivas do indivíduo e por fatores que geram o encontro do usuário com os sistemas de informação ou as consequências de tal confronto. A importância do aspecto cognitivo para o comportamento de busca de informação é apresentada por Kuklthau (2008):

“Vários estudos apoiaram a noção de busca de informação como um processo de construção do conhecimento com diferentes estágios cognitivos e afetivos (Cole 1997; George et al., 2006; Harada 2002; Pitts 1995; Tang e Salomão 1998; Serola e Vakkari 2005; Vakkari 2001; Vakkari et al., 2003, Vakkari e Hakala 2000, Wang e Soergel 1998, Yang, 1997). O modelo do Processo de Pesquisa de Informação muitas vezes causa ansiedade e incerteza estudos confirmaram essa noção de interação entre os fatores cognitivos, afetivos e comportamentais na busca da informação (Bilal 2000; 2001; 2002; Bilal e Kirby 2002; George et al., 2006; Heinstrom 2002; Jiao e Onwuegbuzie 1997; 1999; Jiao et al. 2006, Jiao et al. 1996; Mellon 1988; Onwuegbuzie 1997; Onwuegbuzie e Jiao 1998; 2004; Wang e Soergel, 1998)”.

É preciso atentar para o fato de que não é possível mais se limitar à tarefa de localizar fontes de informação. É necessário levar em consideração tarefas de interpretação, formulação e aprendizagem envolvidas no processo de busca dessa informação. O aumento no acesso à vasta quantidade de informação requer, entretanto, serviços que se centrem no significado da busca mais do que meramente na localização da fonte. Nessa perspectiva, os usuários da informação não podem ser vistos apenas como integrantes do sistema, mas como a “razão de ser” do serviço de informação.

Sistemas de informação organizados na perspectiva tradicional concentram-se prioritariamente na aquisição e administração de grandes coleções de materiais. Assumiu-se, durante décadas, que as atividades técnicas dos sistemas eram o seu ponto estratégico. Considerava-se que os usuários utilizavam o sistema exatamente da maneira como estes tinham sido projetados. Não se imaginava indagar, aos sistemas, questões imprescindíveis sobre a identidade e propósitos principais de seus usuários. Como a informação era considerada algo existente fora das pessoas e passível de ser transferida de uma para outra, parecia ser possível que eficiência e sucesso das operações de um sistema pudessem ser medidos em função do número de fontes de informações recuperadas pelo sistema versus o que realmente foi de interesse do usuário. Isso, na realidade, coloca novamente o usuário como um processador imperfeito da informação, pois é já sabido que nem todas as pessoas se interessam pelas mesmas fontes indicadas. Resulta desse procedimento que hoje em dia se conhece muita coisa sobre planejamento, aquisição, organização, controle e desenvolvimento de coleções, mas muito pouco sobre como as pessoas fazem uso dos sistemas ou para que fins e como a informação, a matéria-prima dos sistemas, está sendo utilizada (Ferreira 1995).

A ORGANIZAÇÃO DA INFORMAÇÃO NA WEB DE DADOS

O impacto do uso da Web na sociedade, nos indivíduos e nas organizações tornou-se objeto de pesquisa, extrapolando o campo especializado da computação aplicada, e atingindo áreas de estudos organizacionais e sociológicos. Por ser essencialmente dinâmica e sem fronteiras, tanto do ponto de vista físico como virtual, é importante que seja conhecida em detalhes, tanto para assegurar sua livre transformação quanto para permitir sua disponibilidade, confiabilidade e acessibilidade por todos.

A denominada Web 2.0 segue uma filosofia com princípios de leitura e escrita de natureza participativa, em que cada usuário pode intervir diretamente na escolha e introdução de dados no âmbito de cada site; é cooperativa, uma vez que compartilha ideias, preferências, informações e conhecimento; é interativa, na medida em que, através de toda a gama de recursos multimídia, é possível um diálogo simultâneo com os usuários; é democrática, pois sob essa filosofia existe liberdade de expressão, de pensamento, e, sobretudo, de trânsito de informações, independentemente dos interesses de cada um; é também sociotécnica, pois, através de todas as suas características, é possível um intercâmbio de culturas, religiões, etnias e outros. Hoje, não só o usuário participa como também adiciona conteúdo que ao serem compartilhados em outros recursos colaborativos passam a receber comentários dos membros vinculados a sua comunidade agregando valor à informação compartilhada, além de criar vários caminhos para a localização do próprio recurso. Assim, temos uma nova relação e interação na qual aumenta o nível sofisticado de atitudes sociais eletrônicas.

A Web Semântica apresenta um novo modelo de interatividade entre o usuário da informação e a internet. Esta deve idealmente apresentar toda sua informação rotulada com indicações de contexto e significado, organizada de forma que tanto humanos como máquinas possam entendê-la. Surge uma nova denominação para diferentes evoluções pretendidas na Web e na interação desta com o usuário: a Web 3.0 (Passarelli 2008). Entre os muitos aspectos incluídos nessa denominação, destacam-se:

- a) Web semântica;
- b) Web geoespacial;
- c) Web 3D e de realidade aumentada;
- d) Conteúdo Web mais dinâmico, agregando inteligência artificial;
- e) Maior compatibilidade da Web com a tecnologia de banco de dados;

Neste ambiente contemporâneo, surge um novo usuário da informação com novas demandas por recursos computacionais e novas capacidades em produzir novos conhecimentos. Considerando-se os principais aspectos da Competência Informacional: a Tecnologia da Informação, as Fontes de Informação, os Processos de Informação, o Controle da Informação, a Construção do Conhecimento, a Extensão do Conhecimento e a Inteligência (Bruce 2003), destaca-se um profissional ativamente envolvido com os processos de ensino e aprendizado em Competência Informacional: o bibliotecário (Souza 2010).

PROCEDIMENTOS METODOLÓGICOS

Esta pesquisa de caráter qualitativo e exploratório tem como objetivo proporcionar uma reflexão sobre a atuação do profissional da informação na organização do conhecimento na web de dados, com vistas a torná-lo mais explícito ou a construir hipóteses. A pesquisa preocupa-se, portanto, com aspectos de realidade que não pode ser quantificada, centrando-se na compreensão e explicação da dinâmica das relações sociais. A revisão da literatura apresenta um panorama da educação com os novos recursos informacionais e digitais, as literacias da informação, a transdisciplinaridade entre o campo da Biblioteconomia, Informação e Comunicação e a crescente tendência das Escolas de Informação - iSchools. Os procedimentos metodológicos adotados para este trabalho compreendem duas dimensões: teórica e aplicada. Do ponto de vista teórico, caracteriza-se como sendo do tipo descritiva exploratória de natureza qualitativa, devido ao fato de ter como objetivo a reunião de dados e informações sobre os temas abordados.

RESULTADOS

Saberes científicos e o acesso a inovações tecnológicas apresentam-se distribuídos de forma muito desigual, considerando-se nesta comparação países, regiões, estratos sociais, faixas etárias, níveis educacionais, dentre outros. Assim, a problemática da diversidade cultural e os estudos sobre ela devem fazer parte da consideração teórica, da investigação empírica e do planejamento de políticas na área de fomento à Competência Informacional. Além disso, o usuário da informação típico não deseja apenas extrair informações específicas e definitivas no menor tempo possível ou, alternativamente, se dispõe a investir esforço para buscar e explorar informação e conhecimento. A verdade é que as pessoas oscilam continuamente entre extrair e explorar, e o uso da informação é um processo confuso, desordenado, sujeito aos caprichos da natureza humana, como qualquer outra atividade (Canclini 2009). Entre os

elementos mais importantes que influenciam o uso da informação estão as atitudes do indivíduo em relação à informação e a sua busca, atitudes essas que são fruto da educação, do treinamento, da experiência passada, das preferências pessoais.

O universo da informação digital se expande no contexto do que hoje chamamos de Big Data, onde a busca por informações fica cada vez mais difícil uma vez que o universo digital está constituído por dados não estruturados que precisam ser organizados, acessados, apropriados e então utilizados na produção de novos conhecimentos. A Web de Dados virá com o objetivo de organizar estas informações para que os usuários tenham mais facilidade na busca da informação. A busca Semântica como é chamada os buscadores da Web de Dados, organizam informações por assuntos determinados, conectando conhecimentos. A busca Semântica da Web de Dados dividirá os resultados em categorias para que a busca da informação seja mais rápida e organizada.

A Figura 1 relaciona o Grau de Conectividade da Informação com o Grau de Conectividade Social de acordo com a evolução tecnológica: Web - Conectando Informação, Web Social - Conectando Pessoas, Web Semântica – Conectando Conhecimento e a Metaweb – Conectando Inteligência.

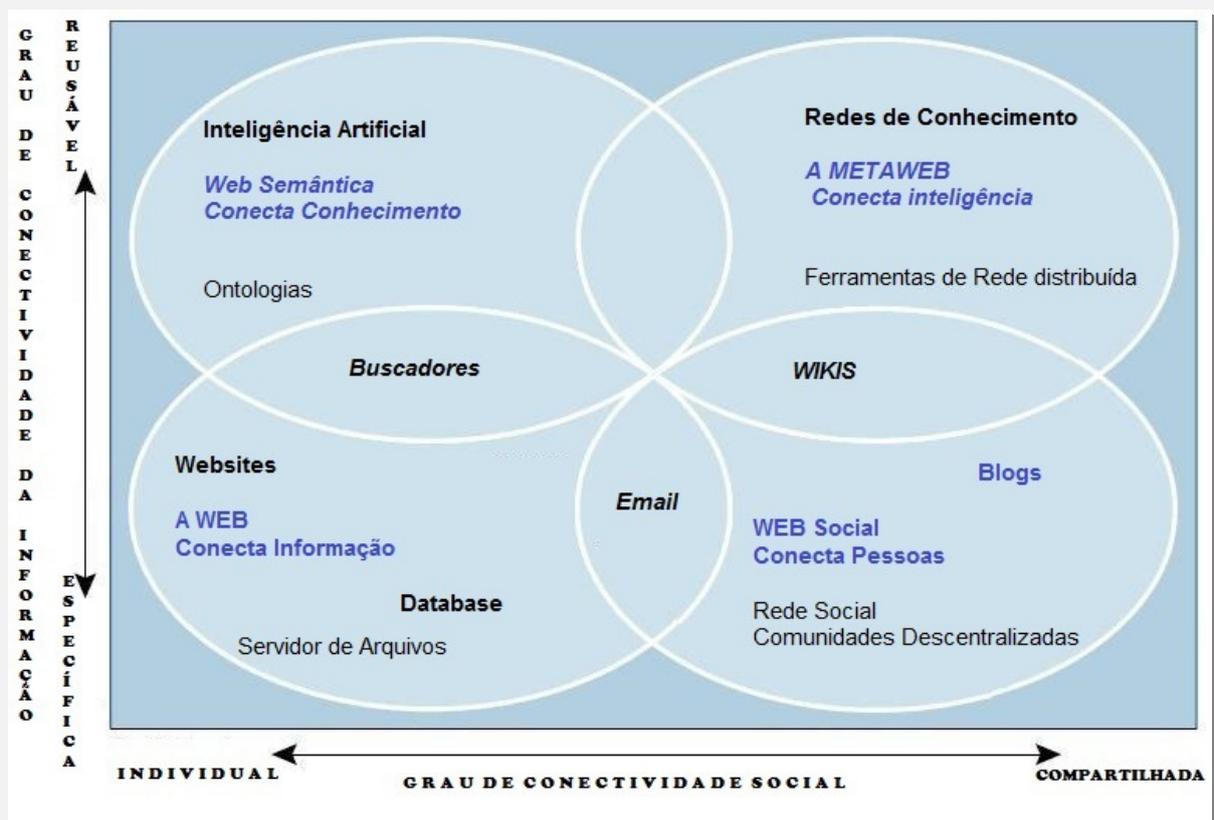


Figura 1. Evolução Tecnológica da Web – Adaptação do Autor

Fonte: <http://www.novaspiavack.com/science/new-version-of-my-metaweb-graph-the-future-of-the-net>

A organização do conhecimento liga os três processos de uso estratégico da informação – a criação de significado, a construção do conhecimento e a tomada de decisões – num ciclo contínuo de aprendizagem e adaptação que podemos chamar de ciclo do conhecimento.

Bibliotecas estão desenvolvendo continuamente processos e serviços de forma a explorar os benefícios da Web 2.0. Library 3.0 refere-se a bibliotecas utilizando tecnologias como a Web Semântica, a

computação em nuvem, dispositivos móveis, e re-imaginando o uso de tecnologias estabelecidas, tais como pesquisa federada (RFID), para facilitar a disseminação do conteúdo gerado pelo usuário e a colaboração para promover e fazer coleções de bibliotecas acessíveis. O resultado final da Biblioteca 3.0 é a expansão da "biblioteca sem fronteiras", onde as coleções podem ser feitas facilmente disponíveis aos usuários da biblioteca, independentemente da sua localização física. Biblioteca 3.0 é um complemento virtual para os espaços físicos da biblioteca (Belling 2013).

É neste contexto de crescente disponibilidade de dados que a “era digital” se apresenta com múltiplas oportunidades de atuação para o profissional da informação no uso das tecnologias da informação e comunicação na gestão e organização da informação na Web de Dados.

CONCLUSÕES

Há atualmente a necessidade de novas propostas para abordagens gerenciais e estratégicas associadas à utilização da Tecnologia da Informação na apropriação e geração de conhecimento no contexto da Web Inteligente.

Diante da emergência de uma economia digital - formação de padrões complexos a partir de uma multiplicidade de interações, cumpre-nos avaliar as contínuas tensões na educação entre a Ciência da Informação e a Biblioteconomia neste novo cenário orientado à formação de recursos humanos: futuros bibliotecários – profissionais da informação. É preciso preparar profissionais da informação habilitados à lidar com a complexidade do mercado de trabalho global, proporcionando uma revisão sistemática nos planos de ensino associada às principais questões que estão à mudar em sua área de conhecimento no contexto social, político, econômico, tecnológico e cultural: o impacto dos dispositivos digitais e redes sociais, a evolução dos serviços bibliotecários incluindo referência virtual, o suporte ubíquo a bibliotecas, acesso digital a repositórios, preservação digital, os novos esforços para organizar o conhecimento, a Web Semântica, o catálogo da biblioteca de próxima geração, o impacto da publicação digital e dos e-books, as questões políticas relacionadas com o acesso em banda larga e a neutralidade da rede, novas interpretações de direitos autorais relacionados com a digitalização em massa de livros e artigos acadêmicos, novas iniciativas para integrar bibliotecas, arquivos e museus (RUBIN, 2010).

O desafio imposto pela Era Digital em entender o papel da informação, e em particular na área da Ciência da Informação, no cenário que os especialistas chamam de a “4 ° Revolução Industrial” - iniciada no início da década de 90 com ascensão da desmaterialização de funções e processos e do trabalho do conhecimento. A formação e capacitação de recursos humanos, cooperação tecnológica, uso em larga escala das tecnologias da informação e comunicação TICs estão à pavimentar o caminho da revolução digital.

Na sociedade em contínua mudança, temos que repensar os modelos de ensino e observar as complexidades do mercado de trabalho 4.0 e suas relações com a Informação, Comunicação e Cultura e estruturar programas que permitam formar o profissional da informação com as competências necessárias para atuar em um mercado de trabalho cada vez mais dependente das competências digitais.

Na Era da Informação e, mais precisamente, na conjuntura de rede onde estamos imersos, muitos são os desafios e os problemas que não podemos apenas enfrentar, mas devemos estudar com minúcia e profundidade a fim de irmos encontrando as respostas globais e específicas mais adequadas (SILVA, 2008).

Neste trabalho, discutimos as tendências nesse campo, em especial interesse aquelas que influenciam o contexto da chamada Biblioteca 3.0, de modo que, cada vez mais, a tecnologia possa proporcionar a criação e obtenção de valor ao usuário da informação. Espera-se que o desenvolvimento dessa área leve a um maior entendimento dos recursos da tecnologia colaborativa utilizados em ambientes informacionais digitais. Em qualquer caso, a inserção dessas tecnologias consiste em inovação que deve estar vinculada à tradição e à missão das bibliotecas e dos repositórios bem como na atuação do profissional da informação na organização e gestão da informação na Web de Dados.

ACKNOWLEDGMENT

FAPESP Research Project – Processo 2016/07358-6

REFERÊNCIAS BIBLIOGRÁFICAS

Belling, A. (2013). Exploring Library 3.0 and beyond. Retrieved 18 June, 2017 from http://www.libraries.vic.gov.au/downloads/20102011_Shared_Leadership_Program_Presentation_Day_exploring_library_3.pdf

Bilal, D. (2000). Children's use of the Yahoooligans! Web search engine: I. Cognitive, physical and affective behaviours on fact-based search tasks. *Journal of the American Society for Information Science*, 51(7), 646-665.

Bilal, D. (2001). Children's use of the Yahoooligans! Web search engine: II. Cognitive and physical behaviors on research tasks. *Journal of the American Society for Information Science and Technology*, 52(2), 118-136.

Bilal, D. (2002). Children's use of the Yahoooligans! Web search engine: III. Cognitive and physical behaviors on fully self-generated search tasks. *Journal of the American Society for Information Science and Technology*, 53(13), 1170-1183.

Bilal, D. & Kirby, J. (2002). Differences and similarities in information seeking: children and adults as Web users. *Information Processing and Management*, 38(5), 649-670

Bruce, C. (2003). Seven Faces of Information Literacy. Retrieved 18 June, 2017 from <http://www.bestlibrary.org/digital/files/bruce.pdf>

Canclini, N, G. (2009). *Diferentes, desiguais e desconectados*. Rio de Janeiro: Editora UFRJ. ISBN 857108291X

CGI.br (2011). *Dimensões e características da WEB brasileira: um estudo do .gov.br*. <http://www.cgi.br/media/docs/publicacoes/2/cgiibr-nicbr-censoweb-govbr-2010.pdf>

Cole, C. (1997). Information as process: The difference between corroborating evidence and “information” in humanistic research domains. *Information Processing and Management*, 33(1), 55-67.

Dudziak, E, A. (2001). *A Information Literacy e o papel educacional das bibliotecas*. Dissertação (Mestrado) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2001.

Dudziak, E. A. (2003). "Information literacy: princípios, filosofia e prática". *Ciência da Informação*, Brasília, v.32, n.1, p. 23-35.

Ferreira, S, M, S, P. (1995). "Novos paradigmas e novos usuários da informação". *Ciência da Informação*, Brasília, v.25, n.2, p. 217-223.

George, C., Bright, A., Hurlbert, T., Linke, E. C., St. Clair, G. & Stein, J. (2006). Scholarly use of information: graduate students' information seeking behaviour. *Information Research*, 11(4), paper 272. Retrieved 18 June, 2017 from <http://informationr.net/ir/11-4/paper272.html> (Archived by WebCite® at <http://www.webcitation.org/5cbDFityo>)

Harada, V. H. (2002). Personalizing the information search process: A case study journal writing with elementary-age students. *School Library Media Research*, 5. Retrieved 18 June, 2017 from http://www.ala.org/aasl/sites/ala.org.aasl/files/content/aaslpubsandjournals/slr/vol5/SLMR_PersonalizingInfoSearch_V5.pdf (Archived by WebCite® at <http://www.webcitation.org/5cbDWgyrj>)

Heinström, J. (2002). Fast surfers, broad scanners and deep divers – personality and information seeking behaviour. Åbo (Turku), Finland: Åbo Akademi University Press. (Doctoral dissertation.) Retrieved 18 June, 2017 from http://www.abo.fi/fakultet/media/21373/thesis_heinstrom.pdf

Jiao, Q. G. & Onwuegbuzie, A. J. (1997). Antecedents of library anxiety. *The Library Quarterly*, 67(4), 372-389.

Jiao, Q. G. & Onwuegbuzie, A. J. (1999). Identifying library anxiety through students' learning modality preferences. *Library Quarterly*, 69(2), 202-216.

Jiao, Q. G., Onwuegbuzie, A. J. & Bostick, S. L. (2006). The relationship between race and library anxiety among graduate students: a replication study. *Information Processing and Management*, 42(3), 843-851.

Jiao, Q. G., Onwuegbuzie, A. J. & Lichtenstein, A. (1996). Library anxiety: characteristics of 'at-risk' college students. *Library and Information Science Research*, 18(2), 151-163.

Mellon, C.A. (1988). Attitudes: the forgotten dimension in library instruction. *Library Journal*, 113(14), 137-139.

Onwuegbuzie, A.J. (1997). Writing a research proposal: the role of library anxiety, statistics anxiety and composition anxiety. *Library and Information Science Research*, 19(1), 5-33.

Onwuegbuzie, A. J. & Jiao, Q. G. (1998). The relationship between library anxiety and learning styles among graduate students: Implications for library instruction. *Library and Information Science Research*, 20(3), 235-249.

Onwuegbuzie, A. J. & Jiao, Q. G. (2004). Information search performance and research achievement: an empirical test of the Anxiety-Expectation Mediation model of library anxiety. *Journal of the American Society for Information Science and Technology*, 55(1), 41-54.

Passarelli, B. (2008). "Do Mundaneum à WEB Semântica: discussão sobre a revolução nos conceitos de autor e autoridade das fontes de informação". *DataGramaZero - Revista de Ciência da Informação* - v.9 n.5, P.1-13.

Passarelli, B.; Junqueira, A, H. Gerações Interativas Brasil. Crianças e Adolescentes Diante das Telas. <http://ccvap.futuro.usp.br/gerinter2012.pdf>

Pitts, J. M. (1995). Mental models of information. The 1993-1994 AASL/Highsmith Research Award Study. *School Library Media Quarterly*, 23(3), 177-184.

RUBIN, R. (2010). *Foundations of library and information science*. New York: Neal Schuman.

Serola, S. & Vakkari, P. (2005). The anticipated and assessed contribution of information types in references retrieved for preparing a research proposal. *Journal of the American Society for Information Science and Technology*, 56(4), 373-381.

SILVA, Armando Malheiro da (2008). Inclusão digital e literacia informacional em ciência da informação. *Prisma.Com*, Porto, 7 (Jul.).p.16-43. ISSN: 1646-3153. Url: <http://revistas.ua.pt/index.php/prisma.com/article/view/683/pdf>

Sousa, R, S, C., Nascimento, B, S. (2010). “Competências Informacionais: uma análise focada no currículo e na produção docente dos cursos de Biblioteconomia e gestão da informação”. *Revista ACB: Biblioteconomia em Santa Catarina, Florianópolis*, v.15, n.2, p.130-150.

Tang, R. & Solomon, P. (1998). Toward an understanding of the dynamics of relevance judgment: an analysis of one person's search behavior. *Information Processing and Management*, 34(2), 237-256.

UNESCO. (2017). *Building Knowledge Societies*. Retrieved March 20, 2017, from <http://en.unesco.org/themes/building-knowledge-societies>

Vakkari, P. (2001). A theory of the task-based information retrieval process: a summary and generalisation of a longitudinal study. *Journal of Documentation*, 57(1), 44-60.

Vakkari, P. & Hakala, N. (2000). Changes in relevance criteria and problem stages in task performance. *Journal of Documentation*, 56(5), 540-562.

Vakkari, P., Pennanen, M. & Serola, S. (2003). Changes of search terms and tactics while writing a research proposal: a longitudinal case study. *Information Processing and Management*, 39(3), 445-463.

Wang, P. & Soergel, D. (1998). A cognitive model of document use during a research project. Study 1. Document selection. *Journal of the American Society for Information Science*, 49(2), 115-133.

W3C Brasil (2011). Missão do W3C. <http://www.w3c.br/Sobre/MissaoW3C>

Yang, S. (1997). Information seeking as problem-solving using a qualitative approach to uncover the novice learners' information-seeking process in a perseus hypertext system. *Library and Information Science Research*, 19(1), 71-92.

Kuhlthau, C.C., Heinström, J. & Todd, R.J. (2008). "The 'information search process' revisited: is the model still useful?" *Information Research*, 13(4) paper 355. Retrieved March 20, 2017, from <http://InformationR.net/ir/13-4/paper355.html>



O ENSINO DA CLASSIFICAÇÃO ARQUIVÍSTICA NA FORMAÇÃO DO PROFISSIONAL DA INFORMAÇÃO: ANÁLISE QUALITATIVA DOS CURRÍCULOS DA UFF E UNESP

Juliana de Mesquita Pazos¹, Clarissa Moreira dos Santos Schmidt²

¹Universidade Federal Fluminense, 0000-0001-8771-1531, juliana_mesquita@id.uff.br

²Universidade Federal Fluminense, 0000-0003-1555-4594, clarissaschmidt@id.uff.br

RESUMO Uma efetiva organização do conhecimento arquivístico depreende, fundamentalmente, do aprofundamento teórico e empírico adquirido pelo profissional da informação no decurso de sua formação acadêmica. Nesse sentido, a classificação arquivística assume papel preponderante na teoria, metodologia e prática da Arquivística, porquanto está intimamente ligada à conformação científica da área e à formulação do seu escopo teórico-metodológico; e, desse modo, auxilia na condução adequada do trabalho de organização da informação e do conhecimento nos arquivos. Apesar de sua importância, a área apresenta problemáticas pouco investigadas que influem no modelo de ensino e aprendizagem do profissional da informação. Diante disso, nosso objetivo, neste trabalho, foi o de verificar se a classificação arquivística é objeto de uma disciplina específica na Universidade Federal Fluminense e na Universidade Estadual Paulista, a fim de analisar se seus currículos e planos de ensino expressam a relevância da operação para a singularidade do profissional da informação. Para tanto, foi empregada metodologia exploratória, qualitativa e analítica para evidenciar o papel da classificação arquivística para a epistemologia Arquivística; estabelecer as especificidades inerentes à função, teoria, terminologia, metodologia e instrumentos produzidos pela operação; explicar o processo de criação da graduação em Arquivologia no Brasil, na UFF e na UNESP, além de analisar seus currículos e planos de ensino, de modo a averiguar possíveis dissonâncias que afetem a formação profissional. A pesquisa demonstrou que ambas as universidades possuem disciplinas voltadas para o estudo exclusivo da classificação arquivística; projetos pedagógicos que determinam a confluência curricular entre as áreas de Arquivologia e Biblioteconomia sob a ótica da Ciência da Informação; e abordam os aspectos relevantes ao estudo da área.

PALAVRAS-CHAVE *Ensino da Classificação Arquivística, Formação do Profissional da Informação, Organização do Conhecimento Arquivístico, Universidade Estadual Paulista, Universidade Federal Fluminense.*

ABSTRACT An effective organization of archival knowledge derives, fundamentally, from the theoretical and empirical deepening acquired by the professional of the information in the course of its academic formation. In this sense, the classification of records assumes a preponderant role in the theory, methodology and practice of Archival Science, because it is closely linked to the scientific conformation of the area and to the formulation of its theoretical-methodological scope; and thus assists in the proper conduct of the work of organizing information and knowledge in archives. Despite its importance, the area presents little investigated issues that influence the teaching and learning model of the information professional. In view of this, our objective, at work, was to verify if the classification of records is object of a specific discipline at the Federal University of Fluminense and the State University of São Paulo, in order to analyze if its curriculum and teaching plans express the relevance of the operation to the singularity of the information

professional. For that, an exploratory, qualitative and analytical methodology was used to highlight the role of classification of records for Archival Science epistemology; Establish the inherent specificities of the function, theory, terminology, methodology and instruments produced by the operation; to explain the process of creation of the undergraduate degree in Archival Science in Brazil, UFF and UNESP, and analyze their curriculums and teaching plans, in order to investigate possible dissonances that affect professional training. The research demonstrated that both universities have disciplines focused on the exclusive study of classification of records; Pedagogical projects that determine the curricular convergence between the areas of Archival Science and Librarianship from the point of view of Information Science; And discuss aspects relevant to the study of the area.

KEYWORDS *Archival Classification Teaching, Archival Knowledge Organization, Fluminense Federal University, Information Professional Training, São Paulo State University.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

1. INTRODUÇÃO

A Arquivística, inserida no escopo da Ciência da Informação, partilha do intento de satisfazer às necessidades informacionais dos indivíduos. Nesse sentido, a classificação arquivística destaca-se por salvaguardar as informações contidas nos documentos, garantindo a transparência e o compartilhamento de informações, o respaldo de decisões organizacionais e o exercício da plena cidadania pelos usuários (Sousa, 2003). Recordamos, ainda, que a obra elementar da área (o “Manual de Arranjo e Descrição de Arquivos”, publicado em 1898) foi sistematizada, em grande parte, pela urgência de um método classificatório que atendessem às especificidades inerentes ao documento de arquivo; e que os princípios da Proveniência e da Ordem Original foram formulados pela mesma razão.

Assim, podemos afirmar que a classificação arquivística é uma operação que fundamenta e distingue todo o labor do profissional da informação em arquivos, pois assume papel preponderante na organização da informação e do conhecimento arquivístico.

Segundo Rousseau e Couture (1998), o surgimento dos Princípios da Proveniência e da Ordem Original ocorreu em reação à aplicação dos métodos de classificação bibliográficos aos acervos arquivísticos, que ocasionavam mutilações nos fundos de arquivo, prejudicando a representação das funções e atividades do organismo produtor. Silva, Ribeiro, Ramos e Real (2009) recordam que as classificações bibliográficas começaram a ser empregadas nos arquivos no final do século XVII, com o despontar do pensamento iluminista, na tentativa de organizar a documentação para satisfazer à demanda do público erudito, comprometendo de forma permanente a organicidade da documentação e seu potencial informacional.

Nesse contexto, Schmidt e Smit (2015) sublinham que as informações contidas nos documentos de arquivo são asseguradas através da classificação arquivística. Lopes (1996) e Sousa (2003) defendem que a classificação arquivística, ao exprimir o contexto orgânico-funcional do fundo, auxilia na conformação do patrimônio histórico-documental institucional/social e funciona como intermediária entre a informação contida no acervo arquivístico e a satisfação das necessidades informacionais dos usuários. E Gonçalves (1998) define a classificação arquivística como uma operação lógica que consiste em criar categorias referentes à estrutura, funções e atividades meio e fim da entidade produtora dos documentos, de modo a evidenciar as relações orgânicas existentes entre ela e seus documentos. O

produto final dessa operação se manifesta num Plano de Classificação/Quadro de Arranjo, no qual os documentos são ordenados hierarquicamente no interior de fundos, em séries e subséries, conforme sua composição estrutural e/ou funcional.

No entanto, Sousa (2006, p. 133) recorda que “o reconhecimento da importância da função classificação não garantiu um aprofundamento e um cuidado maior com a questão”, elencando questões não solucionadas de ordem terminológica e metodológica, tais como: a inexatidão quanto à aplicação de termos distintos para um mesmo processo classificatório (classificação e arranjo); a crença errônea na similaridade entre as atividades de classificação, ordenação e arquivamento; e a aplicação equivocada de métodos classificatórios estranhos às especificidades do documento de arquivo. Nessa linha de pensamento, Lopes (1997, p. 92) declara que “são raros, também, os textos que falam sobre, especificamente, classificação de documentos ou informações na fase ativa e semi-ativa. A bibliografia refere-se, nos grandes manuais e em outras obras, à classificação operada nos arquivos permanentes”.

Este cenário em que a classificação arquivística se apresenta como vital para a área, mas prossegue com problemáticas que são pouco investigadas, influi diretamente no modelo de ensino e aprendizagem dos estudantes dos cursos de Arquivologia.

Diante disso, nosso objetivo, neste trabalho, foi verificar se a classificação arquivística é lecionada como disciplina específica na Universidade Federal Fluminense (UFF) e na Universidade Estadual Paulista (UNESP), de modo a analisar se os planos de ensino demonstram a relevância da operação para a singularidade do profissional da informação em arquivos; se abrangem suas problemáticas teóricas e práticas; e se diferenciam a classificação arquivística da bibliográfica, considerando o histórico de uso das classificações temáticas nos arquivos. Para tanto, buscamos evidenciar, através de uma abordagem teórica e analítica, o papel da classificação para a cientificidade da Arquivística; explicar o processo de institucionalização universitária do curso de Arquivologia no Brasil, na UFF e na UNESP; e examinar os currículos dessas universidades e os planos de ensino das suas disciplinas, para verificar se as especificidades inerentes à função, teoria, terminologia, metodologia e instrumentos produzidos pela classificação arquivística são abordados no modelo disciplinar proposto.

2. METODOLOGIA

Considerando a destacada relevância da função classificação para o tratamento documental arquivístico, as problemáticas conceituais e operacionais que a cercam e a carência de pesquisas na área, julgamos pertinente o presente trabalho de pesquisa que tem como proposta problematizar os preceitos fundamentais que norteiam o processo de classificação em arquivos na raiz da formação acadêmica do profissional da informação em arquivos, mediante um estudo nos currículos e planos de ensino do curso de Arquivologia da UFF e UNESP. Justificamos esse recorte de pesquisa pelo fato de que estas universidades, localizadas no sudeste do país, operacionalizam os currículos de Arquivologia e Biblioteconomia sob a ótica interdisciplinar da Ciência da Informação e, ainda, possibilitam uma visão mais abrangente sobre o ensino da classificação arquivística no âmbito de uma universidade federal e de uma estadual.

Neste sentido, a metodologia caracterizou-se por uma abordagem exploratória, qualitativa e analítica. O método foi aplicado sob uma perspectiva dialética, isto é, a identificação dedutiva dos aspectos gerais que distinguem e orientam a classificação arquivística; e a investigação indutiva das particularidades

do modelo disciplinar nos currículos das graduações em Arquivologia da UFF e UNESP, entendidos como um produto social que reflete um tempo, espaço e contexto determinados.

A consolidação do referencial teórico decorreu de uma revisão bibliográfica a respeito do estudo da classificação arquivística, de modo a identificar e selecionar os elementos mais relevantes ao seu ensino. Para conformação da parte empírica, e satisfação dos objetivos a que essa pesquisa se propõe, realizamos pesquisa documental das legislações que regem os currículos, além do levantamento dos planos de ensino das disciplinas nos sítios eletrônicos da UFF e UNESP, e/ou por contato via e-mail com as respectivas coordenações de cursos.

3. A ARQUIVÍSTICA NAS UNIVERSIDADES BRASILEIRAS

A fundação do Arquivo Nacional brasileiro ocorreu em 1838 (intitulado Arquivo Público do Império) e foi determinante para o desenrolar do pensamento arquivístico no país. De acordo com Marques (2007), o Arquivo Nacional foi responsável pelas primeiras tentativas de organizar cursos de capacitação profissional para tratar dos arquivos, empenhando-se por estabelecer um curso de formação profissional em arquivos em suas dependências. O objetivo foi alcançado com a implantação definitiva do Curso Permanente de Arquivos (CPA), no ano de 1960, baseado numa tradição francesa de gestão e organização de arquivos.

Na década seguinte, foi fundada a Associação dos Arquivistas Brasileiros – AAB (1971-2015), que reuniu profissionais inclusive na defesa da institucionalização universitária do CPA e da regulamentação da profissão de arquivista. Também promoveu os primeiros congressos brasileiros de Arquivologia e elaborou o primeiro periódico científico da área, chamado “Arquivo & Administração” (1972-2015), impulsionando o desenvolvimento da Arquivística brasileira enquanto campo do Fazer e Saber. Em 1977, o CPA transformou-se no curso universitário de Arquivologia da Federação das Escolas Federais Isoladas do Estado do Rio de Janeiro (atual Universidade Federal do Estado do Rio de Janeiro – UNIRIO); e a profissão de arquivista foi regulamentada pela Lei n. 6.546 (1978) (Fonseca, 2005; Marques, 2007).

Nesse sentido, Schmidt (2012) defende que a AAB atuou como mediadora das necessidades dos trabalhadores da área, que ansiavam por uma formação mais completa, uma profissão reconhecida segundo seu mérito, e saberes pensados no *lócus* da universidade e da pesquisa científica. E Marques (2007) argumenta que a institucionalização acadêmica da Arquivística era uma necessidade premente, visto que o aumento do número de arquivos no país não refletia num cuidado maior com o Fazer e Saber arquivísticos.

As décadas seguintes foram marcadas pela consolidação do pensamento arquivístico no âmbito universitário, com a expansão do número de cursos de graduação em Arquivologia no país; qualificação do corpo docente dessas graduações com a entrada dos mesmos na pós-graduação; ampliação das áreas de concentração dos cursos de pós-graduação que permitiam a pesquisa em Arquivologia, com destaque na Ciência da Informação; maior produção e comunicação científicas em periódicos da área por professores universitários; e a saída de profissionais das instituições arquivísticas para a universidade (Fonseca, 2005).

Isto posto, consideramos de suma importância investigar como decorre o ensino da classificação arquivística nas universidades brasileiras e como seus conceitos – fundamentais e, por vezes,

problemáticos – são apreendidos pelo futuro arquivista. Essa opinião é corroborada por Britto (1999) e por Fonseca (2005), que defendem a necessidade de examinar constantemente a adequação das matrizes curriculares às necessidades profissionais e intelectuais do profissional da informação.

3.1 A UNIVERSIDADE FEDERAL FLUMINENSE (UFF)

A Universidade Federal Fluminense é uma instituição pública de ensino superior mantida pelo Governo Federal desde sua criação em 1960. O curso de Bacharel em Arquivologia foi implementado pela Resolução n. 73 (1978), sendo reconhecido pelo Ministério da Educação (MEC) em 1985, junto ao Departamento de Documentação, no Instituto de Artes e Comunicação Social (Niterói, Rio de Janeiro). Atualmente encontra-se vinculado ao Departamento de Ciência da Informação e compartilha um projeto pedagógico comum ao curso de Bacharel em Biblioteconomia e Documentação.

Desde a criação do curso de Arquivologia da UFF, vigoraram três currículos plenos, que sofreram ajustes curriculares pontuais que não serão abordados nesse trabalho, salvo aquele responsável pela introdução da disciplina Classificação Arquivística no terceiro currículo universitário.

A primeira matriz curricular foi instituída pela Resolução n. 53 (1978), sendo composta por 35 disciplinas obrigatórias de núcleos básico e profissional (2.985 horas), 5 disciplinas optativas (300 horas), 2 estágios supervisionados (270 horas) e as disciplinas de Prática desportiva (60 horas) e Estudo de problemas brasileiros [Educação moral e cívica] (60 horas). Os prazos mínimo, médio e máximo para a conclusão do curso eram, respectivamente, 6, 8 e 10 períodos.

O segundo currículo foi introduzido por meio da Resolução n. 84 (1993), de modo a articular graduação e extensão. O currículo básico e profissional passou a ser composto por 37 disciplinas obrigatórias (2.445 horas), 2 disciplinas optativas (120 horas), 2 disciplinas eletivas (120 horas), 1 estágio supervisionado (225 horas) e 1 trabalho de conclusão de curso (120 horas). A duração mínima, média e máxima para integralização mudou para 7, 8 e 16 períodos.

O terceiro currículo, incorporado pela Resolução n. 297 (2006), resulta dos debates que se iniciaram no ano de 2001 a respeito da congruência entre os projetos pedagógicos dos cursos de Arquivologia e Biblioteconomia e Documentação dentro do Departamento de Ciência da Informação. As reflexões foram estimuladas pela necessidade de adequação à política pedagógica da universidade, pautada numa perspectiva interdisciplinar e/ou transdisciplinar; e pela articulação do eixo comum disciplinar entre as áreas de Arquivologia, Biblioteconomia e Museologia realizada pelo MEC. O currículo foi, então, organizado em 3 núcleos: o específico, com 10 disciplinas (600 horas); o geral, com 22 disciplinas (1.350 horas) comuns ao curso de Biblioteconomia e Documentação, incluindo o trabalho de conclusão de curso (120 horas); e o complementar, com 5 disciplinas obrigatórias (350 horas), 2 disciplinas optativas (120 horas), 1 estágio supervisionado (120 horas) e as atividades acadêmicas (120 horas). A duração mínima, média e máxima para a conclusão do curso foi mantida em 7, 8 e 16 períodos.

Ao analisarmos esses três currículos plenos, encontramos em cada um deles alguma disciplina que se refere à classificação (não necessariamente arquivística) em sua nomenclatura, a saber: Introdução à Classificação (1978-1993), Fundamentos da Classificação (1993-2006) e Classificação em Arquivos (2016-). Nos planos de ensino das disciplinas Introdução à Classificação e Fundamentos da Classificação não foram verificadas preocupações com a classificação arquivística, visto que os conteúdos eram voltados puramente para as classificações filosóficas e bibliográficas.

Em relação à classificação arquivística, apesar da sua comprovada importância, inexistiu qualquer disciplina exclusivamente voltada ao seu estudo desde a criação da graduação até o ano de 2016, quando a disciplina Classificação em Arquivos tornou-se obrigatória, após dois períodos letivos como optativa. Cabe ressaltar que, no decorrer desse período, a classificação arquivística foi tratada, ainda que de maneira breve, como matéria das seguintes disciplinas: Gestão de Documentos II (1993-) e Linguagens Documentárias Notacionais (2007-), que dedicam uma unidade do plano de ensino à classificação arquivística; e Representação da Informação (2007-), que trata paralelamente dos preceitos das classificações arquivística e bibliográfica, com ênfase na recuperação da informação.

Assim, a disciplina Classificação em Arquivos (60 horas), incluída no 5º período acadêmico do terceiro currículo pleno através da Resolução n. 175 (2016), volta-se exclusivamente para o ensino da classificação em arquivos. A disciplina exige como pré-requisitos as disciplinas Fundamentos Arquivísticos I e II e Gestão de Documentos I e II, num encadeamento que tem a seguir as disciplinas de Avaliação e Seleção de Documentos, Arquivos Permanentes e Descrição Arquivística, de modo que as principais funções envolvidas no tratamento e organização de arquivos estejam representadas na matriz curricular.

A avaliação do plano de ensino demonstra que os aspectos considerados relevantes para a literatura especializada no ensino da teoria, metodologia e operacionalização da classificação em arquivos são realizados a contento, tais como: a verificação histórica dos métodos de organização e classificação de documentos arquivísticos; a metodologia classificatória baseada nos aspectos estruturais e/ou funcionais do organismo produtor, e não na sua recuperação temática; o debate acerca dos conceitos de classificação e arranjo, seja como procedimentos distintos ou como um mesmo procedimento referido por dois termos; e a operacionalização da classificação através de um instrumento (plano de classificação/quadro de arranjo) e de métodos adjacentes para ordenação e arquivamento.

Quadro 1. Plano de Ensino da Disciplina Classificação em Arquivos (2016-)

Ementa	A classificação na teoria e na prática Arquivística. Teorias e metodologias para classificação em arquivos. Ordenação e arquivamento. Métodos estrutural e funcional. Elaboração de instrumentos de classificação.
Objetivos	Abordar a teoria da classificação em Arquivos. Entender os fatores e processos envolvidos na Classificação em Arquivos. Conhecer os métodos para Classificação e Arranjo. Possibilitar o reconhecimento e classificação de principais tipos documentais, bem como a elaboração de instrumentos de classificação.
Conteúdo Programático	A função Classificação e os documentos de arquivo: a Classificação na trajetória histórico-epistemológica da Arquivologia; teorias da Classificação em Arquivos; Classificação e Arranjo. Teorias e metodologias para classificação em arquivos: ordenação e arquivamento; elementos da Classificação e tipos de Classificação; elaboração e análise de instrumentos de classificação.

Fonte: elaborado pelas autoras.

3.2 A UNIVERSIDADE ESTADUAL PAULISTA (UNESP)

A Universidade Estadual Paulista “Júlio de Mesquita Filho” é uma universidade pública brasileira mantida pelo Governo do Estado de São Paulo desde sua criação em 1976. O curso de Bacharel em Arquivologia foi implementado na Faculdade de Filosofia e Ciências (Marília, São Paulo) através da Resolução n. 26 (2003), sendo reconhecido pelo MEC em 2014. Atualmente está vinculado ao departamento de Ciência da Informação, assim como o bacharelado em Biblioteconomia.

A primeira matriz curricular do curso foi instituída pela Resolução n. 87 (2003), que determinava a composição do currículo pleno por 41 disciplinas obrigatórias de núcleos geral e específico (2.460 horas), 2 disciplinas optativas (60 horas), 1 estágio supervisionado (252 horas) e 1 trabalho de conclusão de curso (sem atribuição de horas). O prazo mínimo e máximo para a conclusão do curso eram, respectivamente, de 4 e 7 anos.

Esse currículo já evidenciava uma preocupação com a função classificação arquivística, haja vista o estabelecimento da disciplina Classificação e Arranjo Documental (2003-2012), no 5º período letivo, composta por aulas teóricas (30 horas) e práticas (30 horas).

A matriz curricular sofreu alterações significativas com a Resolução n. 157 (2012). Foi estabelecido um sistema de pré-requisitos entre as disciplinas e o currículo pleno passou a ter 44 disciplinas obrigatórias de núcleos geral e específico (2.190 horas), 4 disciplinas optativas (240 horas), 1 estágio curricular (210 horas), 1 trabalho de conclusão de curso (120 horas) e atividades complementares (90 horas). O estágio curricular ficou atrelado ao 5º período em diante, sob a seguinte divisão operacional: arquivos correntes e intermediários (60 horas), classificação arquivística (30 horas), descrição (30 horas), preservação (30 horas) e arquivos permanentes (60 horas). A feitura do trabalho de conclusão de curso ficou vinculada às linhas de pesquisa desenvolvidas pelo corpo docente na pós-graduação, no intento de aproximar a graduação em Arquivologia e a pós-graduação em Ciência da Informação. Manteve-se a duração do curso em 4 (mínimo) e 7 (máximo) anos.

Nesse currículo consta a disciplina Classificação Arquivística, composta por aulas teóricas (30 horas) e teórico-práticas (30 horas), estando vinculada ao 4º período letivo, sendo lecionada paralelamente às disciplinas Produção Documental, Arquivos Correntes e Intermediários e Planejamento e Gestão de Unidade Arquivísticas, evidenciando a necessidade da classificação arquivística desde a gênese do documento de arquivo. Não existem pré-requisitos necessários ao curso.

Verifica-se que as disciplinas Classificação e Arranjo Documental (2003-2012) e Classificação Arquivística (2012-) compartilham a mesma ementa, ainda que os objetivos e o conteúdo programático tenham sido aprofundados, o que nos permite inferir que se trata de uma mesma disciplina que teve sua nomenclatura modificada com a reestruturação curricular. Todavia, essa alteração é significativa, pois demonstra o posicionamento do curso quanto à questão terminológica envolvida na classificação arquivística, ou seja, nos parece que ao excluir o termo “arranjo documental” existe a compreensão da classificação e do arranjo como um mesmo processo classificatório no âmbito dos arquivos.

Além disso, a análise dos planos de ensino dessas disciplinas demonstram que: (i) a teoria, metodologia, terminologia e operacionalização da classificação arquivística é realizada sob a defesa da não importação de métodos estranhos às especificidades do documento de arquivo; (ii) existe a preocupação em inserir a função classificação arquivística dentro da política de gestão documental, com vistas ao controle eficiente e eficaz dos documentos produzidos por empresas públicas e/ou privadas; (iii) o Princípio da Proveniência é destacado como fundamento da função ao orientar os métodos de feitura do instrumento de classificação a partir da estrutura e/ou funções da entidade produtora; (iv) a diferença entre a composição do fundo de arquivo e da coleção documental é abordada, possibilitando compreender que a classificação arquivística (manutenção do contexto de produção dos documentos e da sua organicidade) possui objetivos e usos distintos da classificação bibliográfica (disposição física nas estantes para posterior localização e recuperação de documentos diversos); (v) a dicotomia terminológica e operacional existente entre os procedimentos de classificação e arranjo são discutidas pelas diferenças e similaridades dos instrumentos produzidos; (vi) a ordenação interna das séries

documentais dentro do fundo é debatida a partir dos critérios que a orientam, da sua operacionalização e dos instrumentos de trabalho produzidos para controle e acesso; e (vii) o procedimento de arquivamento não é mencionado, mas supomos que seja abordado quando trabalhados os temas de ordenação e codificação no plano de classificação.

Quadro 2. Planos de Ensino das Disciplinas Classificação e Arranjo Documental e Classificação Arquivística

Disciplinas	Classificação e Arranjo Documental (2003-2012)	Classificação Arquivística (2012-)
Ementa	Classificação, ordenamento, controle e destinação dos documentos como elementos para uma política de gestão documental.	Classificação, ordenação, controle e destinação dos documentos como elementos para uma política de gestão documental.
Objetivos	Avaliar e aplicar instrumentos de classificação e arranjo para a elaboração de programas e políticas voltadas para a eficiência e economia da administração pública e privada.	Construir planos de classificação e quadros de arranjo de fundos arquivísticos. Aplicar a classificação na elaboração de programas e políticas voltadas para a eficiência e economia da administração pública e privada.
Conteúdo Programático	Introdução à classificação: princípios de procedência e respeito aos fundos. O fundo de arquivo e sua estrutura interna: as séries documentais. Sistemas de classificação do fundo de arquivo. Ordenação de séries documentais: critérios, operação, unidade, instrumentos de trabalho, controle e acesso.	Introdução à classificação: princípios de procedência, respeito aos fundos e organicidade. Identificação do fundo e sua diferença de coleção. O fundo de arquivo e sua estrutura interna: as séries documentais. Sistemas de classificação do fundo de arquivo: métodos estruturais e funcionais. Plano de classificação e arranjo documental: diferenças e semelhanças. Elaboração e codificação do plano de classificação. Ordenação de séries documentais: critérios, operação, unidade, instrumentos de trabalho, controle e acesso. A classificação arquivística como instrumento da gestão documental: o plano de classificação e a construção da tabela de temporalidade.

Fonte: elaborado pelas autoras.

4. CONCLUSÕES

A Arquivologia consolidou-se como um campo científico autônomo nos fins do século XIX, a partir da publicação de um manual de arquivistas holandeses sobre os preceitos teórico-metodológicos que fundamentam o trabalho de classificação e descrição dos documentos produzidos e acumulados organicamente nos arquivos. Desde então, a área preocupou-se em estabelecer seu *corpus* teórico e prático sob a égide do Princípio de Proveniência, buscando manter a integridade intelectual e física dos documentos no interior de um fundo documental.

Nesse contexto, a Classificação Arquivística assume um papel preponderante na epistemologia da Arquivística e na teoria, metodologia e prática da área em todas as fases do seu ciclo documental. Portanto, em todo o labor profissional do arquivista. Também recebe influências de áreas afins e possui dilemas não solucionados de ordem teórica, metodológica, terminológica e operacional.

A análise dos três currículos plenos que vigoraram no curso de graduação em Arquivologia da Universidade Federal Fluminense revelou que, apesar da comprovada relevância da classificação arquivística para a teoria e a prática arquivísticas, inexistiu qualquer disciplina obrigatória voltada

exclusivamente para o seu estudo desde a implantação da graduação em 1978 até o ano de 2016, quando a disciplina Classificação em Arquivos tornou-se obrigatória. Averiguamos que, no decorrer desse período, a classificação arquivística foi abordada, ainda que de forma pontual e não exclusiva, pelas disciplinas Gestão de Documentos II (1993-), Representação da Informação (2007-) e Linguagens Documentárias Notacionais (2007-).

Por sua vez, a análise dos dois currículos plenos praticados no curso de Arquivologia da Universidade Estadual Paulista Júlio de Mesquita Filho demonstra uma preocupação em distinguir disciplinas voltadas para o ensino exclusivo da classificação arquivística: Classificação e Arranjo Documental (2003-2012) e Classificação Arquivística (2012-).

Assim, esta pesquisa demonstrou que ambas as universidades possuem, atualmente, disciplinas voltadas para o estudo exclusivo da classificação arquivística, além de projetos pedagógicos que determinam a confluência curricular entre as áreas de Arquivologia e Biblioteconomia em departamentos de Ciência da Informação. A análise dos planos de ensino das disciplinas de Classificação em Arquivos (UFF) e Classificação Arquivística (UNESP) demonstra que ambas abordam os conteúdos apontados pela literatura especializada como relevantes ao ensino da classificação arquivística, isto é, a verificação histórica dos métodos de classificação de documentos arquivísticos e sua relação com o nascimento científico da Arquivística; a abordagem teórica e metodológica da classificação arquivística obedecendo aos princípios da Proveniência e da Ordem Original, tomando como base os aspectos estruturais e funcionais da entidade produtora e negando a importação de metodologias estranhas às especificidades dos documentos arquivísticos; a discussão teórica, metodológica e terminológica que envolve os termos classificação, arranjo, ordenação e arquivamento; e a operacionalização técnica do plano de classificação/quadro de arranjo através de métodos adjacentes de ordenação e arquivamento.

Diante do exposto, consideramos que a formação do profissional da informação, na atual Era do Conhecimento, deve abranger conhecimentos teóricos e técnicos que possibilitem uma constante reflexão acerca das teorias que orientam as práticas nos distintos ambientes produtores de informação em relação com o papel social desse profissional na organização e divulgação das informações sob sua guarda. Nas instituições ou unidades arquivísticas, distinguimos a classificação arquivística como a operação que exprime o contexto de produção dos documentos, mantendo a organicidade e a autenticidade dos documentos de arquivo e, portanto, suas informações preservadas nos ambientes analógico e digital. A análise do modelo formativo quanto à classificação arquivística empregado nas universidades analisadas, demonstrou que para a UFF essa foi uma preocupação secundária até o ano de 2016, quando foi instituída uma disciplina voltada para o ensino exclusivo da classificação em arquivos, ainda que o projeto pedagógico do curso determinasse desde 2007 a integração dos currículos de Arquivologia e Biblioteconomia, sob a orientação do Departamento de Ciência da Informação, para uma ampla formação do profissional da informação. Já na UNESP, essa foi uma questão pensada desde a criação do curso, que também possui um projeto pedagógico de tronco-comum disciplinar.

Por fim, acreditamos que o presente trabalho de pesquisa atingiu aos objetivos propostos, pois foi procedida a análise do ensino da função classificação arquivística nas matrizes curriculares interdisciplinares da UFF e da UNESP sob um viés histórico das próprias universidades federal e estadual; do reconhecimento da graduação alinhada ao projeto pedagógico da Biblioteconomia; das reestruturações curriculares; e dos planos de ensino das disciplinas. Por isso, acreditamos que esse estudo pode vir a contribuir positivamente para trabalhos futuros acerca tanto do ensino universitário da função classificação arquivística quanto da institucionalização universitária da Arquivística no Brasil.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- Associação dos Arquivistas Holandeses. (1973). *Manual de arranjo e descrição de arquivos* (M. A. Wanderley, Trad.). (2a ed.). Rio de Janeiro: Arquivo Nacional. (Obra original publicada em 1898).
- Britto, M. T. N. (1999). Cartografia do ensino universitário da Arquivologia nas Américas. In J. M. Jardim, & M. O. Fonseca (Orgs.), *A formação do arquivista no Brasil* (Cap. 1, pp. 9-30). Niterói: EdUFF.
- Fonseca, M. O. (2005). *Arquivologia e ciência da informação*. Rio de Janeiro: Fundação Getúlio Vargas.
- Gonçalves, J. (1998). *Como classificar e ordenar documentos de arquivo*. São Paulo: Arquivo do Estado.
- Lei n. 6.546, de 04 de julho de 1978*. Dispõe sobre a regulamentação das profissões de Arquivista e de Técnico de Arquivo, e dá outras providências. Recuperado de http://www.planalto.gov.br/ccivil_03/leis/1970-1979/L6546.htm
- Lopes, L. C. (1996). *A informação e os arquivos: teorias e práticas* (1a ed.). Niterói, São Carlos: EdUFF, EdUFSCAR.
- Lopes, L. C. (1997). *A gestão da informação: as organizações, os arquivos e a informática aplicada*. Rio de Janeiro: Arquivo Público do Estado do Rio de Janeiro.
- Marques, A. A. C. (2007). *Os espaços e os diálogos da formação e configuração da Arquivística como disciplina no Brasil* (Dissertação de mestrado). Universidade de Brasília, Brasília, Distrito Federal, Brasil.
- Resolução UNESP n. 26, de 22 de maio de 2003*. Dispõe sobre a criação do Curso de Arquivologia junto à Faculdade de Filosofia e Ciências do Campus de Marília. Recuperado de <https://sistemas.unesp.br/legislacao-web/>
- Resolução UNESP n. 87, de 25 de julho de 2003*. Estabelece a estrutura curricular do Curso de Arquivologia junto a Faculdade de Filosofia e Ciências do Campus de Marília. Recuperado de <https://sistemas.unesp.br/legislacao-web/>
- Resolução UNESP n. 157, de 4 de dezembro de 2012*. Estabelece a estrutura curricular do Curso de Arquivologia da Faculdade de Filosofia e Ciências do câmpus de Marília. Recuperado de <https://sistemas.unesp.br/legislacao-web/>
- Resolução UFF n. 53, de 31 de agosto de 1978*. Estabelece o Currículo do Curso de Graduação em Arquivologia.
- Resolução UFF n. 73, de 28 de junho de 1978*. Dispõe sobre a Criação do Curso em Arquivologia.
- Resolução UFF n. 84, de 05 de maio de 1993*. Estabelece o Currículo do Curso de Graduação em Arquivologia.
- Resolução UFF n. 175, de 04 de maio de 2016*. Estabelece o Ajuste Curricular do Curso de Graduação em Arquivologia, Grau: Bacharelado, aprovado pela Resolução n.º 297/2006 e alterada pelas

Resoluções n.º 165/2008, 011/2010, 326/2010 e 32/2015 deste Conselho. Recuperado de <http://www.noticias.uff.br/bs/bs.php>

Resolução UFF n. 297, de 29 de novembro de 2006. Estabelece o Currículo do Curso de Graduação em Arquivologia. Recuperado de <http://www.noticias.uff.br/bs/bs.php>

Rousseau, J., & Couture, C. (1998). *Os fundamentos da disciplina arquivística*. Lisboa: Publicações Dom Quixote.

Schmidt, C. M. S., & Smit, J. (2015, novembro). Organização e representação da informação em arquivos: uma análise a partir da função classificação. *Anais do Congresso ISKO Espanha-Portugal*, Múrcia, Espanha, 2.

Silva, A. M., Ribeiro, F., Ramos, J. & Real, M. L. (2009). *Arquivística: teoria e prática de uma Ciência da Informação* (3a ed.). Porto: Edições Afrontamento.

Sousa, R. T. B. (2003). Os princípios arquivísticos e o conceito de classificação. In G. M. Rodrigues, & I. Leite (Orgs.), *Organização e representação do conhecimento na perspectiva da ciência da informação* (pp. 240-269). Brasília: CID/Thesaurus.

Sousa, R. T. B. (2006). Classificação de documentos arquivísticos: trajetória de um conceito. *Arquivística.net*, 2 (2), 120-142.

Universidade Estadual Paulista Júlio de Mesquita Filho (2003). *Projeto Pedagógico do Curso de Arquivologia*. Marília: Unesp.

Universidade Estadual Paulista Júlio de Mesquita Filho (2012). *Projeto Pedagógico do Curso de Arquivologia*. Marília (SP): Unesp.



AÇÕES PARA O PROCESSO DE ENSINO APRENDIZAGEM EM DISCIPLINAS DE ORGANIZAÇÃO E TRATAMENTO DA INFORMAÇÃO: INOVAÇÃO PEDAGÓGICA DO GRUPO DE PESQUISA RECRI/UFMG

Elisângela C. Aganette¹, Benildes Coura M. S. Maculan², Célia da Consolação Dias³,
Gercina Ângela Lima⁴, Cintia de Azevedo Lourenço⁵

¹*Grupo de Pesquisa Representação de Conhecimento e Recuperação da Informação (RECRI) da Escola de
Ciência da Informação (ECI) da Universidade Federal de Minas Gerais (UFMG), Brasil,
elisangelaaganette@gmail.com*

²*RECRI, ECI, UFMG, Brasil, 0000-0003-4303-9071, benildes@gmail.com*

^{3,4,5}*RECRI, ECI, UFMG, Brasil, ppggoc@eci.ufmg.br*

RESUMO O presente artigo apresenta os resultados alcançados pelo Grupo de Pesquisa em Representação de Conhecimento e Recuperação da Informação (RECRI) da Escola de Ciência da Informação (ECI) da Universidade Federal de Minas Gerais (UFMG), referentes a um projeto que contou com o suporte de dois alunos, um de mestrado e um de doutorado, oriundos de programas de pós-graduação da ECI. Teve como objetivo a reformulação da apresentação dos conteúdos e a produção de material didático para o ensino de disciplinas de Organização e Tratamento da Informação e do Conhecimento, oferecidas aos alunos de graduação do curso de Biblioteconomia da ECI/UFMG. Mais especificamente, este artigo descreve o trabalho inicial da remodelagem dos conteúdos de duas disciplinas: Sistemas de Classificação - CDU e Sistemas de Classificação - CDD. Este trabalho evidenciou a necessidade e importância do desenvolvimento de projetos pedagógicos teóricos que possam refletir e criar inovação nas práticas de ensino-aprendizagem.

PALAVRAS-CHAVE *Formação Docente; Produção de Material Didático; Redesenho Educacional. Organização da Informação e do Conhecimento.*

ABSTRACT The present article presents the results achieved by the Research Group on Knowledge Representation and Information Retrieval, RECRI of the School of Information Science (ECI) of the Federal University of Minas Gerais (UFMG), referring to a project that was supported by Two students, one master's and one doctoral, coming from ECI's graduate programs. It aimed to reformulate the presentation of content and the production of didactic material for the teaching of Information and Knowledge Organization and Treatment courses offered to undergraduate students of the ECI / UFMG Librarianship course. More specifically, this work began to reshape the contents of two disciplines: Classification Systems - CDU and Classification Systems - CDD. This work evidenced the need and importance of the development of practical and theoretical projects that can reflect and create pedagogical innovation for teaching and learning in the field of OTIC disciplines.

KEYWORDS *Teacher Training; Production of Didactic Material; Educational Redesign. Organization of Information and Knowledge.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

O mundo interconectado, globalizado e dinâmico não pode mais se distanciar das tecnologias de informação e comunicação (TICs) e os profissionais que trabalham no ensino superior têm feito uso dos recursos dessas tecnologias para aperfeiçoar o processo de construção do conhecimento dos alunos. No Brasil, nota-se uma tendência em se valorizar as iniciativas sobre investigações de inovação em práticas pedagógicas no ensino superior, a partir das ações e metas estipuladas pelo Ministério da Educação (MEC), um órgão governamental que busca transformar o modelo vigente. A prática universitária não pode mais se basear em uma relação rígida, caracterizada pela memorização do conteúdo, mas, sim, em uma efetiva possibilidade de produção do conhecimento.

Na contemporaneidade, considera-se que o ensino superior oferece oportunidades para a construção do conhecimento de forma crítica e sistematizada, sendo necessário criar condições inovadoras e favoráveis que permitam que professores e alunos possam refletir sobre a teoria e a prática do ensino-aprendizagem. A formação inicial e a formação continuada de professores demonstram o comprometimento das instituições de ensino com o processo educativo, evidenciando a postura crítica e o papel social do docente. Nesse sentido, nota-se a importância de fomentar a qualidade e a inovação do processo de ensino-aprendizagem nos cursos de ensino superior, introduzindo novas metodologias pedagógicas, com uso de meios tecnológicos e com o aprimoramento de material didático.

Pautado nesse desafio, o Grupo de Pesquisa em Representação de Conhecimento e Recuperação da Informação, RECRI, da Escola de Ciência da Informação da Universidade Federal de Minas Gerais (UFMG), no Brasil, desenvolveu um projeto que está ligado ao Programa de Incentivo à Formação Docente (PIFD), oferecido pela Pró-Reitoria de Graduação da UFMG. Esse trabalho teve o suporte de dois alunos, oriundos dos programas da ECI/UFMG, Pós-Graduação em Gestão & Organização do Conhecimento - PPG-GOC e Pós-Graduação em Ciência da Informação - PPGCI, para possibilitar o contato desses alunos com as questões pedagógicas da docência no ensino superior. O objetivo norteador do projeto é a reformulação da apresentação dos conteúdos e a produção de material didático para o ensino no campo da Organização e Tratamento da Informação e do Conhecimento do curso de Biblioteconomia da ECI/UFMG. Mais especificamente, este artigo apresenta os principais resultados alcançados com o trabalho inicial de remodelagem dos conteúdos para duas disciplinas: Sistemas de Classificação - CDD e Sistemas de Classificação - CDU.

O artigo está organizado como segue: na introdução apresenta-se a visão geral sobre o ensino superior e o projeto RECRI; em seguida, descreve a técnica que é aplicada para o ensino-atividade; a seção objetos de aprendizagem caracteriza e aponta os objetivos de diferentes recursos digitais; a seção sistemas de classificação bibliográficos apresenta as características deste tipo de instrumento; a seção metodologia descreve os procedimentos que foram aplicados na concretização do projeto de ensino; a seção de resultados expõe a implementação dos novos recursos digitais de aprendizagem em cada disciplina e termina com as conclusões e ações futuras.

FORMAÇÃO EM SITUAÇÃO

O método de capacitação Formação em Situação, segundo Navarro (2015), é uma técnica que segue a abordagem de Jean-Marie Barbier para a educação de adultos. A técnica é pautada no atendimento à dimensão multidimensional de formação do indivíduo, quando se busca contemplar as perspectivas

cognitiva, política e social do aluno, atrelando-as aos novos conhecimentos técnico-científicos que se quer ensinar, saberes estes que são inerentes às práticas da profissão.

Em entrevista concedida a Boas (2012, p. 261), Barbier declara que seus estudos iniciaram nos anos de 1980, utilizando-se das “situações de trabalho, das novas tecnologias ou das pesquisas em formação”, para “desenvolver uma tipologia das culturas educativas e, em particular, uma distinção entre culturas de ensino, culturas de formação e culturas de desenvolvimento de competências”. Para Barbier, na Formação em Situação, os “conceitos-chave são todos aqueles que giram em torno das relações entre formação e atividade. Por exemplo, a noção de objetivo, de capacidade, de negociação, de alternância” (Boas, 2012, p. 265), que pode ser apresentada da seguinte maneira:

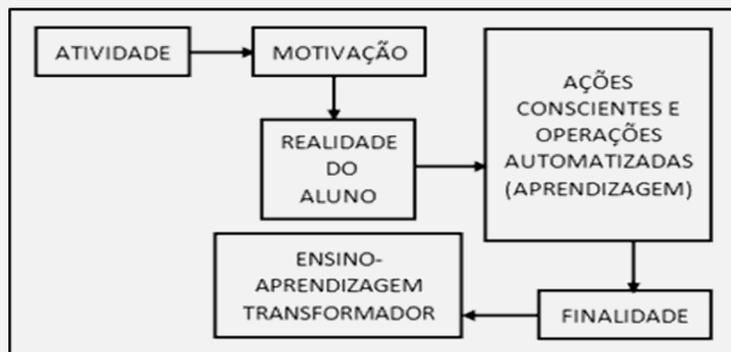


Figura 1 – Formação em Situação

Fonte: as autoras (2017).

Analisando a Figura 1, tem-se a Atividade a ser ensinada, que envolve a Motivação do aluno para a sua realização, que só pode ser alcançada pela competência individual desse aluno em assumir iniciativas (Realidade do Aluno). A partir disso, são determinadas as situações de trabalho, que abarcam as Ações Conscientes, que darão condições ao aluno de compreender, conceitualmente, a atividade, e a aplicação condicionada de seus procedimentos (Operações Automatizadas), que evidenciam a Finalidade de tal atividade e promovem um Ensino-aprendizagem Transformador. Nessa perspectiva, o professor é um mediador e deve levar seus alunos a construir e elaborar a ideia do objeto (atividade) estudado.

A partir de seus estudos, Barbier percebeu a necessidade de se utilizar uma Teoria de Atividade (TA) para ser aplicada no campo da formação e das pesquisas sobre as práticas profissionais (BOAS, 2012). Daniels (2003) afirma que a TA teve origem com os trabalhos de Vigotski, Leontiev e Luria e que foi desenvolvida com base no conceito de mediação. Para o autor, a TA vem sendo utilizada para a compreensão do indivíduo (consciência individual e coletiva) em situações práticas de realização de atividades em diversos sistemas sociais. Segundo Daniels (2003), Vygotsky afirma que a internalização de uma atividade é um processo no qual há uma mediação da linguagem, quando os signos adquirem significado e sentido para o sujeito (agente). Nessa perspectiva, o aluno é o agente que irá fazer uso dos materiais ou conceitos (elementos de mediação) para atingir um resultado, que é o de aprender a realizar uma atividade e compreender a sua finalidade.

Para integrar as inovações pedagógicas nas disciplinas modeladas, foi necessária a reformulação dos conteúdos, criando objetos digitais de aprendizagem que permitissem a interação, colaboração e socialização de informações e experiência entre grupos, para disponibilizá-los em um ambiente virtual.

OBJETOS DIGITAIS DE APRENDIZAGEM

Os objetos digitais de aprendizagem (ODAs), ou “recursos digitais de aprendizagem”, são utilizados por educadores para facilitar a prática pedagógica, utilizando meios tecnológicos de aprimoramento de material didático. Esse termo foi adotado pelo *Learning Technology Standards Committee (LTSC)*, do *Institute of Electrical and Electronics Engineers - IEEE*, referenciando objetos de conteúdos multimídia, *softwares* educacionais, pessoas, organizações, livros, entre outros. Esses recursos permitem que as atividades acadêmicas se tornem mais dinâmicas, despertando o interesse dos alunos por meio de conteúdos criativos e planejados.

Existem diversas definições sobre o que são ODAs na literatura, (Polsani, 2003; Audino e Nascimento, 2010; Churchill, 2007; Filatro; 2008; entre outros), mas ainda não existe uma definição universalmente aceita, provavelmente por ser uma área relativamente nova. A definição mais recorrente é da norma 1484, *Standard for Learning Object Metadata, de 2002 (LTSC/IEEE)*, que descreve os ODAs como uma entidade que pode ser usada, reutilizada ou referenciada durante o ensino com suporte tecnológico. Uma das definições mais referenciadas (Willey, 2002) considera os ODAs como quaisquer recursos digitais que possam ser reusados para apoiar a aprendizagem.

A ideia principal dos ODAs é fragmentar o conteúdo educacional em unidades modulares independentes para serem reutilizados em diferentes ambientes de aprendizagem, em um determinado contexto. Além disso, eles podem ser utilizados individualmente ou combinados com outros para complementar alguma instrução que requer um maior suporte de ensino-aprendizagem. De acordo com Gazzoni (2006), a elaboração do ODA deve levar em conta todos os procedimentos pedagógicos, desde a escolha do conteúdo a ser construído, as estratégias mais adequadas para fazê-lo, até a compreensão do processo de ensino-aprendizagem e das interações entre os alunos.

Singh (2001) sugere que um ODA deve ser estruturado e dividido em três partes: (1) objetivos pedagógicos: que norteiam o uso do objeto, e uma lista dos conhecimentos prévios necessários para aproveitamento do conteúdo; (2) conteúdo instrucional: material didático necessário para que o aluno possa atingir os objetivos propostos; (3) prática e *feedback*: permite ao aluno utilizar o material e receber retorno sobre seu desempenho.

Mendes, Souza e Caragnato (2004), apresentam as características e elementos da estrutura e da operacionalidade dos ODAs: i) reusabilidade: o objeto poderá ser reutilizável em múltiplos contextos para múltiplos propósitos; ii) adaptabilidade: adaptáveis a qualquer ambiente de ensino e aprendizagem; iii) granularidade: “tamanho” de um objeto, podendo ser um conteúdo recortado, como uma imagem, texto ou fragmento de áudio. Um ODA de menor granularidade pode ser uma página *web* inteira, que combina textos, imagens e vídeos; iv) acessibilidade: identificação padronizada que garante a sua recuperação e acesso fácil, para ser usado em diversos locais; v) durabilidade: permanecer intactos independentes das atualizações de *software* ou *hardware*; vi) interoperabilidade: habilidade de operar através de uma variedade de *hardware*, sistemas operacionais e *browsers*, com intercâmbio entre diferentes plataformas e sistemas; vii) metadados: descrever as informações de um objeto, permitindo que sejam facilmente encontrados em um sistema de busca.

Os exemplos mais citados na literatura como objetos digitais de aprendizagem incluem vídeos, jogos, animações, vídeo aulas, infográficos, mapas, imagens, gráficos, sons, ferramentas, conteúdos de aplicações multimídia, conteúdos instrucionais, recursos didáticos na forma de arquivos digitais,

referências a *sites*, emulações interativas, visualização de objetivos de aprendizagem, organizações ou eventos ou outros materiais que possam ser utilizados como ferramenta pedagógica dentro e fora da sala de aula.

SISTEMAS DE CLASSIFICAÇÃO BIBLIOGRÁFICOS: CDD E CDU

Os Sistemas de Classificação Bibliográfica (SCB) têm como principal função organizar acervos de bibliotecas e unidades de informação segundo os assuntos de que tratam os itens do acervo. Esses sistemas se fundamentam em três conceitos: categorias, divisão lógica e relacionamento. As classificações bibliográficas têm hoje uma utilidade indiscutível quando se trata de localizar fisicamente o documento em sua classe de assunto.

Na perspectiva da Biblioteconomia, os SCBs são instrumentos verbais-simbólicos, construídos a partir do mapeamento de conceitos que são organizados conforme critérios e características pré-estabelecidos. Em geral, é constituído por três elementos: (1) esquema de classificação, agrupando elementos, de maneira sistemática, por suas semelhanças e diferenças; (2) notação da classificação (numérico ou alfanumérico); (3) índice alfabético, que facilita a busca do usuário. O processo de classificar é o ato de analisar um item documental para extrair dele os assuntos que o representam e, depois, “dividir [esses conceitos] em grupos ou classes, segundo as diferenças e semelhanças. É dispor os conceitos, segundo suas semelhanças e diferenças, em certo número de grupos metodicamente distribuídos” (PIEIDADE, 1983, p.9), podendo ser organizados segundo um ponto de vista ou atributos classificatórios que atendam às necessidades do usuário. Nessa atividade, é utilizado um SCB para que se possa converter os assuntos do conteúdo do documento em notações classificatórias.

Neste artigo, os SCBs que interessam são os instrumentos Classificação Decimal de Dewey (CDD) e Classificação Decimal Universal (CDU), cujas características gerais são as seguintes:

- 1) Classificação Decimal de Dewey (CDD): foi concebida por Melvil Dewey, em 1873, e publicado pela primeira vez em 1876, tendo por base o sistema de Willians T. Harris para a classificação do conhecimento humano, os princípios da classificação de Bacon (1608, “*The advancement of learning*”), e o sistema decimal de Àmpere, com subdivisões de caráter hierárquico, com notações puras. O sistema CDD é composto por quatro volumes e mais utilizado na classificação de acervos mais gerais, que englobam temas de diferentes áreas do conhecimento;
- 2) Classificação Decimal Universal (CDU): foi desenvolvida por Paul Otlet e La Fontaine a partir de 1895, sendo a primeira edição publicada entre 1904-1907 e tendo por base a CDD, e, assim, é também um sistema de classificação bibliográfica geral. A principal característica do sistema CDU é permitir a síntese (combinação) de conceitos, sendo considerado um instrumento semi analítico-sintético. É mais utilizado na classificação de acervos mais especializados, pois permite a combinação de conceitos (ex.: Bibliotecas Universitárias Especializadas) (SOUZA, 2012). O sistema CDU possui notações mistas e é composto por dois volumes.

Os objetos de aprendizagem apresentados neste artigo foram desenvolvidos para as disciplinas referentes ao uso desses dois instrumentos, que são apresentadas na seção de resultados.

METODOLOGIA

Realizou-se um levantamento sobre os conteúdos da grade curricular do curso de Biblioteconomia da ECI/UFMG buscando verificar os fundamentos teórico-metodológicos no campo de Organização e Tratamento da Informação. A análise desses conteúdos evidenciou que as disciplinas oferecidas nesse campo abarcam dois eixos básicos: (1) Bases teórico-conceituais: respaldam o processo de construção do conhecimento e de socialização da informação (tal como Estudos de Usuário); (2) Bases teórico-procedimentais: suporte ao desenvolvimento de métodos de representação e a criação de instrumentos (linguagens de indexação e padrões de descrição).

Seguindo essas perspectivas, o grupo RECRI tem realizado discussões no sentido de promover ações para fomentar o ensino-aprendizagem dos alunos do curso em questão, buscando criar estratégias para promover uma aproximação entre a teoria e a prática no ensino, na pesquisa e na extensão. Dentre essas estratégias, criou-se o laboratório de pesquisa de práticas inovadoras para a construção de material didático interativo, introduzindo novas metodologias pedagógicas, assistidas por tecnologia de ensino à distância.

O projeto contou com o apoio da Diretoria de Inovação e Metodologias de Ensino – GIZ, uma rede de desenvolvimento de práticas de ensino superior da UFMG, criada em 2008. O GIZ apoia projetos de pesquisas para o desenvolvimento de metodologias pedagógicas e tecnologias educacionais inovativas, de forma colaborativa e personalizada, para diferentes áreas do conhecimento. As etapas a seguir explicitam a forma de definição, execução e registro do referido projeto:

- 1) Elaboração do planejamento do projeto, com definição de cronograma e dos papéis e responsabilidades de acordo com o escopo do projeto.
- 2) Estabelecimento de parceria com o GIZ.
- 3) Qualificação dos bolsistas, a partir da participação em minicursos ofertados pelo GIZ:
 - a) Uso do *Moodle*: trata-se de um ambiente virtual de aprendizagem a distância *Moodle*, amplamente utilizado na UFMG e que permitiu que se colocasse em rede um conjunto de informações interligadas;
 - b) Criação de objetos de aprendizagem.
- 4) Liberação dos ambientes no *Moodle* tanto para reunir a documentação do projeto, bem como para organizar os ambientes de teste com os recursos utilizados em cada disciplina.
- 5) Composição dos protótipos de redesenho educacional:
 - c) Descrição do protótipo, que consiste na listagem e descrição dos Objetos de Aprendizagem previstos para composição do protótipo:
 - Vídeos de apresentação das disciplinas;
 - Aulas introdutórias das disciplinas com power point dinâmico e recurso de gameificação;
 - Mapas Mentais de apresentação das disciplinas, elaborado no *MindMeister* e disponibilizado no *Prezi*;
 - Exercícios em formato de Quiz;
 - Estante Virtual de livros para exercícios de classificação;
 - Tutorial de uso e apresentação do Moodle.

- d) Modelagem do Recurso *Moodle*: criação de canais tais como: i) chat para estimular discussão de conceitos e estabelecer vínculos entre os participantes; ii) fóruns para a disponibilização de portfólio, repositório de atividades e relatório de atividades de campo, assim como para disponibilizar as atividades práticas. Para isso foi necessário:
- Modelar das páginas das disciplinas em módulos que correspondem aos conteúdos de cada disciplina;
- 
- Definir os fóruns sobre frequência e notas, configurado para contato particular do aluno junto ao professor;
 - Definir os fóruns sobre dúvidas referentes à Disciplina;
 - Realizar agendamento online para monitoria da disciplina.
- 6) Elaboração do protótipo das duas disciplinas: CDD e CDU.
- 7) Registro de elaboração do protótipo, por meio do detalhamento das impressões vivenciadas pelos agentes envolvidos no projeto em ambiente Wiki.

RESULTADOS ALCANÇADOS E DISCUSSÕES

A equipe do GIZ contribuiu com sugestões quanto ao uso de recursos apropriados ao redesenho educacional das disciplinas no projeto, com o objetivo de: i) potencializar as aulas em *PowerPoint* inserindo recursos de áudio, temporizador e execução de atividades, pelos alunos; ii) desenvolver material em *podcast*; iii) desenvolver material de áudio e vídeo pré-modelado que possibilite um reaproveitamento em outros semestres; iv) configurar as disciplinas integralmente na plataforma *Moodle* com opção de migração de conteúdo nos semestres seguintes.

O GIZ assessorou as duas alunas bolsistas, juntamente com as professoras do RECRI, responsáveis pelo projeto, por meio do método de capacitação denominado Formação em Situação, descrito neste artigo, que permitiu maior eficiência no processo de construção dos recursos. Ademais, o GIZ fez a seleção do material efetivamente utilizado no protótipo e disponibilizado no ambiente de teste do *Moodle*.

Como resultados deste projeto destacam-se os produtos: i) criação de miniprotótipos: redesenho das disciplinas de CDD e CDU, com os objetos de aprendizagem construídos e disponibilizados no ambiente *Moodle* e passível de replicação a cada semestre; ii) capacitação das bolsistas por meio das oficinas do GIZ e da participação em eventos da área; iii) repasses de conhecimento – RCs: disponibilização das apresentações correspondentes para que as professoras pudessem replicar entre os pares.iv) reforço para os membros do RECRI da necessidade de continuidade do projeto.

MODELAGEM DA DISCIPLINA CDD

A Classificação Decimal de Dewey - CDD, trata-se de um sistema de classificação e ordenação lógica do conhecimento por assuntos para facilitar a recuperação da informação. É utilizada também para ordenar as entradas de um catálogo sistemático e de uma bibliografia (PIEDADE, 1983).

A disciplina “Sistemas de Classificação Bibliográficos: CDD” é oferecida aos alunos do quinto período do curso de Biblioteconomia da ECI/UFMG. Tem como objetivos: i) introduzir o aluno no estudo e uso das classificações bibliográficas; ii) introduzir os conceitos de classificação e de organização da informação; iii) capacitar para a identificação da estrutura e função dos principais sistemas de classificação bibliográfica, formas de utilização e pertinência de uso, segundo objetivos institucionais e características dos documentos e finalmente iv) abordar o uso dos princípios da classificação para a organização de informações em novos contextos. Nesse contexto, busca-se apresentar os fundamentos dos sistemas de classificação bibliográfica, sua estrutura e aplicações nos processos de organização do conhecimento, tomando-se como objetos de estudo um dos sistemas de classificação decimal, a CDD. Os conteúdos são apresentados usando estratégias de aulas teóricas e práticas, por meio de aulas expositivas, leituras, exercícios e avaliações.

Inicialmente, conforme Figura 2, foram compiladas informações básicas aos alunos, uma vídeoaula e um mapa conceitual da disciplina.

Figura 2 – Apresentação da disciplina CDD

Fonte: as autoras (2017)

Para alcançar os objetivos propostos para a disciplina, o ensino foi dividido em três módulos, a saber: i) Módulo I: Classificação bibliográfica (classificação e recuperação da informação; histórico; teorias de classificação bibliográfica; sistemas de classificação bibliográfica: tipos; elementos; ordenação, tabelas, notação, índice; como classificar: plano das ideias, plano verbal e plano notacional); ii) Módulo II: Classificação Decimal de Dewey (histórico e estrutura; classes básicas; tabelas auxiliares; índice; síntese de notação: com as tabelas auxiliares; entre as classes principais; Módulo III: Prática de Classificação Decimal de Dewey (exercícios; laboratório), como pode ser observado na Figura 3:

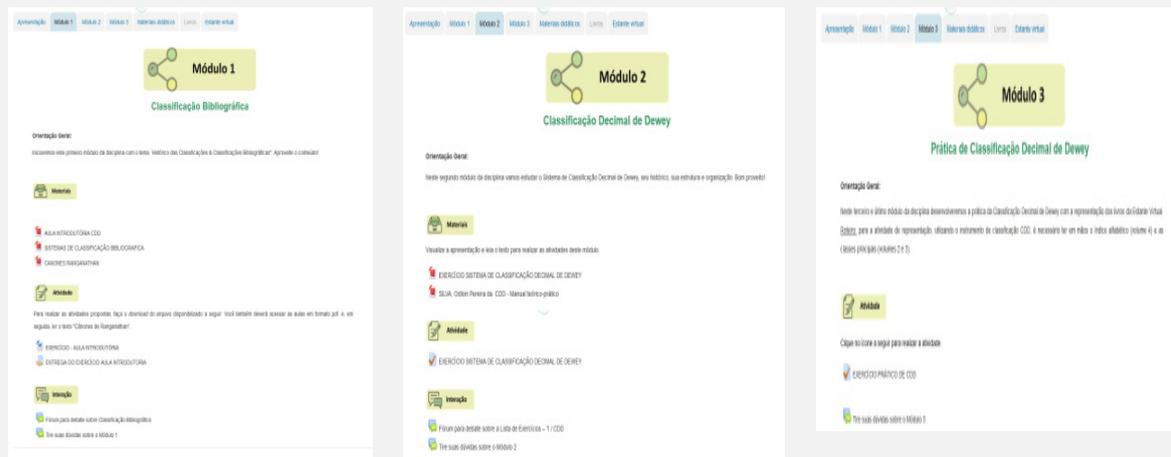


Figura 3 – Módulos da disciplina CDD

Fonte: as autoras (2017)

Além disso, são disponibilizados os Materiais didáticos utilizados na disciplina como: apresentações de aula, referências bibliográficas e apostilas. Por último, há uma Estante virtual para ser utilizado pelos alunos nos exercícios de classificação bibliográfica.

Desta forma, para cada um dos conteúdos dos quatro módulos foram criados objetos de aprendizado que visam o ensino-aprendizagem do sistema CDD.

MODELAGEM DA DISCIPLINA CDU

A disciplina “Sistemas de Classificação Bibliográficos: CDU” é oferecida aos alunos do sexto período do curso de Biblioteconomia da ECI/UFMG. Tem como objetivo o ensino-aprendizagem do sistema CDU na classificação de assuntos de acervos de bibliotecas, como instrumento nos serviços de organização e recuperação da informação contida em recursos documentais. Nesse sentido, tem como escopo fazer conhecer os fundamentos teóricos da classificação (bases teóricas, histórico e evolução) e da CDU (fundamentos, características e funções), assim como compreender o uso e aplicação das suas tabelas. Os conteúdos são apresentados com aulas expositivas, leituras, exercícios e avaliações.

Seguindo essa organização, conforme Figura 4, foram compiladas informações básicas aos alunos, uma vídeo-aula e um mapa conceitual da disciplina.

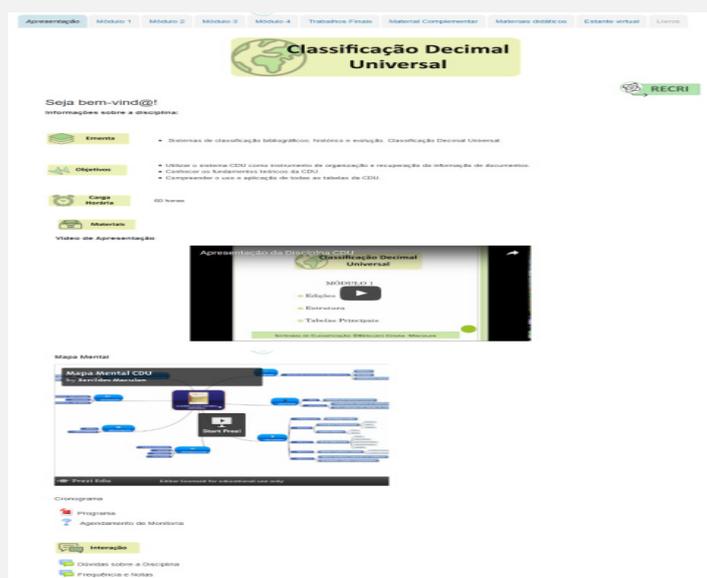


Figura 4 – Apresentação da disciplina CDU

Fonte: as autoras (2017)

Para alcançar os objetivos da disciplina, o conteúdo foi dividido em quatro módulos: (1) Módulo I: introdução, com os fundamentos conceituais, histórico, características e apresentação das dez Tabelas Principais (sendo que a tabela quatro está vazia, aguardando por assuntos novos); (2) Módulo II: apresentação das cinco Tabelas Auxiliares Comuns Relacionais; (3) Módulo III: apresentação das cinco Tabelas Auxiliares Comuns Independentes e das seis Tabelas Auxiliares Comuns Dependentes; (4) Módulo IV: apresentação das Tabelas Auxiliares Analíticas Especiais e dos princípios de ordenação, com a ordem de citação (opcional) e a ordem de arquivamento (obrigatória), conforme Figura 5:

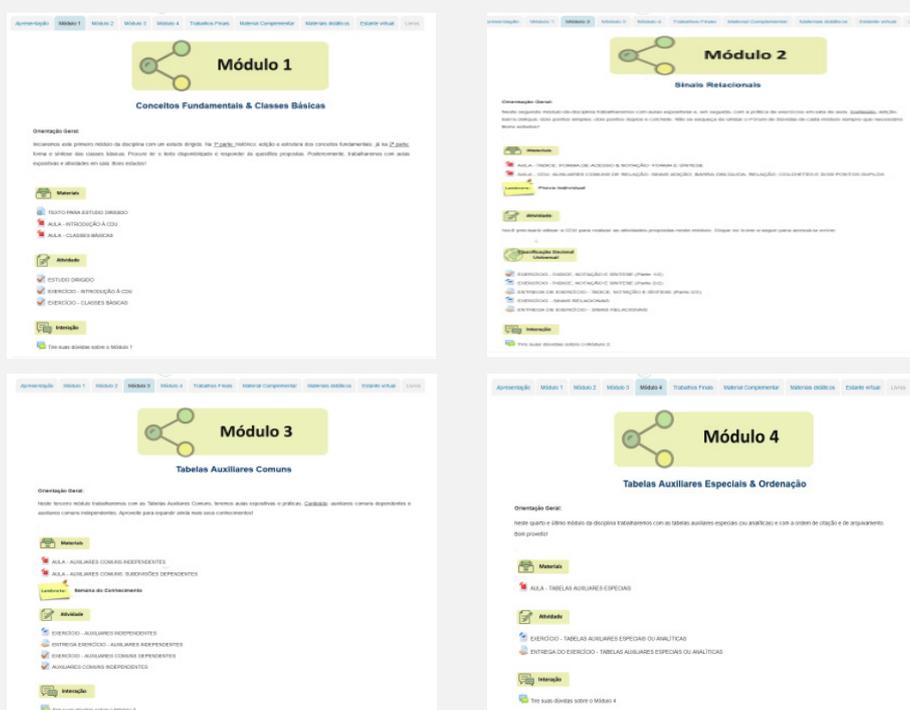


Figura 5 – Módulos da disciplina CDU

Fonte: as autoras (2017)

Além disso, são disponibilizados os Trabalhos finais, Materiais complementares, Materiais didáticos e uma Estante virtual para os exercícios de classificação bibliográfica. Para cada módulo foram criados objetos de aprendizado para o ensino-aprendizagem do sistema CDU.

Em termos gerais, a solução completa do Desenho Educacional das Disciplinas CDD e CDU, contempladas no referido projeto, foi estruturada da seguinte forma:

- Organização do conteúdo didático no *Moodle*: ambiente mais amigável e um *layout* com mais usabilidade;
- Desenvolvimento de objetos de aprendizagem: Mapas Mentais das disciplinas refletindo o conteúdo do programa, os módulos e etapas previstos; potencializar as aulas disponibilizadas em *PowerPoint* com os recursos multimídia: *Podcast* de tópicos específicos; Leitura do conteúdo e inclusão do áudio no próprio .ppt para auxílio dos alunos com necessidades especiais; Vídeo complementar à aula; Recurso/link com dúvidas frequentes, módulo e/ou atividade;
- Criação de Estante Virtual: estante virtual com um acervo, com *e-books* de domínio público em formato .pdf para uso em atividades práticas de classificação com a CDD e da CDU.
- Criação um ambiente ativo no ambiente das disciplinas para tirar dúvidas relativas à frequência, provas, marcação de horário de monitoras, por meio de fóruns (conjunto de fóruns pré-definidos);
- Criação de mecanismos para melhorar a gestão da disciplina;
- Elaboração de exercícios interativos, possibilitando acesso às respostas dos colegas e/ou correção automática;
- Criação um ambiente de perguntas frequentes, com respostas realizadas pelos professores por meio de vídeos;
- Criação de agenda/calendário para que o aluno reserve e acompanhe os horários de agendamento das atividades de monitoria e de atividades paralelas.

CONCLUSÕES

Consideramos que é importante e necessário desenvolver projetos práticos e teóricos que possam refletir sobre o ensino-aprendizagem no campo das disciplinas em Organização e Tratamento da Informação, visando criar condições favoráveis ao aprendizado dos alunos de graduação em Biblioteconomia. Uma vez que a sociedade está inserida na era da informação e do conhecimento, o ensino das disciplinas que compõem a área de organização e tratamento da informação e do conhecimento deve incluir conteúdos para a representação em diferentes contextos. Além disso, deve-se desenvolver mecanismos que possam facilitar a compreensão de tais conteúdos, como também possibilitar aos alunos uma reflexão teórica e prática acerca das possíveis melhorias das atividades técnicas, do uso de tecnologias que ajudam a otimizar a aprendizagem, atuando também como atores neste processo.

Outro aspecto é a validação, por parte dos alunos, dos recursos implementados. Ressaltamos que está programada a continuidade do projeto para este ano de 2017 com o redesenho de outras disciplinas. A ideia é que todos os recursos sejam usados e avaliados do ponto de vista da sua adequação ao foco do projeto e que seja feito o registro da necessidade de ajustes para uma implementação posterior

Em relação aos resultados alcançados neste projeto, o RECRI destaca a importância de que todas as ações dos docentes responsáveis pelas disciplinas permitam, aos futuros profissionais de Biblioteconomia, ter acesso à teoria e à prática, para refletir, de forma crítica, sobre os problemas que se manifestam entre o fazer intelectual e o fazer técnico. E, para finalizar destacamos a importância deste projeto como uma ação complementar para otimizar a formação dos alunos do curso de biblioteconomia, bem como uma oportunidade de trabalhar como um laboratório de prática que oportunize a reflexão tanto dos docentes quanto dos discentes nas teorias e práticas relacionadas ao conteúdo das disciplinas da área de Organização e Tratamento da informação.

AGRADECIMENTOS

Agradecemos à Pró-Reitoria de Graduação da UFMG, pelo apoio financeiro, à equipe do GIZ e aos alunos bolsistas pelo suporte no projeto, assim como à Fundação de Amparo à Pesquisa de Minas Gerais (Fapemig), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

REFERÊNCIAS BIBLIOGRÁFICAS

- Araújo, N. M. S. (2013). A avaliação de objetos de aprendizagem para o ensino de língua portuguesa: Análise de aspectos tecnológicos ou didático-pedagógicos? In J. Araújo & N. M. S. Araújo (Eds.), *EaD em tela: docência, ensino e ferramentas digitais* (pp. 179-207). Campinas: Pontes.
- Audino, D. F., & Nascimento, R. S. (2010). Objetos de aprendizagem: Diálogos entre conceitos e uma nova proposição aplicada à educação. *Revista Contemporânea de Educação*, 5, 128-148.
- Churchill, D. (2007). Towards a useful classification of learning objects. *Education Technology Research and Development*, 55(5), 479-497.
- Daniels, H. (2003). *Vygotsky e a pedagogia*. São Paulo: Loyola.
- Dodebei, V. (2002). Funções das linguagens documentárias. In V. Dodebei, *Tesouro: linguagem de representação da memória documentária* (pp. 24-25). Niterói: Intertexto.
- Filatro, A. (2008). *Design instrucional: na prática*. São Paulo: Pearson Education do Brasil.
- Gazzoni, A., Canal, A. P., Falkembach, G. A. M., Fioreze, L. A., Pincolini, L. B., & Antoniazzi, R. (2006). Proporcionalidade e semelhança: Aprendizagem na via objetos de aprendizagem: *RENOTE: Revista Novas Tecnologias da Educação*, 4(2), 1-9.
- IEEE Learning Technology Standards Committee*. (2000). Fonte: IEEE, Institute of Electrical and Electronics Engineers: <http://www.ieeeltsc.org>
- IEEE Standard for Learning Object Metadata* 1484.12.1. (6 de Sept. de 2002). Fonte: IEEE Computer Society. Acesso em 9 de Jun. de 2017, disponível em <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1032843&isnumber=22180>

Mendes, R. M., Souza, V. I., & Caregnato, S. E. (2004). A propriedade intelectual na elaboração de objetos de aprendizagem. *Anais do V Cinform, Encontro Nacional de Ciência da Informação*, Salvador, BA, Brasil, UFBA, 1-7. Disponível em <http://www.lume.ufrgs.br/bitstream/handle/10183/548/000502901.pdf?sequence=1>

Navarro, L. (2015). *Formação em situação de trabalho para profissionais da saúde mental: A arte como mediação*. Campinas: PUC-Campinas.

Piedade, M. (1983). *Introdução à teoria da classificação*. Rio de Janeiro, RJ: Interciência.

Singh, H. (2001). *Introduction to learning objects*. Paper apresentado em Julho de 2001 no Washington eLearning Forum. Acesso em 9 jun. 2017, disponível em <http://www.elearningforum.com/meetings/2001/july/Singh.pdf>

Souza, S. (2009). *CDU: como entender e utilizar a segunda edição-padrão internacional em língua portuguesa*. Brasília: Thesaurus.

Souza, S. (2012). *CDU: como entender e utilizar a segunda edição-padrão internacional em língua portuguesa (3. ed. rev. atual.)*. Brasília: Thesaurus.

South, J. B., & Monson, D. W. (2000). A university-wide system for creating, capturing, and delivering learning objects The instructional use of learning objects. In D. A. Wiley (Ed.), *The instructional use of learning objects*. Bloomington: AECT. Disponível em <http://reusability.org/read/>

UDC Consortium. (20 de maio de 2017). Universal Decimal Classification Consortium. Disponível em: <http://www.udcc.org/>

Wiley, D. A. (2002). Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. In D. A. Wiley, *The instructional use of learning objects*. Bloomington: AECT. Disponível em <http://reusability.org/read/>



ATUAÇÃO PROFISSIONAL EM AMBIENTES DIGITAIS: PERSPETIVAS PARA O DESENVOLVIMENTO DA COMPETÊNCIA EM INFORMAÇÃO

Gabriela Belmont de Farias¹, Maria de Fátima Oliveira Costa²

¹Programa de Pós-graduação em Ciência da Informação, Universidade Federal do Ceará,
gabibfarias@gmail.com

²Programa de Pós-graduação em Ciência da Informação, Universidade Federal do Ceará,
fatima12oliveiracosta@gmail.com

RESUMO Aborda sobre o desenvolvimento de competências em informação como condição indispensável à formação profissional de qualidade, proporcionando adequada atuação em ambientes digitais. O objetivo da pesquisa foi identificar, por meio da análise do plano de ensino da disciplina *Seminário de Atuação Profissional* e por um roteiro de perguntas aplicado ao docente que ministra a disciplina e discentes matriculados, face às suas habilidades informacionais, no intento de capacitá-los à pertinente atuação profissional, ao eger os principais indicadores que sinalizarão aos alunos do oitavo semestre do curso de Biblioteconomia que julgam ter domínio para atuarem em ambientes digitais, além de analisar e descrever formas de desenvolver e tais indicadores. Configura-se a pesquisa como bibliográfica, documental, com a finalidade de identificar pontos em comum entre os temas: competência em informação, atuação bibliotecária e ambientes digitais. A análise dos dados compreende os seguintes componentes: ementa, objetivos e conteúdo do plano de ensino da disciplina *Seminário de Atuação Profissional* da unidade curricular de Recursos e Serviços de Informação; e as informações coletadas no questionário respondido pelos alunos e pela docente. Concluímos que os alunos participantes da pesquisa possuem habilidades básicas para atuação em ambientes digitais, sendo, entretanto, necessário desenvolver melhor tanto as ambientações quanto, também, as habilidades.

PALAVRAS-CHAVE *Atuação Profissional, Competências em Informação, Ambientes digitais. Mercado de Trabalho.*

ABSTRACT It addresses the development of information skills as an indispensable condition for quality professional training, providing an adequate professional performance in digital environments. The objective of the research was to identify, through the analysis of the teaching plan of the Seminar of Professional Performance and a script of questions applied to the teacher who teaches discipline and students enrolled with their informational skills in order to enable them to perform Professional students choosing the main indicators that will signal to the students of the eighth semester of the Librarianship course that they consider to have mastery to work in digital environments, besides analyzing and describing ways to develop these indicators. The research is set up as bibliographical and documentary in order to identify points in common between the themes: information competence, librarian performance and digital environments. The analysis of the data includes the following elements: contents, objectives and content of the teaching plan of the subject Professional Seminar of the curricular unit of Resources and Information Services; And the information collected in the questionnaire answered by the students and by the teacher. We conclude that the students participating in the research have basic skills to perform in digital environments, however, it is necessary to develop better both the environments and the skills.

KEYWORDS *Professional Performance, Skills in Information, Digital Environments. Job market.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

As universidades públicas em seu tripé - ensino, pesquisa e extensão, mediante as atividades oferecidas aos alunos de graduação e da pós-graduação, cumprem seu papel de promover a cooperação e o intercâmbio de ideias e conhecimentos com outros setores da sociedade. Para tais objetivos possam ser atingidos, é preciso que o ensino seja significativo e de qualidade, e esteja focado no desenvolvimento de competências. Para tanto, se torna indispensável que, no processo de ensino-aprendizagem, o docente e discentes sejam estimulados ao desenvolvimento da Competência em Informação (CoInfo), de modo que as habilidades desenvolvidas durante a formação acadêmica se tornem significativas e adequadas à sociedade.

No entendimento de Belluzzo (2013), a CoInfo é condição *sine qua non* ao desenvolvimento e inovação social de um país. No Brasil, o movimento pela consolidação da CoInfo está sendo realizada mediante o desenvolvimento de atividades com propósito de informar e formar profissionais conscientes da aplicabilidade da CoInfo em suas diversas atividades, sejam elas, em ambientes analógicos ou digitais.

Para que haja sustentabilidade no desenvolvimento de habilidades, no que diz respeito ao acesso e uso da informação, é necessário compreender três requisitos fundamentais: competência em informação para a cidadania - reporta-se ao uso crítico de dados e informação; competência em informação para o crescimento econômico – refere-se ao uso criativo e intensivo do conhecimento e à combinação eficiente dos serviços de informação; e competência em informação para a empregabilidade, relacionada ao desenvolvimento contínuo da pessoa com estratégias necessárias ao acesso e ao êxito econômico. (BELLUZZO, 2013).

Acerca da atuação profissional segundo Valentim (2000, p. 136) :

O profissional da informação precisa, antes de tudo, perceber qual realidade está vivenciando, primeiramente entender o ambiente em que atua, num segundo momento criar mecanismos eficientes de atuação na sociedade e, finalizando, enfrentar as mudanças cada vez maiores, antecipando-se às necessidades futuras da sociedade.

Nesse propósito, percebe-se a necessidade de debater sobre a prática e a atuação profissional, a serem percebidas por intermédio das mudanças que acontecem diuturnamente, e que carecem ser acompanhadas pelo profissional.

Lembrando o pensamento de (Vázquez, 1990, p. 157), quando elucida que:

[...] a prática não fala por si mesma. Como todos e quaisquer fatos, fatos práticos carecem de análise e interpretação racional, pois o critério da verdade está na prática, mas só se descobre numa relação propriamente teórica com a prática mesma.

Isso implica em reforçar alguns aspectos de natureza social da profissão e do mercado de trabalho. Entre os principais atributos do bibliotecário, a mediação se constitui quase uma identidade da profissão. De

fato, as ações e gestões por ele empreendidas têm por escopo auxiliar a busca e obtenção da informação e, de algum modo, a absorção do conhecimento. (COSTA, 2016. p.73)

A complexidade de estabelecer uma conexão da filosofia da CoInfo nos ambientes educacionais é desafiadora, pois a compreensão dessa filosofia, muitas vezes, é limitada, prejudicando o desenvolvimento de programas educacionais ou de capacitação.

Apresentaremos os resultados parciais da pesquisa cujo projeto intitulou-se ‘Competência em Informação e Criatividade: ações pedagógicas na unidade curricular de recursos e serviços de informação’. A finalidade foi identificar por meio do plano de ensino a disciplina Seminário de Atuação Profissional, docentes e discentes juntamente com suas habilidades informacionais suscetíveis para serem desenvolvidas na formação bibliotecária. Os objetivos da pesquisa são: a) Identificar os indicadores adotados pela *Information Literacy Competency Standards of Higher Education - Association of College & Research Libraries* (ACRL, 2000), os quais os alunos do oitavo semestre do curso de Biblioteconomia julgam ter domínio para atuarem em ambientes digitais; e b) Analisar e descrever formas de desenvolver os indicadores não sinalizados no objetivo anterior.

TEORIAS DE APRENDIZAGEM APLICÁVEIS AO ENSINO-APRENDIZAGEM EM CURSOS DE BIBLIOTECONOMIA

A educação contemporânea, inserida em ambiente de ensino considerado tradicional, é levada a ser repensada pela ineficácia que tem sua dinâmica em uma sociedade infoglobalizada. Observamos no contexto social a necessidade de o docente adentrar o mundo cotidiano do aluno e vivenciar suas aptidões culturais e, depois, refletir sobre seu ato de ensinar, construir um novo modelo de aprendizado baseado também no emocional. Facilitando, assim, a atenção na disciplina a ser ministrada em uma época marcada por transição, incertezas e indeterminações, resultantes de mudanças nas estruturas de convivência social, que alteram de forma significativa a cultura do local, acompanhada de novos componentes sociais e tecnológicos que exigem a sua utilização, de maneira que os usuários possam contribuir com conhecimentos que, em sua maioria, fogem do conteúdo probatório ou do rigor científico, causando uma explosão de informações que não podem ser comprovadas, por terem opiniões instantâneas e interativas sem fundamentos teóricos.

De acordo com Farias (2016), pode-se observar que o sistema educacional do saber pronto e imerso nas mudanças sociais, políticas, tecnológicas, econômicas e culturais influenciam nova dinâmica para o ensino superior. Ainda, segundo a autora, as características do docente, somadas às dos alunos, influenciam o desenvolvimento do ensino e da aprendizagem.

Em relação às características dessas influências, destacam-se alguns aspectos: dominação de uma cultura; aumento da velocidade das mudanças tecnológicas, sociais e culturais. Então, a multiplicidade do conhecimento implicará na revisão daquilo que, na realidade, as pessoas deverão aprender, quais sejam: a customização do conteúdo da informação; as aulas agruparão estudantes com interesses e capacidades distintas; potencializar-se-á o ensino à distância; fomentar-se-ão os trabalhos criativos, dentre outros fatores decorrentes dessa sociedade complexa. Vale lembrar, ainda, que o conhecimento não equivale à informação, pois ele se relaciona com a compreensão e o significado que se proporciona à informação (ONTORIA PEÑA; GÓMEZ R.; RUBIO MOLINA, 2004).

A sensação de perplexidade produzida por alunos universitários ao serem estimulados a desenvolver atividades criativas e inovadoras evidencia um problema educacional instalado em todos os níveis de escolaridade - ausência de familiaridade com o processo criativo, somado à necessidade do desenvolvimento da Competência em Informação, ou seja, o acesso e uso da informação de forma inteligente de modo a constituir o conhecimento e a sua aplicação à realidade social; esta última é área emergente no contexto brasileiro, porém, consolidada nos países desenvolvidos, apesar desses fatores incorporarem atribuições prioritárias e fundamentais para o bom desenvolvimento cognitivo. Parece que, não só as universidades, mas, também todo o sistema educacional resiste em sintonizar-se com as demandas atuais da sociedade, mantendo os mesmos e arcaicos métodos e técnicas de ensino consideradas ultrapassadas, “[...] sem o menor indício da nova cultura da aprendizagem” (ONTORIA PEÑA; GÓMEZ R.; RUBIO MOLINA, 2004, p.17).

É preciso refletir sobre como transformar o ensino para adaptar o aprendizado às demandas da sociedade, ou seja, é imprescindível pensar em variadas configurações para promover o aprendizado, utilizando estratégias pedagógicas adequadas à realidade do aluno e do que almeja a sociedade em relação, especialmente, ao que diz respeito ao bibliotecário. A configuração da aprendizagem deve estar situada em algo centrado no entendimento da informação básica que facilita o aperfeiçoamento do conhecimento, assim como o descobrimento dos caminhos que permitem compreender a informação necessária em outros momentos da vida. Outra característica é a de potencializar a capacidade de aprender e de pensar criticamente, facilitando a reflexão ante as variadas fontes de informação, o que poderá ajudar a reconhecer a validade e a precisão dos dados proporcionados. Desse modo, o enfoque do aprendizado é sintonizado com o desenvolvimento da Competência em Informação.

A flexibilidade no ensino-aprendizagem contribui para a potencialização do aprendizado cooperativo, cuja estrutura é mais horizontal e se afina melhor com a autonomia no aprendizado. Na aprendizagem, o aluno torna a base do aprendizado que consiste de envolvimento, compromisso e responsabilidade por parte dele, em ambiente gratificante e agradável, junto ao envolvimento responsável desses alunos, facilitando o aprender e incide na melhoria dos resultados. O aprender a aprender se torna uma habilidade essencial a esse “[...] modelo orientado ao processo, as pessoas e as organizações será o novo campo de possibilidades no que diz respeito à criatividade e ao desenvolvimento humano, e à exploração da interdependência de uns e outros” (HUNT, 1997, p.27).

O estímulo ao pensamento criativo e ao desenvolvimento da Competência em Informação é o grande desafio para o ensino superior, porque o desenvolvimento de pessoas competentes no acesso e uso da informação para criar e inovar é função essencial das escolas de Biblioteconomia, na medida em que visam ao preparo de bibliotecários como cidadãos capazes de realizar transformações de impacto na sociedade. A Competência em Informação e a criatividade adquirem uma função preponderante numa sociedade em permanente transformação, tornando-se uma das chaves do avanço do ensino superior e, conseqüentemente, da Biblioteconomia. É a mais valia que reforça a qualidade do ensino e do bem-estar quer da comunidade universitária, quer da sociedade em geral (WECHSLER; NAKANO, 2011).

Destaca-se que há quase um século o bibliotecário é sujeito de um fazer profissional no Brasil, entretanto, tem na sociedade contemporânea a sua ação e sua permanência como motivos para reflexão em decorrência das mudanças advindas com o uso das tecnologias de informação e comunicação (TIC), bem como as transformações no mundo do trabalho, a ampliação e ressignificação desse campo com o aporte da Ciência da Informação. Vale lembrar que o exercício de uma profissão é o aspecto que caracteriza e define o seu campo de atuação, conferindo aos seus praticantes uma identidade. Nesse

sentido, a produção de bens para uma coletividade é o aspecto que permite a esses profissionais o reconhecimento e a visibilidade, aspectos fundamentais para o reconhecimento social dessa profissão.

Para que os alunos dos cursos de Biblioteconomia no contexto brasileiro cheguem, no entanto, ao mercado de trabalho, com capacidade de intervir crítica e criativamente em suas atividades, é necessário o desenvolvimento da Competência em Informação no ensino e aprendizagem de sua formação básica. Verifica-se que há distintas concepções em cogitar o tema, de acordo com as autoras Dudziak (2003, p.30) e Belluzzo (2005a, p.37; 2005b, p.37) e, podem ser abordadas como: informação digital - ênfase na tecnologia da informação e da comunicação; informação propriamente dita - ênfase nos processos cognitivos; informação social - ênfase na inclusão social, no aprendizado ao longo da vida como exercício de cidadania.

Evidencia-se a necessidade de os alunos dos cursos de Biblioteconomia aprenderem a produzir um conhecimento novo com origem em uma dessas concepções ou em seu conjunto, para que possam aplicá-lo a uma necessidade específica, na resolução de problemas, na tomada de decisão; ou buscando o aprendizado contínuo, desenvolvendo competências que objetivem conhecer e usufruir do complexo ambiente informacional. Assim, a Competência em Informação permite promover a interdisciplinaridade (referindo-se à abordagem epistemológica dos objetos da informação e do conhecimento), e a transversalidade (referindo-se à sua dimensão didática) por se fundamentar na crítica de uma concepção de construção de conhecimento a partir de uma realidade, sujeita a um ato de conhecer isento e distanciado e apontar para a complexidade do real e a necessidade de se observar a teia de relações entre aspectos de diversidade e contradições, se forem considerados os preceitos dos Parâmetros Curriculares Nacionais (BRASIL, 2002, p. 29).

Além disso, é considerada também como sendo tanto transdisciplinar quanto multidisciplinar, por envolver a informação, o conhecimento, a tecnologia e a inteligência, que são elementos indispensáveis à pesquisa, independentemente da área de atuação do profissional bibliotecário, promovendo o aprendizado contínuo e o uso crítico da informação para produzir novos conhecimentos. De modo geral, a Competência em Informação, na literatura da área da Ciência da Informação, estabelece relação entre o acesso e uso da informação de forma crítica e ética, para possibilitar a tomada de decisão e resolver problemas pessoais e coletivos.

A transdisciplinaridade é uma abordagem científica que visa à unidade do conhecimento. Desta forma, procura estimular uma nova compreensão da realidade, articulando elementos que passam entre, além e através das disciplinas, numa busca de compreensão da complexidade. (ROCHA FILHO, 2007). Multidisciplinar são as experiências, determinado campo do conhecimento, englobando várias disciplinas, que se utilizam para fins de estudo, como parte de um currículo escolar ou dentro de um programa específico. (ROCHA FILHO, 2007).

A American Library Association (ALA) (1989) nomeia os requisitos básicos para a pessoa ser competente em informação, sendo eles: saber buscar, avaliar, filtrar e usar a informação quando necessária. É preciso, contudo, que os cursos de Biblioteconomia insiram em sua integralização curricular ou até mesmo em seus programas educacionais, a filosofia do desenvolvimento da Competência em Informação, e que objetivem tornar os alunos aprendizes autônomos e protagonistas no cenário social, enfatizando a integração currículo-sociedade. Nesse momento, entra em destaque o conceito de aprendizagem ao longo da vida.

As habilidades desenvolvidas por via da Competência em Informação podem se tornar fator de vantagem profissional e em todos os âmbitos da vida, pois, de acordo com a American Association of School Librarians (AASL) (2001), a Competência em Informação prepara a pessoa a acessar e usar a informação, de forma que a ela possibilite agir com proveito das oportunidades inerentes à sociedade da informação globalizada. Costa (2014) acrescenta ainda que, essa competência em informação deve ser um conjunto de conhecimentos, ou disciplinas, pertencente à área de Ciência da Informação.

METODOLOGIA

No intuito de perceber e identificar as competências sinalizadas anteriormente, definimos a pesquisa como bibliográfica, documental com a finalidade de identificar pontos em comum entre os temas: competência em informação, atuação bibliotecária e ambientes digitais. A análise dos dados compreende os seguintes aspectos: ementa, objetivos e conteúdo do plano de ensino da disciplina Seminário de Atuação Profissional da unidade curricular de Recursos e Serviços de Informação; e as informações coletadas no questionário respondido pelos alunos e pela docente. A realidade é construída socialmente e entendida como o compreendido, o interpretado, o comunicado. Então, a realidade não é única: existem tantas quantas forem as suas interpretações e comunicações. (Gil, 1999; Triviños, 1992).

Configura-se como uma pesquisa bibliográfica, documental e experimental com a finalidade de identificar pontos em comum entre os temas: competência em informação, atuação bibliotecária e ambientes digitais.

Os instrumentos de coleta de dados foram compostos por um roteiro de perguntas para serem respondidas pelos alunos matriculados na disciplina, em epígrafe, sendo respondidos por dezoito alunos que participaram da pesquisa.

Os procedimentos de análise foram pautados nos seguintes elementos, conforme dito anteriormente: do quadro abaixo de padrões e indicadores da ACRL, ementa, objetivos e conteúdos do plano de ensino da citada disciplina pertencente à unidade curricular de Recursos e Serviços de Informação. Também foram analisadas as informações coletadas do questionário respondido pelos alunos, acreditando assim que esses dados serão significativos para o progresso da disciplina.

Quando 1 - Padrões e Indicadores da ACRL.

PADRÕES	INDICADORES DE DESEMPENHO
1 Determina a natureza e a extensão da necessidade de informação.	1.1 Define/reconhece a necessidade de informação 1.2 Identifica uma variedade de tipos e formatos de fontes de informação potenciais 1.3 Considera os custos e benefícios da aquisição da informação necessária
2 Acessa a informação necessária com efetividade.	2.1 Seleciona os métodos mais apropriados de busca e/ou sistemas de recuperação da informação para acessar a informação necessária. 2.2 Constrói e implementa estratégias de busca delineadas com efetividade. 2.3 Busca a informação via eletrônica ou com pessoas utilizando uma variedade de métodos. 2.4 Retrabalha e melhora a estratégia de busca quando necessário 2.5 Extrai, registra e gerencia a informação e suas fontes
3 Avalia criticamente a informação e as suas fontes.	3.1 Demonstra conhecimento da maior parte das ideias da informação obtida 3.2 Articula e aplica critérios de avaliação para a informação e as fontes 3.3 Compara o novo conhecimento com o conhecimento anterior para determinar o valor agregado, contradições ou outra característica da informação.
4 Usa a informação com efetividade para alcançar um objetivo/obter um resultado.	4.1 É capaz de sintetizar a informação para desenvolver ou completar um projeto 4.2 Comunica os resultados do projeto com efetividade
5 Compreende as questões econômicas, legais e sociais da ambiência do uso da informação e acessa e usa a informação ética e legalmente.	5.1 Demonstra compreensão sobre as questões legais, éticas e socioeconômicas que envolvem a informação e a tecnologia 5.2 Cumpre as leis, regulamentos, políticas institucionais e normas relacionadas ao acesso e uso às fontes de informação 5.3 Indica as fontes de informação nas comunicações do produto ou resultados

Fonte: ACRL, 2000.

PRINCIPAIS RESULTADOS

As informações resultantes da análise dos planos de ensino estão organizadas em três eixos, sendo eles:
 - Entendimento da ementa da disciplina Seminário de atuação profissional em relação ao desenvolvimento de competências para ambientes digitais; Entendimento dos objetivos da disciplina em relação aos padrões e indicadores da ACRL; Indicadores de domínio descritos pelos alunos a fim de facilitar o entendimento da análise dos dados.

Destacamos como principal resultado o seguinte:

A percepção dos estudantes foi manifestada da seguinte forma: - **Indicadores relacionados à habilidade em determinar a natureza e a extensão da necessidade de informação.** Observa-se que os indicadores com maior sinalização são os que estão relacionados a habilidade de definir e reconhecer a necessidade de informação do usuário e de identificar uma variedade de tipos e formatos de fontes de informação potenciais para indicar ao usuário. Verificamos que os alunos têm dificuldade em considerar os custos e benefícios da aquisição da informação nos ambientes digitais. - **Indicadores relacionados à habilidade em acessar a informação necessária com efetividade.** Os alunos identificaram que possuem conhecimento e habilidades para buscar a informação em meio digital utilizando uma variedade de métodos, e, portanto, não reconhecem qual ou quais são os métodos mais apropriados para acessar a informação, bem como utilizar as estratégias de busca. Eles identificaram que sabem extrair,

registrar e gerenciar a informação e suas fontes em ambientes digitais. - **Indicadores relacionados à habilidade em avaliar criticamente a informação e as suas fontes.**

Em relação ao processo de avaliação e a crítica que se refere a informação e as suas fontes no ambiente digital, os alunos sinalizaram que possuem habilidades de conhecer a maior parte das ideias da informação obtida no ambiente digital e conseguem comparar o novo conhecimento com o conhecimento anterior para determinar o valor agregado, contradições ou outra característica da informação. Entretanto, evidenciam que têm dificuldade em articular e aplicar critérios de avaliação para a informação e as fontes em meio digital. - **Indicadores relacionados à habilidades em usar a informação com efetividade para alcançar um objetivo, obter um resultado.** Usar a informação com efetividade é fundamental para obter um resultado bem sucedido. Nesse quesito os alunos evidenciaram que possuem habilidades em sintetizar a informação para desenvolver ou completar um projeto, entretanto possuem baixa habilidade em comunicar os resultados do projeto com efetividade. - **Indicadores relacionados à habilidade em compreender as questões econômicas, legais e sociais da ambiência do acesso e uso da informação de forma ética e legalmente.**

A compreensão das questões econômicas, legais e sociais da ambiência do uso da informação ética e legal é um quesito fundamental para atuação profissional. Observa-se que os alunos têm conhecimento e habilidade de referenciar as informações nas comunicações do produto ou resultados e identificam e cumprem as leis, regulamentos, políticas institucionais e normas relacionadas ao acesso e uso das fontes de informação em ambientes digitais. Entretanto, há uma baixa compreensão sobre as questões legais, éticas e socioeconômicas que envolvem a informação e a tecnologia.

Os indicadores referentes às habilidades informacionais que precisam ser desenvolvidos durante a formação dos alunos para sua atuação em ambientes digitais, são: a) Considerar os custos e benefícios da aquisição da informação necessária; be) Construir e implementar estratégias de busca delineadas com efetividade; c) Retrabalhar e melhorar a estratégia de busca quando necessário; d) Articular e aplicar critérios de avaliação para a informação e as fontes; e) Comunicar os resultados do projeto com efetividade; f) Demonstrar compreensão sobre as questões legais, éticas e socioeconômicas que envolvem a informação e a tecnologia.

Ao analisar as habilidades informacionais acima citadas, verificamos que há necessidade de refletir, de forma mais específica, a questão de leis, regulamentos, normas e políticas como, também, de custo e benefício da aquisição da informação, estratégia de recuperação e critérios de avaliação de fontes. Essas habilidades podem ser desenvolvidas por meio de aulas com situações-problema, bem como com outros métodos que permitam ao aluno se deparar frente à realidade na qual ele precise ativar seu conhecimento em solucionar a situação apresentada.

CONCLUSÃO

Acreditamos que o resultado desta pesquisa permita contribuir de forma significativa para o desempenho da atuação do profissional em ambientes digitais, com olhar humano e social, oferecendo contribuições ao avanço técnico-científico na área da Biblioteconomia e Ciência da Informação. Evidenciou-se a necessidade de estimular os alunos a aprender a produzir um conhecimento novo originado em uma dessas concepções ou em seu conjunto, no sentido de que possa aplicá-lo a uma necessidade específica, na resolução de problemas, na tomada de decisão; ou, ainda, buscando o

aprendizado contínuo, e, pois, desenvolvendo competências que objetivem conhecer e usufruir do complexo ambiente informacional.

Assim, as observações captadas por meio das respostas dos alunos respondentes da pesquisa se coadunam com o pensamento e respostas da docente que ministra a disciplina, que propõe o aumento da carga horária, a fim de que possa trabalhar com mais práticas, palestras de profissionais especialistas, possibilitando igualmente entrevistas com bibliotecários para o conhecimento de sua atuação, a par de poder avaliar a ética ou a falta dela no cotidiano do trabalho dos profissionais.

Concluimos que os alunos participantes da pesquisa possuem habilidades básicas para atuação em ambientes digitais; entretanto, torna-se necessário desenvolver ambientações que possibilitem o desenvolvimento dessas habilidades que demonstraram fragilidades. É necessário que haja mais envolvimento por parte dos discentes para que possam acompanhar a motivação e empenho proporcionado pela docente em relação às possibilidades de desenvolvimento das habilidades informacionais de forma ativa e dinâmica.

REFERÊNCIAS BIBLIOGRÁFICAS

ASSOCIATION OF COLLEGE AND RESEARCH LIBRARIES(2000). **Information literacy competency standards for higher education**. Chicago: ALA/ACRL.

BELLUZZO, R. C. B. C.(2013). Competência em Informação: vivências e aprendizado. In: Regina Celia Baptista Belluzzo; Glória Georges Feres. (Org.). **Competência em Informação: das reflexões às lições aprendidas**. São Paulo: FEBAB, v. 1. p. 58-74.

BRASIL. (2002). Ministério da Educação. Conselho Nacional da Educação. ResoluçãoCNE/CP nº 3 de 18 de dezembro de 2002.Diário Oficial [da] República Federativado Brasil, Brasília, DF, 23 dez.

COSTA, Maria de Fátima Oliveira. (2014). **Concepções dos estudos de usuários na visão dos professores dos cursos de biblioteconomia brasileiros**. 2014. 237 f. Tese (Doutorado em Ciência da Informação) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Marília – SP.

_____. (2016). **Estudos de usuários da informação: ensino e aprendizagem no Brasil**. Fortaleza, CE: Edições UFC.

DUDZIAK, E. A. (2003). Informationliteracy: princípios, filosofia e prática. **Ciência da Informação**, Brasília, v.32, n.1, p.23-35, jan/abr.

FARIAS, G. B. (2014). **Competência em Informação no Ensino de Biblioteconomia:**

por uma aprendizagem significativa e Criativa. 183 f. Tese (Doutorado em Ciência da Informação) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Marília – SP.

GIL, A. C. (1999). **Métodos e técnicas de pesquisa social**. São Paulo: Atlas.

HUNT, T. (1997). **Desarrolla tu capacidad de aprender: la respuesta a los desafíos de la era de la información**. Barcelona: Urano.

ONTORIA PEÑA, A.; GÓMEZ R., J. P.; RUBIO MOLINA, A. (2004). **Potencializar a capacidade de aprender e pensar: o que mudar para aprender e como aprender a mudar**. São Paulo: Madras.

TRIVINOS, Augusto Nivaldo Silva. (1992). **Introdução à pesquisa em ciências sociais: a pesquisa qualitativa em educação**, São Paulo. Atlas.

VALENTIM, Marta Pomim.(2000). Atuação e perspectivas profissionais do profissional da informação. *In: _____ Profissionais da informação: formação, perfil e atuação profissional*. São Paulo: Ed. Polis. p.135-152.

VÁSQUEZ, A. Sanchez. (1990). **Filosofia da práxis**. 4. Ed. Rio de janeiro: Paz e Terra. São Paulo: Ed. Polis. p.135-152.



POLÍTICA DE INDEXAÇÃO PARA BIBLIOTECAS: ESTUDO DE ELABORAÇÃO, IMPLANTAÇÃO E AVALIAÇÃO COM PESQUISA-AÇÃO INTEGRAL¹

Mariângela Spotti Lopes Fujita

*Universidade Estadual Paulista Júlio de Mesquita Filho-Unesp, 0000-0002-8239-7114,
fujita@marilia.unesp.br*

RESUMO A indexação não é processo isolado na organização da informação porque depende de um conjunto de condutas normativas, metodologias, padrões e linguagem sistematizados em Manual de Política de indexação. elaborado conforme itens de infraestrutura, organização e clientela do sistema de informação articulados com as variáveis e elementos que influenciam a recuperação da informação. A recuperação da informação com especificidade e precisão é mais exigida aos catálogos *online* de bibliotecas tendo em vista a facilidade atualmente existente na web. A indexação com metodologias próprias de leitura, análise, identificação e seleção de conceitos, bem como de uso de linguagens de indexação conceituais, apresenta-se como recurso a ser amplamente utilizado na representação temática de registros bibliográficos e de autoridade de catálogos *online*. Com objetivo de sistematização do estudo de política de indexação realizado com uso da metodologia de pesquisa-ação integral com catalogadores de bibliotecas, apresentam-se as etapas de elaboração, implantação e avaliação durante a construção do manual de política de indexação.

PALAVRAS-CHAVE *política de indexação, indexação, catálogos de bibliotecas*

ABSTRACT Indexing is not an isolated process in the organization of information because it depends on a set of normative procedures, methodologies and systematized language and standards. An Indexing Policy Manual was elaborated according to the libraries infrastructure, organization and clientele of the information system taking into account the variables and factors that influence the information retrieval process. Since the web currently provides an easy and fast access, specific and precise information retrieval is more required in online catalogs of libraries. Indexing methodologies for reading, analysis, identification and selection of concepts, as well as the use of conceptual indexing languages, are some of the resources that can be widely used in the thematic representation of bibliographic and authority records of online catalogs. This paper presents the stages of elaboration, implementation and evaluation for the construction of an Indexing Policy Manual aimed to systematize the study of an indexing policy based on the methodology of integral action-research with libraries catalogs.

KEY WORDS *indexing, indexing policy, library catalogs*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

¹ Resultado do projeto de pesquisa “Política de indexação para bibliotecas” realizado com apoio do CNPq

INTRODUÇÃO

Indexação em bibliotecas é uma necessidade cada vez mais presente quando o assunto é recuperação da informação com especificidade, precisão e revocação em catálogos *online*. Mas, como e porque implantar a indexação para catálogos *online*? Para responder a esta questão, é preciso entender que a indexação envolve uma cultura a ser assimilada e seguida e que não diz respeito somente ao processo de indexação ou à linguagem de indexação, mas, sobretudo à definição de uma política de indexação realizada com a equipe da biblioteca.

As bibliotecas têm a necessidade de sistematizar processos e condutas de indexação e para isso dependem de uma política de indexação elaborada em consenso por catalogadores para o planejamento, elaboração e implantação de normas, procedimentos, técnicas e manual da política de indexação com orientações gerais e específicas.

Em estudo sobre política de tratamento da informação documentária com nove bibliotecas universitárias da Universidade Estadual Paulista Júlio de Mesquita Filho (Unesp) para coleta de dados em três áreas do conhecimento – Odontologia, Engenharia Civil e Pedagogia, detectou-se que apesar de possuir um vasto acervo bibliográfico, a recuperação por assuntos não atendia as necessidades dos usuários. Os resultados indicaram, ainda, mudanças significativas no comportamento informacional de acesso e uso do catálogo por usuários que solicitam recuperação por assuntos com mais especificidade, compatibilidade com sua linguagem de busca e disponibilidade de mecanismos de interação (Fujita; Rubi; Boccato, 2009) anteriormente identificados nos estudos de Zumer e Zeng, 1994, Anderson, 1998, Miller, 2004, Gross e Taylor, 2005,

Com a constatação do problema de recuperação, os catalogadores das bibliotecas universitárias da Unesp decidiram, após dois cursos de capacitação sobre fundamentos da indexação para a catalogação em bibliotecas, pela elaboração do manual de indexação do sistema de bibliotecas da Unesp para elaboração, implantação e avaliação da política de indexação com o objetivo de aprimoramento da recuperação por assunto no catálogo *online*. Formou-se, para isso, Grupo de Política de Indexação com catalogadores das bibliotecas universitárias da Unesp e pesquisadores que acompanhou todo o desenvolvimento do Projeto de Pesquisa “Política de indexação para bibliotecas” (FUJITA, 2010).

Com o objetivo de sistematização do estudo de política de indexação realizado com uso da metodologia de pesquisa-ação integral, propõe-se apresentar as etapas de elaboração, implantação e avaliação de construção do manual de política de indexação para a rede de bibliotecas universitárias da Universidade Estadual Paulista.

O desenvolvimento da investigação utilizou metodologia de pesquisa-ação integral de Morin (2004) para elaboração, implantação e avaliação da política de indexação para bibliotecas com o Grupo de Política de indexação.

A POLÍTICA DE INDEXAÇÃO E O CONTEXTO DE BIBLIOTECAS

O funcionamento de uma unidade de informações deve ser considerado na elaboração de uma política de indexação com o objetivo de viabilizar e aprimorar o sistema de recuperação da informação. Esta visão mais ampliada da política de indexação tem enfoque na visão de Gil Leiva (2008, p.297) que a

compreende de dois modos: “[...] tanto a forma de realizar a indexação de uma determinada instituição como o esforço por concretizar, sistematizar e representar em manuais os processos seguidos na indexação.”

A elaboração de uma política de indexação para viabilizar e aprimorar o sistema de recuperação da informação deve considerar: “[...] características e objetivos da organização, determinantes do tipo de serviço a ser oferecido; identificação dos usuários, para atendimento de suas necessidades de informação e recursos humanos, materiais e financeiros [...]” (Carneiro, 1985, p. 221). Conforme Guimarães (2004) e Gomes (1989), estes aspectos estão vinculados a dois eixos, horizontal e vertical, que se configuram como estrutura da elaboração de política de indexação em um sistema de recuperação da informação. No eixo horizontal estão as atividades de gestão, em que recomenda-se a identificação da organização, a clientela e a infraestrutura, as quais são variáveis a serem examinadas em cada ambiente institucional para que seja possível a correlação com o eixo vertical das atividades de organização composta pelos elementos da política de indexação: cobertura de assuntos, seleção e aquisição de documentos-fonte, nível de exaustividade e de especificidade da indexação, escolha da linguagem de indexação, estratégia de busca e tempo de resposta do sistema.

Existem, portanto, duas visões de política de indexação que se complementam: uma visão mais ampliada que considera o contexto da gestão de sistemas de recuperação da informação sobre as atividades de indexação e a visão mais direcionada aos procedimentos e elementos de indexação circunscritos ao manual de indexação, sua operacionalização e avaliação.

As duas visões se complementam porque a política de indexação é um conjunto de procedimentos, materiais, normas e técnicas orientadas por decisões que refletem a prática e princípios teóricos da cultura organizacional de um sistema de informação.

Dessa forma, a política de indexação deve ser elaborada com uma visão mais ampliada que considera o contexto da gestão de sistemas de recuperação da informação sobre as atividades de indexação e a outra visão mais direcionada aos procedimentos e elementos de indexação circunscritos ao manual de política de indexação, necessário para o registro circunstancial das decisões sobre a operacionalização e avaliação.

Em bibliotecas o tratamento da informação inclui a catalogação para representação descritiva de seus recursos informacionais e, para a representação temática, a catalogação de assuntos e a classificação. A indexação, antes mais utilizada para serviços de indexação de resumos, desenvolveu-se mediante processos e sistemas de organização do conhecimento, tais como tesouros, ontologias, taxonomias (Hjørland, 2008), além de medidas de profundidade da indexação, determinadas pela exaustividade e especificidade passíveis de avaliação (Cleveland, Cleveland, 2001), atualmente utilizadas por qualquer tipo de unidade de informação cujo objetivo seja representar para recuperar informação.

A exaustividade e especificidade são alguns dos elementos que, articulados ao tipo de organização institucional, clientela e infraestrutura definem a política de indexação. Golub (2015, p.14) exemplifica a diferença entre a política de indexação da biblioteca escolar e da biblioteca universitária ao utilizar essa articulação:

Uma coleção de vários milhares de recursos de informação em temas gerais, como em uma pequena biblioteca escolar, por exemplo, não precisaria ser muito profunda ou específica; Uma biblioteca universitária com centenas de milhares de recursos de informação teria uma política de indexação mais exaustiva e específica.

Em estudos sobre a existência (Fujita, Santos, 2016, Fernandes, Prudencio, 2016) e ausência (Nunes, 2004) de política de indexação em bibliotecas observa-se que no Brasil as bibliotecas não articulam os conhecimentos teórico-práticos da indexação à gestão da informação e que a formalização da política de indexação é inexistente, o que justifica a ausência (Nunes, 2004). No estudo com bibliotecas do município do Rio de Janeiro (Fernandes e Prudêncio, 2016, p.121) observou-se que “em 56 bibliotecas mapeadas, observou-se que 73% não adotam política de indexação formalizada” e que os resultados das “entrevistas com gestores de 4 unidades sugerem investigar a ausência de política de indexação como decorrente da frágil formação administrativa de bibliotecários.” O diagnóstico realizado com 36 bibliotecas das regiões sul e sudeste do Brasil na investigação de Fujita e Santos (2016) comprovou que existem estágios de política de indexação entre as bibliotecas: completo, parcialmente completo e incompleto com deficiências. Esses estágios demonstram que a maioria das bibliotecas apresenta política de indexação completa e parcialmente completa, porém, a política regulamentada em manual de indexação são poucas que possuem e outras possuem manual de rotinas que poderiam ser transformados em manuais mais completos de política de indexação.

METODOLOGIA

O objetivo de formar equipes de catalogadores de bibliotecas para elaborar, implantar e avaliar política de indexação foi realizado mediante aplicação da metodologia de pesquisa-ação para elaboração e implantação e avaliação de política de indexação com intervenção na realidade dos participantes e das bibliotecas.

A escolha da metodologia da pesquisa-ação se deveu à necessidade do Projeto de Pesquisa “Política de indexação para bibliotecas” (Fujita, 2010) vivenciar, junto à equipe de profissionais das bibliotecas universitárias da Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP), a ação de elaboração, implantação e avaliação da política de indexação como forma sistematizada de melhor observar os problemas e as mudanças.

O universo da pesquisa-ação integral são 32 bibliotecas universitárias da Unesp, presentes em 23 cidades do interior do Estado de São Paulo no Brasil e o objeto de estudo é o catálogo *online* Athena. Os participantes do desenvolvimento da pesquisa-ação integral são 11 catalogadores indicados pela Coordenadoria Geral de Bibliotecas da Unesp, pesquisadores e estudantes de graduação e pós-graduação bolsistas que compõem os Grupos de Política de Indexação e Linguagem Unesp.

Tendo em vista a complexidade que envolvia a elaboração da política de indexação mediante resultados obtidos em pesquisas anteriores acerca da importância da política de indexação em catálogos coletivos on-line de bibliotecas universitárias, realizou-se estudo de observação participante de pesquisadores em política de indexação com um grupo de catalogadores de bibliotecas universitárias, visando elaborar, implantar e avaliar a política de indexação a partir da metodologia da pesquisa-ação.

A aplicação da metodologia foi realizada basicamente pela vivência do grupo com a realidade dos catalogadores e destes com a dos pesquisadores em uma contínua e espiralada troca de conhecimentos para a formulação e solução dos problemas e das restrições. A vivência trouxe a observação do prisma da pesquisa e entregou a análise com propostas de processos, métodos e instrumentos. Nessa perspectiva, a vivência foi desenvolvida durante reuniões quinzenais, ocasião em que se realizavam as discussões e também a elaboração, implantação e avaliação das propostas. As soluções idealizadas pelo

grupo aconteciam, também, por intermédio de cursos de capacitação (Quadro 1), elaborados e ministrados pelos pesquisadores aos catalogadores do grupo que, em seguida, capacitavam os demais catalogadores das bibliotecas, ou, eram cursos eram elaborados pelos catalogadores proficientes do grupo quando era necessário o conhecimento mais técnico (itens 4, 5 e 6 do Quadro 1). As avaliações dos processos, métodos e instrumentos da política de indexação foram sempre realizadas após período de tempo da realização dos cursos de capacitação.

Quadro 1: Cursos de capacitação para elaboração, implantação e avaliação da Política de indexação

Período – primeira etapa: elaboração da Política de indexação	
1. “Elaboração da Terminologia de Assuntos Unesp: possibilidades e perspectivas” para o Grupo de Política de Indexação da Rede de Bibliotecas da Unesp	17 e 18 de novembro de 2010
2. “Modelo de Leitura Documentária para Indexação na Catalogação de Assuntos de Livros em Bibliotecas” para o Grupo de Política de Indexação da Rede de Bibliotecas da Unesp	1º de junho de 2011
Período – segunda etapa: implantação e avaliação da Política de indexação	
3. “Modelo de Leitura Documentária para Indexação na Catalogação de Assuntos de Livros em Bibliotecas” para os catalogadores das bibliotecas darem início à implantação da proposta de Política de Indexação	16 e 17 e 23 e 24 de abril 2012, e 25 e 26 de junho de 2012
4. “Catalogação do registro de autoridade em formato Marc21” para o Grupo de Política de Indexação da Rede de Bibliotecas da Unesp	24 de agosto de 2012
5. “Catalogação do registro de autoridade em formato Marc21” realizado pelo Grupo de Política de Indexação para os catalogadores da Rede de Bibliotecas da Unesp	12 a 14 de novembro de 2012 e 03 e 04 de dezembro de 2012
6. “Ferramenta MARC EDIT” para o Grupo de Política de Indexação da Rede de Bibliotecas da Unesp	10 de abril de 2013
7. “Terminografia” para o Grupo de Linguagem	28 de abril de 2014
8. “Modelo de Leitura para indexação de textos narrativos de ficção” pela Profª Drª Deise Maria Antonio Sabbag para os Grupos “Política de Indexação” e “Catalogação” e catalogadores da Unesp	4 e 5 de setembro de 2014

Fonte: Fujita, 2016.

A pesquisa-ação integral é estruturada, de acordo com Morin (2004, p.61), em uma grade de análise do processo com 5 componentes: *contrato, participação, mudança, discurso e ação*. Morin (2004, p.61) esclarece a função de cada componente de modo objetivo, para facilitar sua aplicação que deriva do próprio entendimento da abordagem socioconstrutivista, e destaca que esses componentes são interdependentes e que a forma de compreendê-las deve ser dinâmica e sistêmica como expressado abaixo:

A *participação* é essencial. O *contrato* é uma das condições que a asseguram. A *mudança* é a finalidade. Os efeitos dessa participação ou ação negociada estão presentes no *discurso* ou na transformação dos espíritos e na *ação* encarada para resolver ou equacionar um problema da melhor maneira possível.

RESULTADOS

A metodologia qualitativa da pesquisa-ação subsidiou a análise de resultados e demonstrou o contexto e o panorama de complexidade de um sistema de bibliotecas universitárias com vistas ao estabelecimento de diretrizes para a construção de política e de manual de indexação.

O desenvolvimento da política de indexação pelo Grupo de Política de Indexação da Rede de Bibliotecas da Unesp compreendeu duas etapas cronologicamente demarcadas: na primeira, ocorreu a elaboração de uma proposta de política de indexação e, na segunda, a implantação da proposta, como segue.

DIAGNÓSTICO E ELABORAÇÃO DA PROPOSTA DE POLÍTICA DE INDEXAÇÃO

A primeira pesquisa “Avaliação da linguagem de indexação pela recuperação no catálogo” foi realizada a fim de se obter comprovação dos índices de precisão na recuperação por assuntos, campo 650, dos registros bibliográficos do catálogo *online* Athena mediante avaliação comparada das linguagens de indexação utilizadas por bibliotecas no Brasil, “Linguagem de Cabeçalhos de Assuntos da Rede Bibliodata” (LCARB), Vocabulário Controlado do Sistema de Bibliotecas da Universidade de São Paulo (VocaUSP) e linguagem natural.

A linguagem natural foi caracterizada pela linguagem que o usuário utilizou no momento da realização da busca por assunto. A Lista de Cabeçalhos de Assunto da Rede BIBLIODATA, elaborada pela Fundação Getúlio Vargas, era a linguagem de indexação utilizada até 2013 para a indexação e recuperação da informação no catálogo coletivo Athena que integra os acervos das trinta e quatro bibliotecas da Rede UNESP. A LCARB é uma linguagem pré-coordenada, de cabeçalhos de assunto autorizados e não autorizados. A Lista de Cabeçalhos de Assunto da Rede BIBLIODATA nunca foi disponibilizada pelo sistema de recuperação da informação na busca por assunto.

O Vocabulário Controlado do SIBi/USP foi desenvolvido pelo Sistema Integrado de Bibliotecas da Universidade de São Paulo (SIBi/USP). O VocaUSP é uma linguagem pós-coordenada, constituída de um repertório terminológico formado a partir da linguagem de especialidade e da linguagem natural de termos descritores e não descritores de várias áreas do conhecimento. Os relacionamentos hierárquicos e de equivalência existentes entre os termos constituem a macroestrutura do VocaUSP, possibilitando a representação e a recuperação por assunto dos documentos no catálogo *online* Dedalus da USP –que agrega o acervo das quarenta e quatro bibliotecas do SIBi/USP.

Tendo em vista os resultados e conclusões obtidas na primeira pesquisa de avaliação ficou claro que a linguagem influencia a busca, porém não tem alcance sobre uma grande quantidade de documentos recuperados pela linguagem natural, o que nos levou a considerar que existia, então, outro problema com maior influência do que a linguagem, o processo de indexação ~~ou a falta dele~~, o que mostrou a necessidade de realização de uma nova pesquisa.

A segunda pesquisa foi realizada com o objetivo de avaliar a consistência do processo de indexação a partir da linguagem de indexação LCARB, realizado conforme elementos de política de indexação em discussão para elaboração de proposta abaixo relacionados:

- Avaliar a linguagem em um processo de indexação determinado por uma política;
- Comparar os índices de precisão na recuperação da informação com uma amostra de livros indexados com e sem política de indexação.

A análise dos resultados obtidos revelou que a base reindexada conforme proposta de política de indexação para as bibliotecas universitárias da Unesp teve melhor desempenho na recuperação de assuntos dos temas selecionados, concluindo-se que:

- A definição de metodologia para o processo de indexação é necessário durante a catalogação;
- O uso de uma metodologia de indexação aumenta a exaustividade e a especificidade dando visibilidade aos conteúdos documentais das coleções das bibliotecas;

Considerando os aspectos de precisão e revocação, foi possível identificar e comprovar o problema de recuperação da informação no catálogo *online* Athena com os resultados obtidos pela primeira e segunda pesquisa. Após discussão e análise dos resultados, o Grupo de Política de Indexação recomendou que o problema de indexação na Rede de Bibliotecas da Unesp fosse solucionado mediante abordagem de dois aspectos:

Da linguagem:- a linguagem deve ser disponibilizada diretamente no sistema Aleph, para o bibliotecário, e no catálogo, para o usuário, para que o programa possa remeter o termo da linguagem natural diretamente para o termo autorizado e permitir a visualização dos respectivos registros bibliográficos.

Da Política de indexação:- deve ser adotada uma política de indexação para a Rede, com processos bem definidos e padronizados, para alcançar um grau plenamente satisfatório de precisão, exaustividade e consistência na indexação.

A primeira e segunda recomendações ensejou a realização, respectivamente, de dois cursos de capacitação ao Grupo de Política de Indexação: “Elaboração da Terminologia de Assuntos Unesp: possibilidades e perspectivas” e “Modelo de Leitura Documentária para Indexação na Catalogação de Assuntos de Livros em Bibliotecas” (Itens 1 e 2 do Quadro 1) que colaboraram para a elaboração da Política de Indexação.

A partir dos resultados das pesquisas, discussões do Grupo de Política de Indexação e das recomendações sobre a linguagem e a política de indexação foi elaborada a “Proposta de Política de Indexação para as Bibliotecas Universitárias da Unesp” (Universidade Estadual Paulista Júlio de Mesquita Filho, 2012) da Rede de Bibliotecas da Unesp conforme estudo elaborado pelo Grupo de Política de Indexação.

A partir de então se tornou vigente a adoção da “Política de Indexação para as Bibliotecas Universitárias da Unesp” que definiu padrões e procedimentos de indexação de registros de assunto no campo 650 do formato MARC21 de registros bibliográficos do catálogo *online* Athena.

IMPLANTAÇÃO E AVALIAÇÃO DA “PROPOSTA DE POLÍTICA DE INDEXAÇÃO PARA AS BIBLIOTECAS UNIVERSITÁRIAS DA UNESP” – O MANUAL DE POLÍTICA DE INDEXAÇÃO PARA BIBLIOTECAS

O início da implantação se deu a partir de capacitação aos catalogadores das bibliotecas (Item 3 do Quadro 1), acerca da proposta de política de indexação com especial enfoque para o uso do Modelo de Leitura Documentária para a indexação de textos científicos, parte essencial da política de indexação que se refere ao processo de análise de conteúdo documentária utilizando uma metodologia sistematizada para o catalogador.

A partir dessas capacitações (Itens 1 a 6 do Quadro 1) houve uma mudança no processo de catalogação visando uma melhor indexação de assuntos do campo 650 dos registros bibliográficos com relação à: identificação dos registros modificados e acompanhamento da elaboração de registros de autoridade.

A partir de capacitação sobre “Formato de autoridade MARC21”, os catalogadores passaram a inserir novos registros de autoridade para nomes pessoais, corporativos e títulos uniformes (campos 100, 110 e 130 do formato MARC21), de assuntos tópicos (campos 150) e de assuntos geográficos (campos 151) na base principal (UEP10) de forma padronizada à medida que os registros bibliográficos eram criados ou modificados.

Com a aplicação da nova metodologia de trabalho foi possível a definição de um universo de 2.130 registros bibliográficos reindexados (marcados com campo IND) do qual foi retirada uma amostra de 600 registros para realizar uma terceira avaliação com o objetivo de verificar o uso do processo de indexação por meio das linguagens “Terminologia de Assuntos da Fundação Biblioteca Nacional”, “*Library Congress Subject Headings*” (LCSH) e “*Medical Subject Headings*” (MeSH). Essa avaliação utilizou a ferramenta MarcEdit para realizar por lotes a verificação e correção dos registros bibliográficos com campo IND.

Os resultados dessa terceira avaliação demonstraram que a política de indexação havia sido internalizada pelos catalogadores tendo em vista que a amostra de registros bibliográficos observada revelou que, de fato, eles realizaram os procedimentos e utilizaram os elementos e instrumentos da proposta de política de indexação. Além disso, foi observado o aparecimento de novos termos que necessitavam de controle, levando à ampliação da Linguagem Unesp.

Esses dois aspectos foram decisivos para a institucionalização da “Proposta de Política de Indexação para as Bibliotecas Universitárias da Unesp” e criação de um Grupo responsável pela Linguagem Unesp.

Nessa perspectiva foi criado o Grupo de Linguagem Unesp com a finalidade de construir a Linguagem Unesp a partir da importação dos registros de autoridade da Terminologia de Assuntos da Fundação Biblioteca Nacional do Brasil (TBN), *Library of Congress Subject Headings* (LCSH) e *Medical Subject Headings* (MeSH). O modelo terminográfico foi adotado para a criação de termos novos a partir do curso de capacitação “Terminografia” (item 7 do Quadro 1)

A institucionalização da “Proposta de Política de Indexação para as Bibliotecas Universitárias da Unesp” ocorreu com a publicação do “Manual de política de indexação para as bibliotecas universitárias da Unesp” (Universidade Estadual Paulista Júlio de Mesquita Filho, 2014) após ajustes decorrentes de avaliações e negociações originadas da aplicação da Proposta.

Em função da mudança que a política de indexação proporcionou foi solicitado pelos catalogadores o curso de capacitação “Modelo de Leitura para indexação de textos narrativos de ficção” (item 8 do Quadro 1) , sobre como indexar textos narrativos de ficção com metodologia adequada às estruturas textuais narrativas, idealizado por Sabbag (2013), pesquisadora do grupo, ..

Considerando os desafios enfrentados e a experiência adquirida, o Grupo de Política de Indexação organizou e publicou livro sobre todo o conhecimento adquirido na elaboração, implantação e avaliação da política de indexação com a finalidade de divulgá-lo às demais redes de bibliotecas (Fujita, 2016).

CONSIDERAÇÕES FINAIS

A elaboração, implantação e avaliação da política de indexação dependeu da participação ativa do Grupo de Política de indexação, cujos resultados são as avaliações e seus resultados, o Manual de política de indexação para bibliotecas e o livro que relata todo o processo realizado.

A pesquisa-ação integral realizada durante o processo de elaboração, implantação e avaliação promoveu a reflexão com a finalidade de identificar problemas ou avaliar mudanças. Os seminários de vivências profissionais, capacitações ou compartilhamento de experiências e dúvidas foram realizados para apresentar: sistematização teórica e metodológica sobre indexação e experiências de elaboração de políticas de indexação para a elaboração do plano de **ação**. A implementação e avaliação da Política de Indexação (plano de ação) pelas bibliotecas foi passo-a-passo discutido e analisado tanto pelo Grupo de Política de Indexação quanto pelos catalogadores e gestores das bibliotecas universitárias em reuniões, seminários e capacitações.

A partir deste relato da experiência de intervenção de pesquisa-ação integral foi possível observar a presença dos cinco conceitos da pesquisa-ação integral: o contrato, a participação, a mudança, o discurso e a ação. Com o avanço das avaliações e da elaboração da Proposta de Política de Indexação que se tornou a Política de indexação em Manual, passou a existir a **participação** de todos os atores com um **discurso** comum que culminou na elaboração da Proposta que se tornou o Manual de Política de indexação considerado, dentro da concepção da Pesquisa-ação integral, como **o contrato**. A Proposta e, depois o Manual, contém, portanto, **o discurso** da problematização, acompanhado dos resultados da avaliação, dos argumentos principais e, principalmente, das estratégias de **ações** para realizar a **mudança** de conduta.

REFERÊNCIAS BIBLIOGRÁFICAS

Anderson, S. (1998). A new horizon: an evaluation of a library online public access catalogue. *Library & Information Research News*, 22(72), 15-24.

Carneiro, M. V. (1985) Diretrizes para uma política de indexação. *Revista da Escola de Biblioteconomia da UFMG*, Belo Horizonte, 14(2), p.221-241.

Cleveland, D. B., Cleveland, A. D. (2001). *Introduction to indexing and abstracting*. 3.ed. Greenwood Village, Co.: Libraries Unlimited. 283p.

Fernandes, G. C., Prudencio, D. S. (2016) Política de indexação em bibliotecas cariocas: mapeamento exploratório. *Ponto de Acesso*, Salvador, v.9, n.1,p. 122-142, abr. 2015. Recuperado de: <https://portalseer.ufba.br/index.php/revistaici/article/viewFile/9885/9633>.

Fujita, M. S. L. (2010) *Política de indexação para bibliotecas*. Marília: Faculdade de Filosofia e Ciências, Universidade Estadual Paulista. (Projeto de pesquisa de Produtividade em pesquisa do CNPq).

Fujita, M.S.L. (2016) *Política de indexação para bibliotecas: elaboração, avaliação e implantação /* Marília: Oficina Universitária; São Paulo: Cultura Acadêmica. 142p

Fujita, M. S. L.; Santos, L.B.P. dos (2016) Política de indexação em bibliotecas universitárias: estudo diagnóstico e analítico com pesquisa participante. *Transinformação*, 28, p.59 - 76.

Gil Leiva, I. (2008) *Manual de indización*. Teoría y práctica. Gijón: Trea.

Golub, K. Subject (2015) *Access to information: an interdisciplinary approach*. Santa Barbara, CA: Libraries unlimited. 165p. Recuperado de: https://books.google.com.br/books?id=z_o7BQAAQBAJ&pg=PA14&lpg=PA14&dq=indexing+policy+definition&source=bl&ots=40yCOXP37g&sig=usYTVOG8KxBm7OKqw_vBl-S3dsI&hl=pt-BR&sa=X&ved=0ahUKEwjK1bCYz7bUAhWED5AKHe4JBcoQ6AEIczAJ#v=onepage&q=indexing%20policy%20definition&f=true.

Gomes, H. E. (1989) O indexador face às novas tecnologias da informação. *Transinformação*, Campinas, 1(1/2), p. 161-174.

Gross, T., Taylor, A. G. (2005). What have we got to lose? the effect of controlled vocabulary on keyword search-ing results. *College & Research Libraries*, 66(3), 212-230.

Guimarães, J. A. C. (2004) As políticas de indexação como elemento para a gestão do conhecimento nas. In: VIDOTTI, S. A. G. (Coord.) *Tecnologia e conteúdos informacionais: abordagens teóricas organizações e práticas*. São Paulo: Polis. p. 43-52.

Hjørland, B. (2008) What is knowledge organization? *Knowledge Organization*, 35(2/3), p.86-101.

Miller, D. H. (2004). User perception and the online catalogue: public library OPAC users “think aloud”. In: McIlwaine, I. A.C. (ed.). *Knowledge organization and the global information society: proceedings of the 8th Inter-national ISKO Conference, 13-16 July 2004, London, UK*. London: Ergon Verlag. V. 9, p. 275-280.

Morin, A. (2004) *Pesquisa-ação integral e sistêmica: uma antropedagogia renovada*. Trad. de Michel Thiollent. Rio de Janeiro: DP&A. 232 p.

Nunes, C. O. (2004) Algumas considerações acerca da ausência de políticas de indexação em bibliotecas brasileiras. *Biblos*, Rio Grande, 16, p. 55-62. Recuperado de: <https://www.seer.furg.br/biblos/article/view/411>.

Sabbag, Deise Maria Antonio. (2013) *Análise documental em textos narrativos de ficção: subsídios para o processo de análise*. 2013. Fls 138, 140-142. Tese (doutorado) - Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Filosofia e Ciências de Marília. Recuperado de: <<http://hdl.handle.net/11449/103384>>..

Universidade Estadual Paulista Júlio de Mesquita Filho. Coordenadoria Geral de Bibliotecas. (2014) *Manual de política de indexação para as bibliotecas universitárias da Unesp*. Mariangela Spotti Lopes Fujita (coord). São Paulo: Unesp. Recuperado de: <<http://www.biblioteca.unesp.br/portal/arquivos/manual-politica-indexacao>>

Universidade Estadual Paulista Júlio de Mesquita Filho. Coordenadoria Geral de Bibliotecas. (2012) *Proposta de política de indexação para as bibliotecas universitárias da Unesp*. Mariangela Spotti Lopes Fujita (coord). São Paulo: Unesp.

Zumer, M.; Zeng, L. (1994). Comparison and evaluation of OPAC end-user interfaces. *Cataloging & Classification Quarterly*. // ISSN 0163-9374. 19(2), 67-98.



A ESTRUTURA/CARGO EM BIBLIOTECA E ARQUIVO: O PODER ORGANIZACIONAL

Samir Elias Kalil Lion¹, Zeny Duarte²

¹Instituto de Ciência da Informação da Universidade Federal da Bahia, Brasil (ICI/UFBA), Doutor em
Ciência da Informação (ICI/UFBA), samireliaslion@gmail.com

²Instituto de Ciência da Informação da Universidade Federal da Bahia, Brasil (ICI/UFBA), Professora Titular
(ICI/UFBA) e Pós-doutora em Ciência da Informação (U.Porto, FCT-PT), Zeny.duarte@gmail.com

RESUMO Demonstra os resultados de uma Tese acerca do poder organizacional (bases de poder) pautado na estrutura hierárquica das bibliotecas universitárias (BUs) do Sistema de Bibliotecas da Universidade Federal da Bahia – SIBI/UFBA e da estrutura hierárquica do Arquivo Público do Estado da Bahia – APEB. A amostra foi constituída, no caso do SIBI/UFBA, pelos 4 (quatro) bibliotecários-chefe das 4 (quatro) Bibliotecas Universitárias e, no caso do APEB, foi constituída pelos 4 (quatro) coordenadores das 5 (cinco) Coordenações e o tratamento dos dados foi feito em Planilha Excel. Conclui-se que há nas BUs do SIBI/UFBA fracos indícios no sentido de ser o poder distribuído pelas equipes de trabalho, que conduz a uma reprodução e uso menos equitativa desse poder, aproximando-se de um poder bastante concentrado na autoridade do cargo. E há nas coordenações do APEB uma baixa tendência no sentido de ser o poder distribuído pelas equipes de trabalho, que conduz a uma reprodução e uso menos equitativa desse poder, aproximando-se também de um poder relevantemente concentrado na autoridade do cargo.

PALAVRAS-CHAVE *Chefe, Coordenador, Hierarquia, Arquivo Público do Estado da Bahia – APEB, Sistema de Bibliotecas da Universidade Federal da Bahia - SIBI/UFBA.*

ABSTRACT Demonstrates the results of a thesis about the organizational power (bases of power) based on the hierarchical structure of university libraries (BUs) of the Library System of the Federal University of Bahia - SIBI / UFBA and the hierarchical structure of the Public Archive of the State of Bahia - APEB. In the case of SIBI / UFBA, the sample consisted of the 4 (four) chief librarians of the 4 (four) University Libraries and, in the case of the APEB, was composed of 4 (four) coordinators of the 5 (five) coordinations and the data processing was done in Excel Worksheet. It is concluded that there is weak evidence in the BUs of SIBI / UFBA to be the power distributed by the work teams, which leads to a less equal reproduction and use of this power, approaching a power quite concentrated in the authority of the position. And there is a low tendency in the coordination of the APEB to be the power distributed by the work teams, which leads to a less equal reproduction and use of this power, approaching also a power concentrated in the authority of the position.

KEY WORDS *Chief, Coordinator, Hierarchy, Public Archives of the State of Bahia - APEB, Library System of the Federal University of Bahia - SIBI / UFBA.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Neste artigo estão presentes os resultados acerca das Bases de Poder que são utilizadas pelos chefes (ocupantes de cargos legalmente instituídos na estrutura hierárquica) para gerenciar suas equipes. Pois, se os fatores constituintes do poder organizacional nas unidades de informação (locais de organização do conhecimento) estiverem centrados no cargo, significa que ali se exerce mais a chefia e menos a liderança porque o poder é exercido principalmente através do exercício da autoridade.

Um dos estudos que levaram à decisão de pesquisar na tese em tela (LION, 2015) a perspectiva política da administração em unidades de informação, que tem relação direta com a formação e atuação profissional em organização do conhecimento, foi o de Lion (2010). O autor encontrou fragilidades em relação ao ser humano, principalmente nas interações verticais e horizontais do Sistema de Bibliotecas da Universidade Federal da Bahia - SIBI/UFBA (que dizem respeito à estrutura de poder, que é o poder centrado na autoridade do cargo). A Bahia é um dos Estados Federados que compõem o Brasil.

O objetivo é demonstrar os resultados de uma Tese acerca do poder organizacional (bases de poder) pautado na estrutura hierárquica das bibliotecas universitárias do Sistema de Bibliotecas da Universidade Federal da Bahia – SIBI/UFBA e da estrutura hierárquica do Arquivo Público do Estado da Bahia – APEB, ambos localizados no Estado da Bahia, Brasil.

METODOLOGIA

A precariedade das relações humanas (interações verticais e horizontais) encontradas na pesquisa de Lion (2010) diz respeito a atitude de aliar estrategicamente, que significa cooptar os diversos saberes dispersos nos setores da biblioteca ou nas bibliotecas do SIBI/UFBA para obter cooperação.

O outro significado de aliar estrategicamente, dentro da perspectiva estratégica, refere-se à construção de parcerias múltiplas entre os setores da biblioteca e nas bibliotecas do SIBI/UFBA para obter cooperação construindo parcerias e compartilhando experiências exitosas: e aqui a pesquisa também encontrou certa precariedade nas relações humanas.

Tais achados justificaram a tese, cujos resultados parciais são trazidos para o presente artigo, principalmente por corroborarem com a importância das políticas e dos estudos voltados para os Recursos Humanos que laboram em unidades que tem como essência a organização e disponibilização do conhecimento, algo bastante importante nos estudos organizacionais tanto na área da Ciência da Informação como principalmente na formação e atuação profissional em organização do conhecimento.

Destaque-se que em decorrência das fragilidades em relação ao ser humano, encontradas na pesquisa de Lion, principalmente nas interações verticais e horizontais (que dizem respeito à estrutura de poder, que é o poder centrado na autoridade do cargo), decidiu-se pesquisar na tese a perspectiva política da administração em unidades de informação, que tem relação direta com o poder organizacional. O objetivo geral da tese foi o de analisar o poder organizacional em Unidades de Informação sob a perspectiva política da Administração. Isso porque os fatores constituintes do poder organizacional, (voltados para questões de liderança, poder da estrutura/cargo/autoridade, conflitos/retaliação e comunicação/informação entre membros das equipes de trabalho), nas unidades de informação, que em

essência organizam o conhecimento, podem orbitar entre a chefia e a liderança. Se centrado no cargo, indica o poder exercido através do exercício da autoridade, se distribuído pelos membros das equipes de trabalho, indica a utilização dos princípios de liderança. Ou seja, se buscarmos respostas para as questões de liderança, poder da estrutura/cargo/autoridade, conflitos e retaliação e comunicação/informação entre membros das equipes de trabalho, com efeito, construiremos uma proposta de formação de pessoas que terão uma melhor atuação profissional na área de Organização do Conhecimento.

Dito de outra forma, os elementos para a análise do poder organizacional em unidades de informação e de organização do conhecimento estão: na Liderança; na Estrutura de Poder (poder centrado na autoridade do cargo); nos Conflitos de Poder; e na dependência de tarefas e resultados, que está na comunicação e informação entre os membros das equipes de trabalho.

Assim, baseado em semelhanças do discurso/enunciado (fala) do indivíduo (arquivista-chefe, bibliotecário-chefe, museólogo-chefe, chefe de atendimento, chefe do arquivo circulante ou protocolo, chefe do serviço de referência, curador etc) é possível indicar, na perspectiva social (poder na perspectiva da cultura organizacional), o discurso do grupo (equipes de trabalho), bem como a formação social à qual pertence o sujeito (estilos de funcionamento coletivamente aceitos). (FONTANELLA et al., 2008).

Para explicar o poder da estrutura/cargo representado pela autoridade do cargo, utilizou-se na tese principalmente Motta P. (2007), Clegg e Hardy (2001) e Martins (2008), este último que adaptou e validou a EBPS de French e Ravem (Escala de Bases de Poder do Supervisor) no Brasil, usada na tese para mensurar o poder da estrutura/cargo em Unidades de Informação e identificar as bases de poder organizacional. Os seguintes autores da área da Ciência da Informação foram utilizados: Barbalho (2012), Oliveira (2008), Campos e Baptista (2008). A Escala de Bases do Poder do Supervisor (EBPS) foi criada por French e Ravem, e adaptada e validada no Brasil por Martins (2008), sendo usada na tese para mensurar o poder da estrutura/cargo em Unidades de Informação e identificar as bases de poder organizacional.

Entendemos que gerentes de unidades de informação que utilizam a autoridade do cargo, estão usando o poder na perspectiva da cultura organizacional, ou seja, o poder que é aceito, ratificado ou retificado pelo coletivo. Isto é, através da Escala de Bases do Poder do Supervisor (EBPS), é possível identificar as bases sobre as quais se assentam a autoridade que é coletivamente aceita nos arquivos, bibliotecas e museus, indo além do poder centrado no chefe. Isso porque o funcionamento organizacional supõe formas coletivas de sentir, pensar e agir a organização, presentes nas equipes de trabalho; e que prova disso é a própria existência da organização, pois se a unidade de informação abre suas portas todos os dias, significa que os chefes exercem na verdade a autoridade que é aceita pelos subordinados dentro de uma negociação constante das dimensões aceitáveis, significa, portanto que há um exercício constante de distribuição do poder centrado no cargo.

A amostra foi constituída, no caso do SIBI/UFBA, pelos 4 (quatro) bibliotecários-chefe das 4 (quatro) Bibliotecas Universitárias e, no caso do APEB, foi constituída pelos 4 (quatro) coordenadores das 5 (cinco) Coordenações. E o critério de escolha da amostra da tese foi não-probabilístico e fundamentada no seguinte: os arquivistas-chefe, ou equivalentes, e os bibliotecários-chefe cotidianamente vivenciam as problemáticas ligadas ao poder organizacional (liderança, autoridade do cargo, conflitos e comunicação nas equipes). A coleta de dados se deu através da aplicação de questionários e o tratamento dos dados foi feito em Planilha Excel.

RESULTADOS

No âmbito da tese em tela (LION, 2015), entende-se que gerentes (bibliotecário-chefe e arquivistas-chefe) quando utilizam a autoridade que o cargo a eles confere, estão usando o poder na perspectiva da cultura organizacional, ou seja, o poder que é aceito, ratificado ou retificado pelo coletivo. Através da Escala de Bases do Poder do Supervisor (EBPS), é possível identificar as bases sobre as quais se assentam a autoridade que é coletivamente aceita, indo além do poder centrado no chefe. Isso porque o funcionamento organizacional supõe formas coletivas de sentir, pensar e agir a organização, presentes nas equipes de trabalho; e que prova disso é a própria existência da organização, isto é, se esta bibliotecas e arquivos abre suas portas todos os dias, significa que os chefes exercem na verdade a autoridade que é aceita pelos subordinados dentro de uma negociação constante das dimensões aceitáveis, significa portanto, que há um exercício constante de distribuição do poder centrado no cargo.

Tabela 1 - Distribuição percentual das afirmativas da Escala de Bases do Poder do Supervisor (EBPS) para mensurar o poder da estrutura/cargo nas BUs do SIBI/UFBA

AFIRMATIVAS	Discordo totalmente		Discordo parcialmente		Não tenho opinião formada/Não sei		Concordo parcialmente		Concordo totalmente		Total	
	Resp.	%	Resp.	%	Resp.	%	Resp.	%	Resp.	%	Resp.	%
1. Meu chefe pode aumentar meu salário.	4	100%	0	0%	0	0%	0	0%	0	0%	4	100%
2. Meu chefe pode me fazer sentir que tenho compromissos a cumprir.	0	0%	1	25%	0	0%	0	0%	3	75%	4	100%
3. Meu chefe pode me dar boas sugestões técnicas sobre meu trabalho.	0	0%	0	0%	0	0%	0	0%	4	100%	4	100%
4. Meu chefe pode tornar meu trabalho difícil para mim.	0	0%	3	75%	0	0%	0	0%	1	25%	4	100%
5. Meu chefe pode me fazer perceber como eu deveria cumprir as exigências do meu trabalho.	0	0%	1	25%	0	0%	1	25%	2	50%	4	100%
6. Meu chefe pode influenciar a organização para conseguir um aumento de salário para mim.	3	75%	0	0%	0	0%	0	0%	1	25%	4	100%
7. Meu chefe pode dividir comigo suas experiências ou treinamentos importantes.	0	0%	0	0%	0	0%	0	0%	4	100%	4	100%
8. Meu chefe pode conseguir benefícios especiais para mim.	1	25%	1	25%	0	0%	0	0%	2	50%	4	100%
9. Meu chefe pode influenciar a organização para me dar uma promoção.	2	50%	0	0%	0	0%	0	0%	2	50%	4	100%
10. Meu chefe pode fornecer para mim conhecimento técnico necessário ao trabalho.	0	0%	0	0%	0	0%	1	25%	3	75%	4	100%
11. Meu chefe pode tornar as coisas desagradáveis para mim.	0	0%	3	75%	0	0%	0	0%	1	25%	4	100%
12. Meu chefe pode tornar meu trabalho desagradável.	0	0%	3	75%	0	0%	0	0%	1	25%	4	100%

13. Meu chefe me faz perceber que eu tenho responsabilidades para cumprir.	2	50%	0	0%	0	0%	0	0%	2	50%	4	100%
14. Meu chefe pode me fazer reconhecer que eu tenho tarefas para realizar.	0	0%	1	25%	0	0%	0	0%	3	75%	4	100%
15. Meu chefe pode me dar dicas relacionadas ao trabalho.	0	0%	0	0%	0	0%	1	25%	3	75%	4	100%
TOTAL	12	20%	13	22%	0	0%	3	5%	32	53%	60	100%

Referência: LION, 2015.

Nas BUs do SIBI/UFBA, os chefes podem fazer com que seus subordinados sintam que têm compromissos a cumprir, pois 75% dos próprios bibliotecários-chefe respondentes se posicionaram em concordância total quanto a isso. Os chefes das BUs do SIBI/UFBA também podem fazer com que os membros das equipes de trabalho reconheçam que têm tarefas para realizar, igualmente com 75% de concordância total. Adicionalmente, a investigação destas BUs aponta uma certa relativização quanto a capacidade do chefe de fazer com que os membros das equipes percebam como deveriam cumprir as exigências do trabalho (25% dos respondentes discordam parcialmente e outros 25% concordam parcialmente), mesmo assim, 50% dos bibliotecários-chefe concordam totalmente com isso, que se somado com os 25% que concordam parcialmente, alcançará uma zona de concordância de 75%. Essa relativização aumenta quando se trata do poder do chefe em fazer com que os membros das equipes de trabalho percebam que têm responsabilidades a cumprir, pois os dados ficaram polarizados entre 50% dos respondentes que discordam totalmente e outros 50% que concordam totalmente, mesmo assim, os dados indicam que em metade das BUs pesquisadas os chefes têm esse poder. Nas BUs pesquisadas apenas 25% dos chefes podem tornar as coisas desagradáveis para os subordinados. A análise desta unidade de informação encontrou fracos indícios de que o chefe torna o trabalho difícil (25% dos respondentes concordam totalmente) e desagradável (também com 25% de concordância total). Nas bibliotecas universitárias do SIBI/UFBA, os bibliotecário-chefe não podem aumentar o salário de seus subordinados, pois nenhum dos pesquisados concordou com isso (0%). Mas 50% dos pesquisados concordam que os chefes podem influenciar estas unidades de informação a dar promoções a seus subordinados. A investigação encontrou fracos indícios de que os chefes são influentes para conseguir aumentos de salário para os membros das equipes de trabalho (25% de concordância total). Mas os dados indicam que o chefe pode conseguir benefícios especiais para os subordinados, pois além de 50% concordarem totalmente, apenas 25% discordam totalmente quanto a isso. Nestas unidades de informação pesquisadas, todos os chefes podem dar boas sugestões técnicas sobre o trabalho a seus subordinados (100% de concordância total) e todos eles dividem com a equipe suas experiências ou treinamentos importantes (também com 100% de concordância total). A análise destas BUs encontrou fortes indícios que os chefes fornecem para a equipe conhecimento técnico necessário ao trabalho (75% de concordância total e 25% de concordância parcial) e fornecem dicas relacionadas ao trabalho (também com 75% de concordância total e 25% de concordância parcial).

Tabela 2 - Distribuição percentual das afirmativas da Escala de Bases do Poder do Supervisor (EBPS) para mensurar o poder da estrutura/cargo das coordenações do Arquivo Público do Estado da Bahia - APEB

AFIRMATIVAS	Discordo totalmente		Discordo parcialmente		Não tenho opinião formada/Não sei		Concordo parcialmente		Concordo totalmente		Total	
	Resp.	%	Resp.	%	Resp.	%	Resp.	%	Resp.	%	Resp.	%
1. Meu chefe pode aumentar meu salário.	0	0%	1	25%	2	50%	1	25%	0	0%	4	100%
2. Meu chefe pode me fazer sentir que tenho compromissos a cumprir.	1	25%	0	0%	2	50%	1	25%	0	0%	4	100%
3. Meu chefe pode me dar boas sugestões técnicas sobre meu trabalho.	1	25%	0	0%	0	0%	0	0%	2	50%	3	75%
4. Meu chefe pode tornar meu trabalho difícil para mim.	1	25%	0	0%	2	50%	1	25%	0	0%	4	100%
5. Meu chefe pode me fazer perceber como eu deveria cumprir as exigências do meu trabalho.	1	25%	0	0%	2	50%	1	25%	0	0%	4	100%
6. Meu chefe pode influenciar a organização para conseguir um aumento de salário para mim.	0	0%	0	0%	2	50%	2	50%	0	0%	4	100%
7. Meu chefe pode dividir comigo suas experiências ou treinamentos importantes.	0	0%	1	25%	0	0%	1	25%	2	50%	4	100%
8. Meu chefe pode conseguir benefícios especiais para mim.	0	0%	0	0%	3	75%	1	25%	0	0%	4	100%
9. Meu chefe pode influenciar a organização para me dar uma promoção.	0	0%	0	0%	3	75%	1	25%	0	0%	4	100%
10. Meu chefe pode fornecer para mim conhecimento técnico necessário ao trabalho.	1	25%	0	0%	0	0%	1	25%	2	50%	4	100%
11. Meu chefe pode tornar as coisas desagradáveis para mim.	1	25%	0	0%	3	75%	0	0%	0	0%	4	100%
12. Meu chefe pode tornar meu trabalho desagradável.	1	25%	0	0%	3	75%	0	0%	0	0%	4	100%
13. Meu chefe me faz perceber que eu tenho responsabilidades para cumprir.	1	25%	0	0%	2	50%	1	25%	0	0%	4	100%
14. Meu chefe pode me fazer reconhecer que eu tenho tarefas para realizar.	1	25%	0	0%	2	50%	1	25%	0	0%	4	100%
15. Meu chefe pode me dar dicas relacionadas ao trabalho.	1	25%	0	0%	1	25%	0	0%	2	50%	4	100%
TOTAL	10	17%	2	3%	27	46%	12	20%	8	14%	59	100%

Referência: LION, 2015.

Nas coordenações do Arquivo Público do Estado da Bahia, os coordenadores podem dar boas sugestões técnicas sobre o trabalho a seus subordinados (50% de concordância total) e a maioria dos chefes dividem com a equipe suas experiências ou treinamentos importantes com 75% de zona de concordância (somando-se 25% que concordam parcialmente com 50% de concordância total). A análise das coordenações do APEB encontrou fortes indícios que os chefes fornecem para a equipe conhecimento técnico necessário ao trabalho (50% de concordância total e 25% de concordância parcial) e fornecem dicas relacionadas ao trabalho (com 50% de concordância total). Os chefes das coordenações do APEB

não podem aumentar o salário de seus subordinados, pois nenhum dos pesquisados concordou totalmente com isso (0%). Mas 50% dos pesquisados concordam parcialmente que os chefes podem influenciar esta unidade de informação a conseguir aumento de salário para seus subordinados. A investigação encontrou fracos indícios de que os chefes podem conseguir benefícios especiais para os membros das equipes de trabalho (25% de concordância parcial), além disso, os dados indicam que os chefes das coordenações do APEB podem influenciar pouco para que promoções sejam dadas aos subordinados (25% de concordância parcial).

Nas coordenações do APEB, são fracos os indícios de que os chefes podem fazer com que seus subordinados sintam que têm compromissos a cumprir (25% de concordância parcial). Também existem fracos indícios de que os chefes podem fazer com que os membros das equipes de trabalho reconheçam que têm tarefas para realizar (igualmente com 25% de concordância parcial). Adicionalmente, a investigação destas aponta fracos indícios quanto a capacidade do chefe de fazer com que os membros das equipes percebam como deveriam cumprir as exigências do trabalho (25% de concordância parcial). Quando se trata do poder do chefe em fazer com que os membros das equipes de trabalho percebam que têm responsabilidades a cumprir, os fracos indícios permanecem (25% de concordância parcial). Nas coordenações, os chefes não tornam as coisas desagradáveis para os subordinados, pois ninguém concordou com isso, o que é reforçado pelos 25% de discordância. A análise destas encontrou fracos indícios de que o chefe torna o trabalho difícil para os subordinados (25% dos respondentes concordam parcialmente) e nenhum indicio de que os chefes tornem o trabalho desagradável para os membros das equipes de trabalho (0% de zona de concordância).

Oliveira (2008) destaca que o desenho organizacional condiciona a estrutura de cargos nas organizações, influenciando a distribuição, configuração e o grau de especialização destes. Ou seja, a autoridade do cargo de bibliotecário-chefe, arquivista-chefe, ou equivalente, é o elemento objetivo da estrutura hierárquica do Arquivo Público do Estado da Bahia e do Sistema de Bibliotecas da UFBA, algo que por ser materializado no plano de cargos dos servidores e funcionários, nos Regimentos e normativas, pode ser racionalmente apropriado pelas pessoas das equipes de trabalho como base de poder, por isso haver a partir daí as disputas pelo acesso ao cargo, que assume a representação, consciente ou inconsciente, de acesso a autoridade e conseqüentemente ao poder.

Campos e Baptista (2008), em sua pesquisa, concluíram, em relação às características dos bibliotecários que trabalham em instituições de ensino superior particulares no Distrito Federal (Instituições de Ensino Superior - IES credenciadas ao MEC), num total de 46 profissionais, que: em relação ao cargo que ocupam na unidade de informação, a maioria afirmou que ocupa o cargo de gerente e de profissional de processo técnico (21 respondentes em cada cargo), concentrando suas atividades nestes dois itens.

O cargo de gerente de arquivo ou biblioteca, baseado em Campos e Batista (2008), tem uma forte conexão com a base estrutural de poder legítimo, referenciado na tese através das seguintes informações para as BUs do SIBI/UFBA: nestas os chefes podem fazer com que seus subordinados sintam que têm compromissos a cumprir; os chefes destas BUs também podem fazer com que os membros das equipes de trabalho reconheçam que têm tarefas para realizar. Adicionalmente, nestas BUs há uma certa relativização quanto a capacidade do chefe de fazer com que os membros das equipes percebam como deveriam cumprir as exigências do trabalho. Essa relativização aumenta quando se trata do poder do chefe em fazer com que os membros das equipes de trabalho percebam que têm responsabilidades a cumprir.

No caso das coordenações do APEB, são fracos os indícios de que os chefes podem fazer com que seus subordinados sintam que têm compromissos a cumprir. Também existem fracos indícios: de que os chefes podem fazer com que os membros das equipes de trabalho reconheçam que têm tarefas para realizar; de que o chefe tenha a capacidade de fazer com que os membros das equipes percebam como deveriam cumprir as exigências do trabalho; ou quanto ao poder do chefe em fazer com que os membros das equipes de trabalho percebam que têm responsabilidades a cumprir.

Isso demonstra que nas BUs do SIBI/UFBA e nas coordenações do APEB a autoridade do cargo é relevantemente exercida pelo reconhecimento de que as chefias têm o poder legal de influenciar os membros das equipes de trabalho através da capacidade de fazer com que os subordinados sintam que têm compromissos a cumprir, de fazer com que os liderados reconheçam que têm tarefas para realizar, de fazer com que os membros das equipes de trabalho percebam como deveriam cumprir as exigências do trabalho e de fazer os chefiados perceberem que têm responsabilidades para cumprir.

Nesse sentido, Barbalho (2012) traz o sistema de autoridade como um dos importantes componentes para a formação da estrutura organizacional em unidades de informação, pois, segundo Motta P. (2007), a estrutura ainda é a fonte principal do poder porque toda ação de poder advém em algum nível da autoridade que essa estrutura confere a seu ocupante.

A partir desses autores, o Sistema de Autoridade do SIBI/UFBA e do APEB é o revestimento do direito ou poder de se fazer obedecer, de dar ordens, de tomar decisões, enfim, de agir. É o poder, é a autoridade, é o direito de tomar decisões, de dar ordens, de requerer obediência ou simplesmente o direito de desempenhar um trabalho que foi designado. Barbalho acrescenta que devem ainda ser observados, no delineamento de uma estrutura organizacional em arquivos e bibliotecas, os sistemas de responsabilidade, de decisão e de comunicação.

A base estrutural de poder de perícia, que tem uma forte conexão com a atividade de processamento técnico e organização da informação apontadas na pesquisa de Campos e Batista, reforça o sistema de autoridade, pois a perícia é a base dominante tanto no SIBI/UFBA quanto no APEB. Na tese, as indagações sobre poder de perícia nas BUs do SIBI/UFBA se revelam assim: 3. Meu chefe pode me dar boas sugestões técnicas sobre meu trabalho (100% de concordância total), 7. Meu chefe pode dividir comigo suas experiências ou treinamentos importantes (100% de concordância total), 10. Meu chefe pode fornecer para mim conhecimento técnico necessário ao trabalho (75% de concordância total) e 15. Meu chefe pode me dar dicas relacionadas ao trabalho (75% de concordância total).

Quanto ao Arquivo Público do Estado da Bahia, boa parte dos seus coordenadores podem dar boas sugestões técnicas sobre o trabalho a seus subordinados e dividem com a equipe suas experiências ou treinamentos importantes. Os coordenadores do APEB fornecem para a equipe conhecimento técnico necessário ao trabalho e fornecem dicas relacionadas ao trabalho.

Isso demonstra que nas coordenações do APEB e nas BUs do SIBI/UFBA a autoridade do cargo é predominantemente exercida pelo reconhecimento por parte dos membros das equipes de trabalho de que os chefes dominam certo conhecimento através de boas sugestões técnicas dadas por estes sobre o trabalho dos subordinados, da capacidade de o chefe dividir suas experiências ou treinamentos importantes com os liderados, de o chefe fornecer conhecimento técnico necessário ao trabalho dos chefiados e da capacidade de o chefe dar dicas relacionadas ao trabalho.

Este poder de perícia representa a presença de um Sistema de Responsabilidade que, segundo Barbalho, é o componente antecessor ao sistema de autoridade no delineamento da estrutura organizacional em bibliotecas e arquivos, e está relacionado a atribuição de encargos inerentes àquilo que é dimensionado para que o colaborador execute as funções que ocupa. Desse modo, os chefes e subordinado das coordenações do APEB e das BUs do SIBI/UFBA, conseguem, através do sistema de responsabilidade, reconhecer os limites dos processos que estão sob sua responsabilidade.

Tabela 3 - Distribuição percentual dos totais para o Poder da Estrutura Organizacional das BUs do SIBI/UFBA

Bases Estruturais de Poder Organizacional	Discordo totalmente		Discordo parcialmente		Não tenho opinião formada/Não		Concordo parcialmente		Concordo totalmente		Total	
	Resp.	%	Resp.	%	Resp.	%	Resp.	%	Resp.	%	Resp.	%
Total para base de poder perícia	0	0%	0	0%	0	0%	2	13%	14	88%	16	100%
Total para base de poder legítimo	2	13%	3	19%	0	0%	1	6%	10	63%	16	100%
Total para base de poder de recompensa	10	63%	1	6%	0	0%	0	0%	5	31%	16	100%
Total para base de poder de coerção	0	0%	9	75%	0	0%	0	0%	3	25%	12	100%
Total para o poder da estrutura	12	20%	13	22%	0	0%	3	5%	32	53%	60	100%

Referência: LION, 2015.

Considera-se, quanto ao poder da estrutura/cargo, que nas Bibliotecas Universitárias do Sistema de Bibliotecas da Universidade Federal da Bahia o poder organizacional é fortemente pautado na estrutura (autoridade do cargo), pois a zona de concordância atingiu 58% (somando-se 53% dos entrevistados que concordam totalmente com 5% que concordam parcialmente). Porém não se deve deixar de observar que os dados indicam uma zona de discordância de 42% (somando-se 20% dos bibliotecários-chefe entrevistados que discordam totalmente com 22% que discordam parcialmente), o que significa uma tendência a distribuição do poder da autoridade do cargo esculpido no organograma.

Em relação a identificação das bases estruturais de poder organizacional, considera-se: a) a base estrutural de poder legítimo, baseado no poder legítimo, ficou em segundo lugar; b) a base estrutural de poder de coerção, baseado no poder de coerção, é o poder menos utilizado nas BUs do SIBI/UFBA; c) a base estrutural de poder de recompensa, baseado no poder de recompensa, é a terceira forma de exercer poder nas BUs pesquisadas; e d) a base estrutural de poder de perícia, baseado no poder de perícia, é o poder dominante nas BUs do SIBI/UFBA.

Tabela 4 - Distribuição percentual dos totais para o Poder da Estrutura Organizacional das coordenações do Arquivo Público do Estado da Bahia – APEB

Bases Estruturais de Poder Organizacional	Discordo totalmente		Discordo parcialmente		Não tenho opinião formada/Não sei		Concordo parcialmente		Concordo totalmente		Total	
	Resp.	%	Resp.	%	Resp.	%	Resp.	%	Resp.	%	Resp.	%
Total para base de poder perícia	3	20%	1	7%	1	7%	2	13%	8	53%	15	100%
Total para base de poder recompensa	0	0%	1	6%	10	63%	5	31%	0	0%	16	100%
Total para base de poder de legítimo	4	25%	0	0%	8	50%	4	25%	0	0%	16	100%
Total para base de poder de coerção	3	25%	0	0%	8	67%	1	8%	0	0%	12	100%
Total para o poder da estrutura	10	17%	2	3%	27	46%	12	20%	8	14%	59	100%

Referência: LION, 2015.

Considera-se, quanto ao poder da estrutura/cargo/autoridade, que nas coordenações do Arquivo Público do Estado da Bahia o poder organizacional é relevantemente pautado na estrutura (autoridade do cargo), pois a zona de concordância atingiu 34% (somando-se 14% que concordam totalmente com 20% que concordam parcialmente). Porém não se deve deixar de observar que os dados indicam uma zona de não posicionamento de 46% (ou seja, 46% dos coordenadores não se posicionaram ou não sabem), o que significa uma baixa tendência a distribuição do poder da autoridade do cargo esculpido no organograma e isso é confirmado pelo baixo índice de pessoas que discordam totalmente (apenas 17%) quanto a centralização do poder.

Em relação a identificação das bases estruturais de poder organizacional, considera-se: a) a base estrutural de poder de perícia, baseado no poder de perícia, é o poder dominante nas coordenações do APEB; b) a base estrutural de poder de recompensa, baseado no poder de recompensa, é a segunda forma de exercer poder nas coordenações pesquisadas; c) a base estrutural de poder legítimo, baseado no poder legítimo, ficou em terceiro lugar; b) a base estrutural de poder de coerção, baseado no poder de coerção, é o poder menos utilizado nas coordenações do APEB.

CONCLUSÕES

As análises conclusivas da seção anterior só são validas para as BUs do SIBI/UFBA, ou seja, são validas apenas para a amostra. A partir da análise do poder organizacional nas Bibliotecas Universitárias (BUs) do Sistema de Bibliotecas da Universidade Federal da Bahia (SIBI/UFBA) sob a perspectiva política da Administração, que é o objetivo geral da tese aplicado às BUs do SIBI/UFBA, conclui-se em relação as bases de poder que: há nas bibliotecas universitárias pesquisadas fracos indícios no sentido de ser o poder distribuído pelas equipes de trabalho, que conduz a uma reprodução e uso menos equitativa desse poder, aproximando-se de um poder bastante concentrado na autoridade do cargo.

Quanto ao Arquivo Público do Estado da Bahia, as análises conclusivas também só são validas para as Coordenações do APEB, ou seja, são validas apenas para a amostra. A partir da análise do poder

organizacional sob a perspectiva política da Administração, que é o objetivo geral desta tese aplicado ao APEB, conclui-se em relação as bases de poder que: há nas coordenações pesquisadas uma baixa tendência no sentido de ser o poder distribuído pelas equipes de trabalho, que conduz a uma reprodução e uso menos equitativa desse poder, aproximando-se de um poder relevantemente concentrado na autoridade do cargo.

A autoridade do cargo de arquivista-chefe ou bibliotecário-chefe, ou profissional da informação gestor, é o elemento objetivo da estrutura hierárquica tanto no SIBI/UFBA quanto no APEB. As nuances do exercício da autoridade do cargo é algo que, por estar materializado no plano de cargos dos servidores da UFBA ou Governo do Estado da Bahia, nos seus respectivos regimento internos e normativas, pode ser racionalmente apropriado pelas pessoas das equipes de trabalho como base de poder e isso pode gerar disputas pelo acesso ao cargo, que assume a representação, consciente ou inconsciente, de acesso a autoridade e consequentemente ao poder.

A base de poder dominante no SIBI/UFBA e no APEB é o poder de perícia e a literatura indica que a maioria dos profissionais da informação ocupam cargos de gestão, em algum nível, e de processo técnico. Este poder de perícia sugere a presença de um Sistema de Responsabilidade, que é o componente antecessor ao sistema de autoridade no delineamento da estrutura organizacional em bibliotecas e arquivos e está relacionado a atribuição de encargos inerentes àquilo que é dimensionado para que o colaborador execute as funções que ocupa.

REFERÊNCIAS BIBLIOGRÁFICAS

Barbalho, C.R.S. (2012). Estrutura organizacional de bibliotecas universitárias: subsidio para sua composição. *Seminário nacional das bibliotecas universitárias – SNBU 2012*, Gramado, Brasil.

Campos, G.N.; Baptista, S.G. (2008). Perfil dos bibliotecários das bibliotecas de instituições de ensino superior privadas do Distrito Federal e as expectativas dos empregadores. *Encontro nacional de pesquisa em ciência da informação – ENANCIB 2008*, São Paulo, Brasil.

Clegg, S. R.; Hardy, C. (2001). Alguns ousam chamá-lo de poder. In Clegg, S. R.; Hardy, C.; Nord, W.R. *Handbook de estudos organizacionais: Reflexões e novas direções*, (2), (pp. 260-289), São Paulo, Brasil: Atlas.

Fontanella, B.J.B.; Ricas, J.; Turano, E.R. (2008). Amostragem por saturação em pesquisas qualitativas em saúde: contribuições teóricas. *Cad. Saúde Pública*, 24 (1), 17-27, Rio de Janeiro, Brasil.

Lion, S.E.K. (2010). A eficiência estratégica do Sistema de Bibliotecas da Universidade Federal da Bahia. *Dissertação, mestrado profissional, Escola de Administração, Universidade Federal da Bahia*. Retrieved from <https://repositorio.ufba.br/ri/handle/ri/7651>.

Lion, S.E.K. (2015). A perspectiva política: um estudo sobre o poder organizacional em unidades de informação. *Tese, Doutorado em Ciência da Informação, Instituto de Ciência da Informação, Universidade Federal da Bahia*. Retrieved from <https://repositorio.ufba.br/ri/handle/ri/19542>.

Martins, M.C.F. (2008). Bases do poder organizacional. In SIQUEIRA, M. M. M. *Medida do comportamento organizacional: Ferramentas de Diagnósticas e de Gestão* (pp. 21-26). Porto Alegre: Artmed.

Motta, P.R. (6 reimpressão). (2007). *Transformação organizacional: a teoria e a prática de inovar*. Rio de Janeiro: Quality Mark.

Oliveira, J.L.R. (2008). Profissionais de Sistemas de Informação: descrição da profissão e notas sobre as dificuldades de uma análise comparativa com os Profissionais da Informação. *Encontro nacional de pesquisa em ciência da informação – ENANCIB 2008*, São Paulo, Brasil.



GESTÃO POR COMPETÊNCIA EM BIBLIOTECAS UNIVERSITÁRIAS

Andreza Nadja Freitas Serafim¹, Izabel França de Lima² Giulianne Monteiro Pereira³,
Fernanda Mirelle A. Silva⁴, Bernardina Maria Juvenal Freire Oliveira⁵

¹Programa de Pós-Graduação em Gestão nas Organizações Aprendentes da Universidade Federal da Paraíba, 0000-0001-9924-546x, andrezanfs@yahoo.com.br

²Programa de Pós-Graduação em Gestão nas Organizações Aprendentes da Universidade Federal da Paraíba, 0000-0003-2701-5432, belbib@gmail.com

³Universidade Estadual da Paraíba, 0000-0002-8400-0715, giulianne.monteiro@gmail.com

⁴Departamento de Ciência da Informação da Universidade Federal da Paraíba, 0000-0003-1814-3600, fmirelle@gmail.com

⁵Programa de Pós-Graduação em Ciência da Informação da Universidade Federal da Paraíba, 0000-0002-6836-3102, bernardinafreire@gmail.com

RESUMO Objetiva verificar na literatura as ações de gestão por competência existentes que são aplicadas em bibliotecas universitárias. O conceito de gestão por competência utilizado, baseia-se na matriz conceitual estabelecida por Le Boterf (2003) e complementada por Dutra (2004), que apontam a competência como capacidade de mobilizar o conhecimento, as habilidades e aptidões individuais em diferentes contextos. Dessa forma, não deve-se vincular a competência ao cargo, mas sim ao colaborador que terá que compreender as demandas do contexto sobre ela e saber mobilizar seu repertório para atendê-las de forma adequada (Dutra, 2004). Diante dessa perspectiva, verifica-se que a gestão por competência pode ser aplicada as Bibliotecas Universitárias (BUs), tendo em vista a necessidade de adaptação em seguir essa convergência, buscando adotar uma gestão mais flexível, estratégica e empreendedora, e também instigar o estímulo e criação de mecanismos que favoreçam o desenvolvimento profissional, promovendo a aprendizagem e proporcionando aos colaboradores oportunidades reais de crescimento profissional. Em termos metodológicos trata-se de uma pesquisa exploratória, delimitada pelas abordagens qualitativa e quantitativa, que de acordo com seus objetivos adota procedimentos respaldados pela tipologia documental e descritiva. Para o alcance dos objetivos, as pesquisas documental e bibliográfica foram utilizadas como instrumento de coleta de dados. Os resultados apresentados, com base na literatura, contribuem para o avanço dos estudos sobre as competências requeridas aos bibliotecários para atuar em bibliotecas universitárias e possibilita sugerir ações para promoção e aperfeiçoamento das competências dos bibliotecários.

PALAVRAS-CHAVE *Gestão por competência. Tecnologia da Informação e Comunicação. Inovação profissional. Bibliotecas Universitárias.*

ABSTRACT It aims to verify in the literature the existing competency management actions that are applied in university libraries. The concept of competence management is based on the conceptual matrix established by Le Boterf (2003) and complemented by Dutra (2004), which points out competence as the capacity to mobilize knowledge, individual abilities and skills in different contexts. In this way, the competence should not be linked to the position, but rather to the collaborator, who will have to understand the demands of the context on it and to know how to mobilize his repertoire to serve them adequately (Dutra, 2004). From this perspective, it can be verified that management by competence can be applied to University Libraries (BUs), considering the necessary adaption to follow this convergence, seeking to adopt a more flexible, strategic and entrepreneurial management, and instigate the stimulus and creation of mechanisms that foster

professional development, promoting learning and providing employees with real opportunities for professional growth. In methodological terms, this is an exploratory research, delimited by the qualitative and quantitative approaches, which according to its objectives adopts procedures backed by the documentary and descriptive typology. In order to reach the objectives, the documentary and bibliographical researches were used as instrument of data collection. The results presented, based on the literature, contribute to the advancement of studies on the librarians skills required to work in university libraries and makes it possible to suggest actions to promote and improve the skills of librarians.

KEYWORDS *Management by competence. Technology of Information and Communication. Professional innovation. University Libraries.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Na contemporaneidade no âmbito das organizações, torna-se necessário discutir a Gestão por Competência (GC) como um instrumento que facilita a gestão do capital humano, tendo em vista que o capital humano agrega valor a organização, através da efetivação das competências organizacionais. Dessa forma, analisar as competências existentes em uma organização é um fator primordial para alocar bem os colaboradores e aproveitar melhor suas competências, habilidades e atitudes.

A gestão por competência caracteriza-se por ser uma das atuais técnicas que auxilia no processo de avaliação e orienta as ações de desenvolvimento profissional. O conceito apresentou-se de forma estruturada, pela primeira vez, em 1973 por David McClelland. Essa abordagem foi disseminada no Brasil por volta dos anos 80 e no início dos anos 90 associada ao perfil de conhecimentos (Dutra, 2006). Diversas definições são apresentadas para o conceito de competência, autores como Woodruffe (1991), Boyatsis (1982), Le Boterf (1995), Zarifian (2012) e Dutra (2004) desenvolveram conceitos de competências aplicados a gestão de pessoas. A definição de competência é múltipla, permitindo ser interpretada de diversas maneiras, conforme verifica-se na literatura, no entanto, o conceito de competência explorado nesta pesquisa é apresentado por Le Boterf (2003, p. 25):

A competência não é um estado ou um conhecimento que se tem, nem é resultado de treinamento, na verdade competência é mobilizar conhecimentos e experiências para atender as demandas e exigências de determinado contexto, marcado geralmente pelas relações de trabalho, cultura da empresa, imprevistos, limitações de tempo e de recursos.

Nesse sentido, a competência consiste em mobilizar o conhecimento, as habilidades e aptidões individuais em diferentes contextos. Dessa forma, não deve-se vincular a competência ao cargo, mas sim ao colaborador que terá que compreender as demandas do contexto sobre ela e saber mobilizar seu repertório para atendê-las de forma adequada (Dutra, 2004).

A gestão por competências diferencia-se do modelo tradicional de gestão que incorporava um ponto de vista restrito, que se refere à função do indivíduo nas organizações. Pois na sociedade atual que se vivencia o conhecimento, a informação e as tecnologias como força motriz de sua economia, a competência torna-se um recurso indispensável promovendo o potencial criativo e inovador dos trabalhadores.

Diante da perspectiva apresentada, verifica-se que a gestão por competência pode ser aplicada nas Bibliotecas Universitárias (BUs), tendo em vista a necessidade de adaptação desse tipo de organização em seguir essa convergência, procurando adotar uma gestão mais flexível, estratégica e empreendedora. E tem como desafio o estímulo e criação de mecanismos que favoreçam o desenvolvimento profissional, promovendo a aprendizagem e proporcionando aos colaboradores oportunidades reais de crescimento profissional.

Conforme a 5ª Lei de Ranganathan (1931), a “[...] Biblioteca é um organismo em crescimento”, e, portanto, deve acompanhar as mudanças que ocorre a todo momento na Sociedade, buscando formas de cumprir sua missão institucional. O bibliotecário, inserido nesse contexto e na sua função de administrador, deve prevê, controlar e planejar a melhor maneira que esse crescimento ocorre. Nesse sentido, a Gestão por competência se apresenta como um recurso para que o Gestor da Biblioteca Universitária possa acompanhar de maneira eficiente o desenvolvimento dessa Unidade de Informação (UI).

A motivação para o desenvolvimento deste estudo surgem a partir de percepções proporcionadas pela prática profissional, pois como bibliotecária gestora de uma Biblioteca Universitária, visualizo a necessidade da BU em acompanhar as mudanças que emergem constantemente, como também pela necessidade que a referida UI possui em adotar as novas práticas de Gestão de pessoas com vistas a melhorar os seus processos internos e a oferta de seus produtos e serviços.

Embora a literatura sobre Gestão de Pessoas e Gestão por Competência seja extensa e bastante rica, este trabalho justifica-se ainda pela escassa quantidade de literatura específica que trate, especificamente, da Gestão por Competência aplicada em Bibliotecas Universitárias.

Assim, a relevância desta pesquisa se dá em vista da urgente necessidade de investigar as práticas, serviços e produtos das BUs a partir do modelo de Gestão por Competências, visando atender dinâmicas recorrentes da gestão de pessoas, alinhando com as perspectivas da GC.

A partir das problematizações e justificativas apresentadas, definiu-se como objetivo desta pesquisa: discutir sobre as ações de Gestão por Competência encontradas através da literatura que existem e que são aplicadas nas Bibliotecas Universitárias.

REFERENCIAL TEÓRICO

O Sistema Nacional de Bibliotecas Públicas Brasileiras (2006) descreve a biblioteca universitária como uma Instituição que:

Tem por objetivo apoiar as atividades de ensino, pesquisa e extensão por meio de seu acervo e dos seus serviços. Atende alunos, professores, pesquisadores e comunidade acadêmica em geral. É vinculada a uma unidade de ensino superior, podendo ser uma instituição pública ou privada. A Biblioteca Universitária dá continuidade ao trabalho iniciado pela Biblioteca Escolar.

Esta descrição é amplamente aceita, pois ele inclui os tipos de usuários que a biblioteca universitária atende e especifica seus objetivos. Corroborando nessa formação conceitual da Biblioteca Universitária, Anzolin e Sermann (2006, p.7) afirmam que:

A biblioteca universitária pode ser definida como aquela a que atua em Instituições de ensino superior, como centros universitários, universidades e faculdades. Tem por finalidade dar suporte informacional, complementando as atividades curriculares dos cursos, oferecendo recursos para facilitar a pesquisa científica. Sua missão é prover informação para o ensino, a pesquisa e a extensão, de acordo como a política, projeto pedagógico e programas da universidade a qual está inserida. As diretrizes do ensino superior reforçam a necessidade de participação ativa das bibliotecas em programas de ensino, pesquisa e extensão.

Segundo Nunes e Carvalho (2016, p. 175), a evolução das bibliotecas universitárias está atrelada ao surgimento das universidades, pois

Estão ligadas historicamente ao desenvolvimento humano e social, e neste sentido também exercem uma importante tarefa para a mediação da informação, acompanhando não apenas a evolução da produção escrita e da circulação do conhecimento, mas também a evolução tecnológica que favorece o processo comunicacional.

As bibliotecas universitárias perpassam por várias transformações no processo de gestão onde muitas delas são decorrentes das diversas mudanças e novas necessidades da Sociedade e, alinhar as BUs à Gestão por Competência (GC) se mostra como uma forma de gerir melhor os recursos humanos. A GC caracteriza-se por ser uma das atuais técnicas que auxilia no processo de avaliação e orienta as ações de desenvolvimento profissional. Existem várias sobre o conceito de competência tanto no âmbito organizacional, quanto no acadêmico. Neste estudo, optou-se por mencionar autores que relacionam o conceito ao ambiente profissional, tendo em vista o foco da pesquisa.

De acordo com Dutra (2004, p. 10), o termo competência corresponde a “um conjunto de conhecimentos, habilidades, atitudes, posturas e ações integradas utilizadas pelas pessoas para obtenção de resultados”. Corroborando com essa definição, Lustri, Miura e Takahashi (2007, p. 187, tradução nossa) afirmam que:

Um competente profissional demonstra uma compreensão prática baseada no conhecimento adquirido através das experiências e esse conhecimento é alterado e aumentado à medida que ele é capaz de mobilizar uma rede de autores e fazer com que compartilhem ações e responsabilidades em torno das mesmas situações.

O conceito de competência explorado nesta pesquisa é apresentado por Valentim (2002, p. 122) como um “[...] conjunto de habilidades, destrezas, atitudes e de conhecimentos teórico-práticos necessários para cumprir uma função especializada de um modo socialmente reconhecível e aceitável”. O conceito propõe que a competência está na capacidade do indivíduo saber articular a sua gama de conhecimentos e experiências para executar uma atividade de forma eficiente.

Essas abordagens conceituais passam por questões como: processo de formação, comportamento, aptidão, aprendizagem, ação, valores, interação, perspectiva dinâmica, resultado e desempenho. Mesmo havendo uma variedade de ênfases, percebe-se uma convergência no sentido da competência buscar um desenvolvimento pessoal e profissional que leva a um melhor resultado. Trata-se de um conceito em construção, o que não significa que seja um conceito incompleto, mas que, pela complexidade do tema, conduz à ideia de um conceito dinâmico (Carbone, 2006).

A Gestão por Competência (GC) favorece a alocação dos recursos humanos de forma que os colaboradores sejam alocados de acordo com suas competências, viabilizando uma melhor execução das atividades. Segundo Cruz et al. (2013, p. 6):

A gestão por competências trata-se de um processo contínuo que toma como referência a estratégia da instituição e direciona as ações de recrutamento e seleção, treinamento, gestão de carreira, avaliação de desempenho e formalização de alianças estratégicas. Captação e desenvolvimento de competências necessárias para atingir assim os objetivos, de tal forma que, gera um impacto positivo sobre o desempenho organizacional.

A gestão por competência propõe-se a “[...] gerenciar o *gap* ou lacunas de competências eventualmente existente na organização, procurando eliminá-lo ou minimizá-lo. [...]” (Carbone et al., 2006, p. 69).

Além disso, a identificação das lacunas das competências individuais existentes na organização possibilitará a promoção de práticas que visem capacitar e ampliar os conhecimentos dos colaboradores, favorecendo o desenvolvimento das competências necessárias para a organização.

Percebe-se que a gestão por competência evidencia uma nova perspectiva para a área de gestão de pessoas. Por isso, alinhar as bibliotecas universitárias a esse tipo de gestão potencializa o uso do capital intelectual existentes nessas organizações. E viabiliza o melhor aproveitamento do conhecimento na aplicação prática das atividades desenvolvidas nesse espaço, favorecendo a eficiência e a eficácia dos objetivos a serem alcançados pelas organizações.

Segundo Rodrigues e Muniz (2010, p. 7), a GC nas bibliotecas universitárias viabiliza:

A reestruturação dos serviços, possibilitam a implementação de mudanças, sempre que necessário, propiciam ao Bibliotecário ou Profissional da Informação (PI) o planejamento de sua atuação, possibilitando as inovações essenciais a unidade de informação para que ela possa atender as expectativas e necessidades do usuário. A Gestão por Competências propicia ao PI aumentar sua capacidade de valorizar as diferenças e promover mudanças significativas nas Unidades de Informação para que ela funcione de forma eficaz.

Dessa forma, verifica-se que é importante que o bibliotecário gestor busque um modelo de gerenciamento que possibilite o desenvolvimento das competências dos colaboradores, para favorecer o desenvolvimento da biblioteca. A gestão por competência é uma estratégia que auxiliará na tomada de decisão e no aperfeiçoamento das práticas administrativas.

As bibliotecas universitárias devem se adequar para se manter no mercado competitivo. O capital humano é um fator de suma importância, uma vez que é responsável por executar as estratégias dentro das organizações. Dessa forma, a atuação dos colaboradores tornou-se um fator de competitividade, haja vista que o desempenho das instituições decorrem de sua capacidade de desenvolver competências e de agregá-las em torno de seus objetivos.

METODOLOGIA

A escolha da metodologia foi realizada de acordo com o fenômeno estudado, portanto os procedimentos metodológicos escolhidos são os mais adequados para essa pesquisa. Para Almeida (2015, p. 32), a metodologia adequada é “aquela que viabiliza a trilha investigativa objetivada, incluindo as bases teóricas fundamentais, sem restringir a criatividade científica e sem se distanciar da realidade estudada”. Esta seção encontra-se subdividida nas seguintes partes: caracterização da pesquisa, campo e população da pesquisa, técnicas para coleta dos dados e procedimentos de análise e interpretação dos dados.

A metodologia empregada para o desenvolvimento da pesquisa está baseada na revisão de literatura acerca do tema em questão. Através de pesquisas bibliográficas, eletrônicas, anais de eventos e outras fontes de informação pertinentes ao assunto.

Em conformidade com os objetivos esta pesquisa caracteriza-se como exploratória uma vez que se faz viável um estudo específico com enfoque nas ações. Segundo Gil (2010, p. 27), “as pesquisas exploratórias têm como propósito proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a construir hipóteses.”

Ainda de acordo com os objetivos, a pesquisa também se caracteriza como documental e descritiva, pois se propõe a conhecer a natureza do fenômeno estudado, sua composição e processos que o constituem ou nele se realizam (Rudio, 2001).

RESULTADOS

As bibliotecas universitárias vêm sofrendo mudanças no processo de gestão, sendo públicas ou privadas, devido a necessidade de se adequar as exigências do mercado, que tem como um de seus principais valores o capital humano.

Por sua vez, essa nova concepção baseia-se na tríade conhecida como CHA (Quadro 1) que significa Conhecimentos, Habilidades e Atitudes Proativas que os colaboradores das instituições devem ter diante da realização de suas tarefas e que a própria organização deve seguir no seu dimensionamento estratégico na área de recursos humanos.

Quadro 1 - As três dimensões da competência

Conhecimentos (Saber)	Habilidades (Saber fazer)	Atitudes Proativas (Saber ser)
Conhecimentos técnicos, escolaridade, cursos, especializações, etc.	Experiência, capacidades do indivíduo.	Ter atitudes compatíveis para atingir eficácia em relação aos conhecimentos e habilidades adquiridas.

Fonte: Ruzzarin (2002).

As três dimensões da competência apresentadas representa o uso adequado pelo colaborador dos seus atributos de competência, ou seja, conhecimentos, habilidades e atitudes que devem ser compatíveis com a função que ele desempenha.

Nesse contexto, percebe-se que a gestão por competência evidencia uma nova perspectiva para a área de gestão de pessoas. Por isso, alinhar as bibliotecas universitárias a esse tipo de gestão potencializa o uso do capital intelectual existentes nessas organizações e viabiliza o melhor aproveitamento do conhecimento na aplicação prática das atividades desenvolvidas nesse espaço, favorecendo a eficiência e a eficácia dos objetivos a serem alcançados pelas organizações.

Dessa forma, verifica-se que é importante que o bibliotecário gestor busque um modelo de gerenciamento que possibilite o desenvolvimento dos conhecimentos, habilidades e atitudes dos colaboradores, esses fatores favoreceram o desenvolvimento da biblioteca, pois a CG é uma estratégia que auxiliará na tomada de decisão e no aperfeiçoamento das práticas administrativas nas BUs.

Nessa perspectiva, o primeiro passo para a implantação da Gestão por Competência em uma organização é a realização do mapeamento das competências, que, Amaral (2006, p. 68) define como um “processo consiste na elucidação das competências para os cargos da organização, por meio do levantamento dos indicadores de competência, os quais são reunidos por semelhança e convertidos nas competências”.

O processo de mapeamento propõe a descrição das competências dos cargos da organização, através do levantamento dos atributos de competência (conhecimento, habilidade e atitude), os quais são reunidos por semelhança e convertidos em competências. Existem várias abordagens para o mapeamento de competências, dentre as quais se destacam as aplicações de entrevistas aos atuais trabalhadores, identificando situações e desafios que eles já enfrentaram, observação (análise do trabalho, de seus resultados e do que é necessário para alcançá-los), além de simulações realizando suposições sobre o ambiente e o futuro do trabalho.

A gestão por competência, segundo Dutra (2004), permite que vários benefícios sejam alcançados tanto para a instituição quanto para o colaborador esses benefícios podem ser aplicados as bibliotecas universitárias:

Quadro 2 – Benefícios da implantação da Gestão por competência

PARA A INSTITUIÇÃO	PARA O COLABORADOR
Simpleza, flexibilidade e transparência;	Horizontes profissionais claros com critérios de acesso definidos;
Otimização de recursos e da massa salarial;	Remuneração compatível com a complexidade das atribuições e das responsabilidades e com mercado;
Direcionamento e otimização dos investimentos no desenvolvimento profissional;	Estímulo ao autodesenvolvimento e à ampliação do espaço de atuação;
Capacidade de atração, retenção e potencialização de talentos;	Condições claras e objetivas para a mobilidade entre as carreiras abrangidas pelo modelo de gestão de pessoas por competências;
Flexibilidade do modelo para adaptar-se às mudanças de estrutura, organização do trabalho e tecnologia	
Equilíbrio entre remuneração e agregação de valor.	

Fonte: Dutra (2001).

Diante dos benefícios apresentados, percebe-se que a GC favorece tanto a organização, quanto o colaborador e que, além disso, permite um melhor planejamento das atividades, pois, através do mapeamento das competências, a biblioteca poderá identificar quais as competências dos seus colaboradores e quais tipos de cursos de capacitação esses precisam para atender a demanda dos serviços oferecidos na BUs.

Desse modo, traçar um perfil das competências necessárias para o bibliotecário atuar em bibliotecas universitárias, possibilita identificação dos conhecimentos, habilidades e atitudes apontados pela literatura que deve fazer parte das competências dos bibliotecários.

As bibliotecas universitárias são espaços que favorecem o processo de aprendizagem dos seus usuários, pois

[...] atende às necessidades de informação dos corpos docente, discente e administrativo, tanto para apoiar as atividades de ensino, quanto de pesquisa e extensão. Pode ser uma única biblioteca ou várias organizadas como sistema ou rede (Cunha & Cavalcanti, 2008, p. 53).

Este conceito mostra o quanto a BU é um espaço dinâmico e que o bibliotecário que atua nesse espaço necessita de competências específicas para prestar um atendimento de qualidade. A Classificação Brasileira de Ocupações (CBO) apresenta bibliotecários como profissionais que

Disponibilizam informação em qualquer suporte; gerenciam unidades como bibliotecas, centros de documentação, centros de informação e correlatos, além de redes e sistemas de informação. Tratam tecnicamente e desenvolvem recursos informacionais; disseminam informação com o objetivo de facilitar o acesso e geração do conhecimento; desenvolvem estudos e pesquisas; realizam difusão cultural; desenvolvem ações educativas. Podem prestar serviços de assessoria e consultoria (Brasil, 2007).

Para identificar as competências necessárias para atuação do bibliotecário nas BUs é preciso apontar os principais serviços de uma biblioteca universitária, tendo em vista que esses podem variar de acordo com as necessidades de cada Instituição de Ensino Superior (IES). Vários autores indicam quais são os serviços básicos prestados nesse tipo de unidade de informação, Macedo (1990), Amboni, (2002), Borges (2007), Freitas, Bolsanello e Viana (2008), conforme mostra o quadro a seguir:

Quadro 3 – Serviços básicos prestados em Bibliotecas Universitárias

SERVIÇOS DE INFORMAÇÃO	
SERVIÇOS	DESCRIÇÃO
CAPACITAÇÃO	Treinamentos focados na utilização do sistema, bases de dados, normalização, direito autoral, inclusão digital, educação ao usuário, entre outros;
COMUTAÇÃO BIBLIOGRÁFICA	Permite a obtenção, mediante pagamento antecipado, de cópias de documentos técnico-científicos disponíveis nos acervos de bibliotecas conveniadas.
LEVANTAMENTO BIBLIOGRÁFICO	Recuperação de materiais bibliográficos específicos para a necessidade informacional.
NORMALIZAÇÃO BIBLIOGRÁFICA	Visa padronizar os documentos de acordo com uma determinada norma
PROCESSO DE REFERÊNCIA/ASSISTÊNCIA INFORMACIONAL	Entrevista com o usuário para identificar e buscar as necessidades informacionais, resposta a informações bibliográficas factuais, auxílio na busca e na recuperação de materiais bibliográficos (pessoal, <i>online</i> ou telefone).

Fonte: Macedo (1990), Amboni, (2002), Borges (2007) e Freitas, Bolsanello e Viana (2008).

Além desses serviços apresentados, existem outros que variam de acordo com a instituição, como por exemplo elaboração de ficha catalográfica, emissão de nada constas, atividades culturais e entre outras. Para atender a demanda desses serviços os bibliotecários devem ter conhecimentos, habilidades e atitudes que favoreçam o desenvolvimento dessas atividades.

De acordo com Vendrel e Miranda (1999), os profissionais da informação devem adquirir seis tipos de competências necessárias para atuação em unidades de informação:

- a) **Competência intelectual** – referente aos processos cognitivos internos necessários para simbolizar e representar ideias, imagens, conceitos ou outras abstrações;
- b) **Competência prática** – referente a um *saber-fazer*, ou seja, é saber articular as tomadas imediatas de decisão, colocando-as em ação;
- c) **Competência interativa** – implica na capacidade dos sujeitos de participar como membro de um grupo de referência como a família, os pares etc;
- d) **Competência social** – envolve no uso do consenso, no exercício de aceitação da liderança, da capacidade de ensinar e aprender com os outros;
- e) **Competências éticas** – implica no discernimento entre o bem e o mal, no amplo e complexo espaço em que os indivíduos estão em processo de inter-relacionamento;
- f) **Competência estética** – que aludem a capacidade de distinguir entre o que há de bom e de ruim e, no plano de valores, entre o belo e o feio.

Do ponto de vista de Vendrel e Miranda (1999), as competências estão inseridas em um ciclo de relevância que, ao longo do tempo sofre alterações sendo possível classificá-las em quatro categorias que são:

- a) **Emergentes** – surgem com mudanças significativas que ocorrem no mercado de trabalho e estão relacionadas, por exemplo, ao domínio de um idioma, ao eficaz uso de novas tecnologias, entre outros fatores;
- b) **Declinantes** – estão relacionadas às mudanças de tecnologias ou do modo de operacionalização das atividades da organização;
- c) **Estáveis ou essenciais** – são aquelas que permanecem relevantes ao longo do tempo;
- d) **Transitórias** – embora essenciais em momentos críticos de transição, podem ser necessárias por um determinado momento.

A Gestão por Competência segundo Carbone et al. (2006, p. 69) permite “[...] gerenciar o *gap* ou lacuna de competências eventualmente existente na organização, procurando eliminá-lo ou minimizá-lo. [...]” Por isso, esse tipo de gestão deve ser pensado no âmbito das bibliotecas universitárias.

CONCLUSÕES

Diante do exposto, percebe-se que a literatura referente ao tema ainda é pouco estudada, precisando de mais estudos aprofundados. Os resultados apresentados, com base na literatura, contribuem para o avanço dos estudos sobre as competências requeridas aos bibliotecários para atuar em bibliotecas universitárias e possibilita sugerir ações para promoção e aperfeiçoamento das competências dos bibliotecários.

Considera-se que o profissional da informação que atua em Bibliotecas Universitárias pode se adequar a essa nova perspectiva da gestão por competência, uma vez que o capital humano existente nas bibliotecas é um fator preponderante, tendo em vista o papel dos bibliotecários, dos auxiliares e usuários, que tem potencial para a criação de ativos intangíveis como o conhecimento.

Por fim, pode-se afirmar que a característica interdisciplinar da Ciência da Informação proporciona novos horizontes, possibilidades e olhares, e com isso possibilidades para se (re)pensar a atuação do profissional bibliotecários, como foi apresentado nesse trabalho.

REFERÊNCIAS BIBLIOGRÁFICAS

Almeida, J. L. S. (2015). *A biblioteca como organização aprendente: o desenvolvimento de competências em informação no Instituto Federal de Ciência e Tecnologia da Paraíba*. (Dissertação de mestrado). Centro de Educação, Universidade Federal da Paraíba, João Pessoa, PB, Brasil.

Amaral, R. M. (2006). *Desenvolvimento e aplicação de um método para o mapeamento de competências em inteligência competitiva*. (Dissertação de mestrado). Engenharia de Produção, Universidade Federal de São Carlos, São Carlos, SP, Brasil. 2006. Recuperado de <https://repositorio.ufscar.br/bitstream/handle/ufscar/3807/832.pdf?sequence=1&isAllowed=y>

Amboni, N. F. (2002). *Qualidade em serviços: dimensões para orientação e avaliação das bibliotecas universitárias federais brasileiras*. (Tese de doutorado). Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, SC, Brasil. Recuperado de <https://repositorio.ufsc.br/bitstream/handle/123456789/83805/188852.pdf?sequence=1>

Borges, M. E. N. (2007). O essencial para a gestão de serviços e produtos de informação. *Revista Digital de Biblioteconomia e Ciência da Informação*, Campinas, v. 5, n. 1, p. 115-128, jul./dez. Recuperado de <http://www.sbu.unicamp.br/seer/ojs/index.php/rbci/index>

BRASIL. (2007). *Classificação Brasileira de Ocupações 2612: profissionais da informação*. Brasília. 2007. Recuperado de <http://www.mtecbo.gov.br/cbosite/pages/pesquisas/BuscaPorTituloResultado.jsf>

Carbone, P. P. et al. (2006). *Gestão por competências e gestão do conhecimento*. 2. ed. Rio de Janeiro: FGV.

Cunha, M. B. & Cavalcanti, C. R. (2008). *Dicionário de biblioteconomia e arquivologia*. Brasília: Briquet de Lemos.

Dutra, J. S. (2004). *Competências: conceitos e instrumentos para a gestão de pessoas na empresa moderna*. São Paulo: Atlas.

Dutra, J. S. (2013). *Competências: conceitos e instrumentos para a gestão de pessoas na empresa moderna*. São Paulo: Atlas.

Dutra, J. S. et al. (2006). Absorção do conceito de competência em gestão de pessoas: a percepção dos profissionais e as orientações adotadas pelas empresas. *Anais do Encontro da ANPAD*, 30, Salvador, Brasil. Recuperado de <http://www.anpad.org.br/enanpad/2006/dwn/enanpad2006-gpra-2578.pdf>

Dutra, J. S. et al. *Gestão por competências: um modelo avançado para o gerenciamento de pessoas*. São Paulo: Gente, 2001.

Dutra, J. S., Hipolito, J. M. A. & Silva, C. M. (2000). Gestão de pessoas por competências: o caso de uma empresa do setor de telecomunicações. *Revista de Administração Contemporânea*, v.1, jan./mar. Recuperado de <http://www.scielo.br/pdf/rac/v4n1/v4n1a09>

Dutra, J. S., Fleury, M. T. L. & Ruas, R. (2013). *Competências: conceitos, métodos e experiências*. São Paulo: Atlas.

Freitas, A. L. P., Bolsanello, F. M. C. & Viana, N. R. N. G. (2008). Avaliação da qualidade de serviços de uma biblioteca universitária: um estudo de caso utilizando o modelo Servqual. *Ciência da Informação*, Brasília, v. 37, n. 3, p. 88-102, set./dez. Recuperado de <http://www.scielo.br/pdf/ci/v37n3/v37n3a07.pdf>

Gil, A. C. (2010). *Como elaborar projetos de pesquisa*. 5. ed. São Paulo: Atlas.

Le Boterf, G. (2003). *Desenvolvendo a competência dos profissionais*. 3. ed. Porto Alegre: Artmed.

Macedo, N. D. (1990). Princípios e reflexões sobre o serviço de referência e informação. *Revista Brasileira de Biblioteconomia e Documentação*, São Paulo, n. 1, v. 23, p. 9-37, jan./dez. Recuperado de <http://basessibi.c3sl.ufpr.br/brapci/index.php/article/download/18785>

McClelland, D. C. (1973). Testing for Competence Rather Than for Intelligence. *American psychologist*. Jan. Recuperado de <https://www.therapiebreve.be/documents/mcclelland-1973.pdf>

Richardson, R. J. (2012). *Pesquisa social: métodos e técnicas*. 3. ed. São Paulo: Atlas.

Rodrigues, D. B., & Muniz, E. M. P. (2010). Gestão por competências em bibliotecas universitárias. *Anais do Seminário Nacional de Bibliotecas Universitárias - SNBU*, 16, São Conrado/RJ. Recuperado de http://www.sibi.ufrj.br/snbu2010/pdfs/orais/final_507.pdf

Rudio, F. V. (2001). Pesquisa descritiva e pesquisa experimental. In F. V. Rudio, *Introdução ao projeto de pesquisa científica*. Petrópolis: Vozes.

Ruzzarin, R. et al. (2002). *Gestão por competências: indo além da teoria*. Porto Alegre: SEBRAE/RS.

Vendrell, B. B., & MIRANDA, M. J. (1999). Competências del profesional de la información: un andén em el Mercosur. *Anais do Encontro de Directores y Docentes de Escuelas de Bibliotecologia del Mercosur*. Santiago, Chile, UTEM, p. 99-104. Recuperado de http://argos.fhycs.unam.edu.ar/bitstream/handle/123456789/510/Be_nitez-Competencias.pdf?sequence=1

Woodruffe, C. (1991). Competent by any other name. *Personnel Management*, p. 30- 33, Sept.

Zarifian, P. (2012). *Objetivo competência: por uma nova lógica*. São Paulo: Atlas.



RECONSTRUCTING NEWS SPREAD NETWORKS AND STUDYING ITS DYNAMICS

Elisa Mussumeci¹, Flávio Codeço Coelho²

¹Fundação Getúlio Vargas, 0000-0002-7882-6465, elisamusumeci@gmail.com

²Fundação Getúlio Vargas, 0000-0003-0635-8989, fcocoelho@fgv.br

ABSTRACT News spread can be seen as a contagious process in internet media outlets. This process can be transformed into a temporal network which represents the influence between published articles and between media outlets. In this article, we propose a methodology based on the application of natural language analysis of the articles to reconstruct the latent network through which news spread. From the reconstructed network, we analyze the network dynamic and then show that the dynamics of the news spread can be approximated by a classical SIR epidemiological dynamic upon the network. From the results obtained we argue that the methodology proposed can be used to make predictions about media repercussion, and also to detect viral news in news streams.

KEYWORDS News, SIR, Epidemics, Temporal Networks

COPYRIGHT This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

The Internet is the main channel for information dissemination in the 21st century. Information of any kind is posted online and spreads via recommendations or advertisement (Hermida et al. (2012), Romero et al. (2011)).

Studying the spread dynamics of information through the internet, is a very relevant and challenging activity, since it can help the understanding of the factors which determine how far and how fast a given information can go. The most common way to observe information flow on the web, is by tracking how many times a given piece is replicated by different users over a period of time. Sometimes the content is modified as it is replicated, making it harder to track. This type of dynamics is common to news articles, social media posts, and a wide variety of content which gets replicated in digital networks.

In the specific case of news articles, a number of factors influence their spread. Among the most important are the reputation of the original publisher -- though that is not easy to measure, and the size of readership of a particular publisher, which will determine the initial spread of any piece. However, the topology of the resulting network associated with dissemination of news is hard to predict and will depend on the subject of each news piece and its resonance with public interests.

In this work, we decided to look at the spread of news stories over the Internet characterizing the resulting spread network and the dynamics of the spread. We start by looking at an actual case of news spread, and reconstructing the spread network by applying ideas of temporal networks and topic modeling, connecting similar articles within the bounds of temporal window of influence. Then we postulate that the spread dynamics approximates an epidemic process and model it using a Network SIR model (Pastor-Satorras et al. (2015)). Modeling the spread of ideas as an epidemic process is not a new idea (Bettencourt et al. (2006)), but here we propose new tools to estimate the spread network from data and compare it with simulated networks produced by an SIR epidemic model. By studying the topologies of these networks and the typical dynamics, we believe it is possible to detect anomalies in the propagation of news in digital media.

METHODOLOGY

DATA SOURCES

The data used for this study was obtained from the Media Cloud Brasil project (MCB) which collects news articles from thousands of sources in the Brazilian Internet since 2013. From the MCB database we obtained 2129 articles talking about the Charlie Hebdo terrorist attack in February 2015. The corpus of articles spans from the day of the attack to the end of march of 2015. The data include the full text of the article, the URL of publication and the date and time of the publication. The choice of topic was guided by the need to have a well defined news topic that was not very likely to be confounded with other stories being published in the same period.

ARTICLE SIMILARITY

To calculate a measure of similarity between text documents one can rely on a number of metrics for textual distance described in the literature (Mihalcea et al. (2006)). Most of these metrics are based on a bag-of-words representation of texts, meaning that texts are defined in terms of which words they contain and their respective frequencies in the documents. Such representation completely disregards higher level linguistic features of texts such as syntactic and semantics. In this analysis, we want to use semantic similarity to describe the association between articles. In order for a news article to influence another, they must talk about the same concepts.

In order to capture the semantics of the articles we started by building a word vector representation for every word in our corpus' vocabulary, taking into account the co-occurrence of words within a sentence. This model is built from a larger corpus of news articles (approximately 2.6 million articles) according to the Skip-gram model, which has been shown to map the words to a vector space where semantic similarity is highly correlated with the cosine distance between word vectors (Mikolov et al. (2013)). This larger corpus corresponded to the total collection of articles of the MCB project at the time this analysis was done. The importance of training the word vector model on a corpus as large as possible, is that one gets a more accurate semantic representation of each word as a vector. It is important that the larger corpus represents a similar informational space as the sample we are trying to analyze.

Table 1. Parameters of the skip-gram model used to construct the word vectors.

Parameter	Value	Meaning
Minimum word count	10	Word minimum frequency in the corpus
Number of features	300	Dimension of word vectors
Context	10	Text window around word

The word vector model, was trained with the parameters described in table 1. The fitted word vector model consists of a matrix of m word vectors (w_i) as rows. Each row represents a n -dimensional feature vector, with $n = 300$:

$$\begin{matrix} Pm & f_1 & f_2 & \dots & f_n \\ w_1 & [a_{11} & a_{12} & \dots & a_{1n}] \\ w_2 & [a_{21} & a_{22} & \dots & a_{2n}] \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_m & [a_{m1} & a_{m2} & \dots & a_{mn}] \end{matrix}$$

From the word vectors obtained, we created document vectors defined as a weighted sum of word vectors. For a document d containing k distinct words, its vector representation D is given by:

$$\vec{D} = \sum_{i=1}^k w_i \times W_{w,d} \tag{1}$$

where $w_{w,d}$ is the weight of the word w_i in the document d . This weight can be calculated in different ways, for this work we used the TFIDF score (Hiemstra (2000)) of the word in the document. Another possibility would be to use the frequency of the word in the document.

The TFIDF weight is the product of two numerical statistics: term frequency and inverse document frequency. There are several ways of calculating this weight, the one used in this article was the below:

$$f_{ij} * \log\left(\frac{D}{df_i}\right),$$

where f_{ij} is the frequency of the word i in document j , D is the corpus, and df_i is the number of documents where the term i appears. The second term of this equation is the IDF, that measures how much information the word provides, that is, it's shows us if it's a common word or a rare one.

From the weighted sum, we obtain document vectors which can also be represented by a matrix:

$$\begin{matrix} & f_1 & f_2 & \dots & f_n \\ d_1 & [a_{11} & a_{12} & \dots & a_{1n}] \\ d_2 & [a_{21} & a_{22} & \dots & a_{2n}] \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_M & [a_{M1} & a_{M2} & \dots & a_{Mn}] \end{matrix}$$

Now we can define the similarity between two documents $\{A, B\}$ as the cosine of the angle θ between their vector representations:

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (2)$$

TEMPORAL ASSOCIATION

Once the similarity of two articles is calculated, their temporal association must be determined in order to consider the probability of the older article being the *infector* or influencer of the more recent one. To determine the most-likely *infector* of an article, we ranked all articles by date of publications and looked within a fixed time window preceding the publication of each article, for the articles which are most semantically similar. The choice of the size of the time window was determined in order to encompass the majority ($\geq 95\%$) of previous similar articles (see figure 4).

RECONSTRUCTING THE SPREAD NETWORK

To reconstruct the spread network of the news, we defined the nodes of our network as the articles published on the subject chosen, and the edges as the infection events, i.e., for every article after the first one, it must have been influenced (infected) by a previously published article. To qualify as an

infector, an article must precede the infected article by less than γ hours, and have a score of similarity (defined by equation (2)) to the infected article of at least ρ . The reconstruction procedure is summarized in the four steps below.

- 1) Rank all articles in ascending publication time. Let p_i denote the publication date of article i .
- 2) Create upper triangular matrix D , where $d_{ij} = H(\gamma \delta_{ij}) * \delta_{ij}$ and $\delta_{ij} = p_j - p_i$. H is the Heaviside function.
- 3) Create similarity matrix S . Where s_{ij} is the similarity defined by equation (2) whenever $d_{ij} \neq 0$ and 0 otherwise.
- 4) For each article j , we define its influencer i as the article corresponding to $\max(s_j)$.

	j=1	2	3	4	5	6
i=1	0	0.85	0.78	0.82	0.82	0.88
2	0	0	0.82	0.89	0.89	0.90
3	0	0	0	0.78	0.77	0.91
4	0	0	0	0	0.99	0.90
5	0	0	0	0	0	0.92
6	0	0	0	0	0	0

Figure 1: This figure is a slice of similarity matrix S , it shows the first 6 articles. In red are the maximum similarity score for each column, which we use to define its infector, per example, the article 4 has been infected by the article 2.

SIMULATION MODEL

To test the hypothesis that news spread follows an epidemic process, we proposed an SIR model for the spread, following the formalism of (Pastor-Satorras et al. (2015)). In this formalism, instead of modeling the status of a given individual as Susceptible (S), Infectious (I) or Recovered (R), we model the probability of each article being in each of the states, in this case, an **S** article would be one which has yet to be published, and **I** one which is published and has been infected by the story and an **R** is one which is too old to influence new articles. This modeling leads us to equations (3).

$$\begin{cases} \frac{d\rho_i^I}{dt} = -\rho_i^I(t) + \lambda\rho_i^S(t) \sum_{j=1}^N a_{ij}\rho_j^I(t) \\ \frac{d\rho_i^S}{dt} = -\lambda\rho_i^S(t) \sum_{j=1}^N a_{ij}\rho_j^I(t) \end{cases} \quad (3)$$

In equation (3), $\rho_i^I(t)$ is the probability of article i being in the infectious state at time t , similarly for $\rho_i^S(t)$; a_{ij} is the probability of article j being influenced by i and comes from the adjacency matrix of the network. λ is an adimensional transmission parameter given by $\lambda = \frac{\beta}{\mu}$. Time (t) in these equations is also adimensional as it is scaled by μ .

The network for the simulation is built from the same node set of the empirical data. The adjacency matrix A is given by

where N_{XY} is the number of times an article from publisher X (the publisher of article i), has infected an article from publisher Y (the publisher of article j) and N_Y is the total number of articles from publisher Y that have been infected, regardless of publisher. These counts are derived from the empirical dataset.

$$\begin{cases} i = j : a_{ij} = 0 \\ i \neq j : a_{ij} = \frac{N_{XY}}{N_Y} \end{cases} \quad (4)$$

The solution of this model generates the temporal dynamics of the probabilities described in (3). From the solutions, $\rho_i^S(t)$ and $\rho_i^I(t)$ we can derive realizations of states for each article, $S_i(t)$, $I_i(t)$, and $R_i(t)$.

To reconstruct the states, we must sample from the probability distribution the states at each time t , conditioning on the previous state. We follow the procedure:

- 1) Let S_t , I_t and R_t be binary state vectors from article states at time t , where 1 means the article is in that state.
- 2) Iterate from $t = 0$ until the final time step available.
- 3) For each time $t > 0$ generate a newly infected I_t^* vector, in which each element i is a realization of a Bernoulli event with probability given by $\rho_i^I(t) \times S_{t-1}[i]$.
- 4) Similarly to the previous step, sample a new R_t vector, in which each element i is a realization of a Bernoulli event with probability given by $\rho_i^R(t) \times I_{t-1}[i]$.
- 5) Update $I_t = I_{t-1} - R_t + I_t^*$.
- 6) Update $S_t = S_{t-1} - I_t^*$.

At the end of the procedure above we'll have three matrixes of states: I , S and R . Each matrix is composed by the states vectors (S_t, I_t, R_t) and represents the states dynamics of the model.

CONSTRUCTING THE SIMULATED SPREAD NETWORK

From the state matrix I we have which articles get infected at each time t . To create a spread network for the simulation we need to define the infectors for each time. For that we used the probability matrix A defined by equation (4). The following steps describe the entire procedure.

- 1) Let I_t be binary state vectors for articles at time t , where 1 means the article is infected.
- 2) Iterate from $t = 1$ until the final time step available.
- 3) For each article i infected at I_t , obtain its probable infectors, P_i , by multiplying I_{t-1} by the column k of matrix A , where $k = i$ and the values are the probability of each article j from I_{t-1} has to infect i (a_{ij} of the matrix A).
- 4) Define the infector of i by sampling from a multinomial distribution with $p = P_i$.

The figure 2 shows the procedure, where the rows t_i are the states vector and are composed by the state value for each document in the step i :

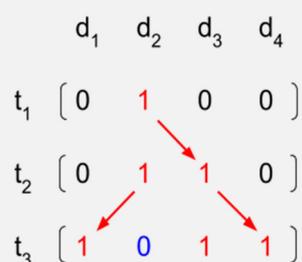


Figure 2: The arrows indicates the infector for each article. The red articles are the ones infected, that is, the ones that can spread the infection, and the blues are the ones that had recovered.

FINDINGS

The dataset used is the result of a very specific search on a news articles database, therefore we can expect to the articles to display a great similarity among themselves.

Looking to the distribution of pairwise similarities that were used to construct the empirical influence network, we can notice that for almost every article there is at least one other with similarity equal or greater than 0.8. Identical articles (similarity equals to 1) were not considered for edge formation.

Figure 3 shows the similarity threshold of the influence network. In order to have a giant component in the network that contains at least 80% of the articles, we need to consider a minimum of 0.8 score similarity. Therefore, we defined $\rho = 0.8$.

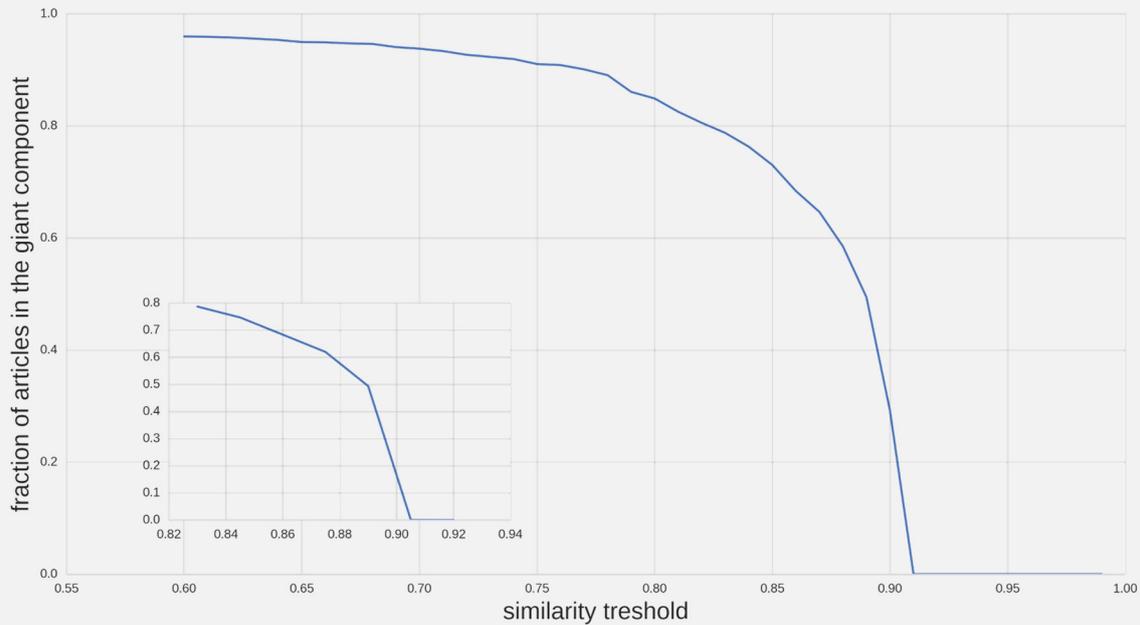


Figure 3: Similarity threshold for the reconstruction of the influence network.

To determine the optimal time window γ in which to search for influencers, we looked at the distribution of time lags from the most similar article (most likely influencer) at various window lengths (figure 4). Even for time windows as long as 15 days, 95% of the influencers were within 7 days of the articles they influenced.

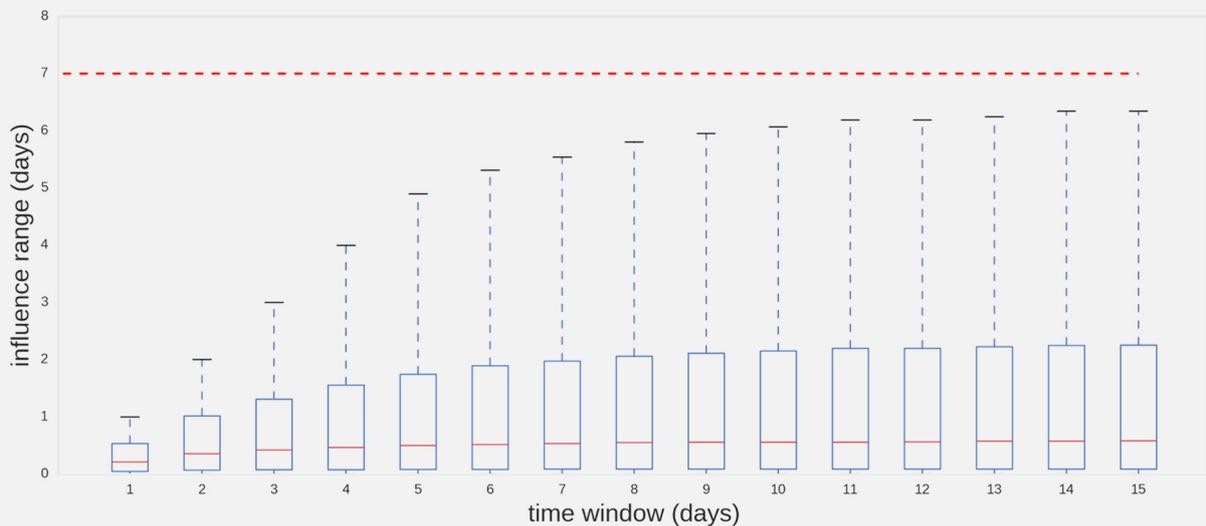


Figure 4: Distribution of time lag from influencer for multiple time window lengths. Notice that no article lags more than seven days from its influencer.

To create the spread network, we defined influence based on the time lag from each pair of articles and from their similarity. Following the previous analysis, we defined $\rho = 0.8$ and $\gamma=168$, which means the infector must precede the infected article by less than 168 hours (7 days) and have at least a 0.8 score of similarity.

Looking at the publication date distribution we notice that the maximum number of articles published in a day was between 250 and 300. We derive the simulation parameters from this distribution. For example, on figure 5 we plot the peak of the infection for a range $([0,0.00005])$ of λ values. From that distribution of peak magnitudes we selected a λ to match the empirical peak: $0.00002 < \lambda < 0.00003$.

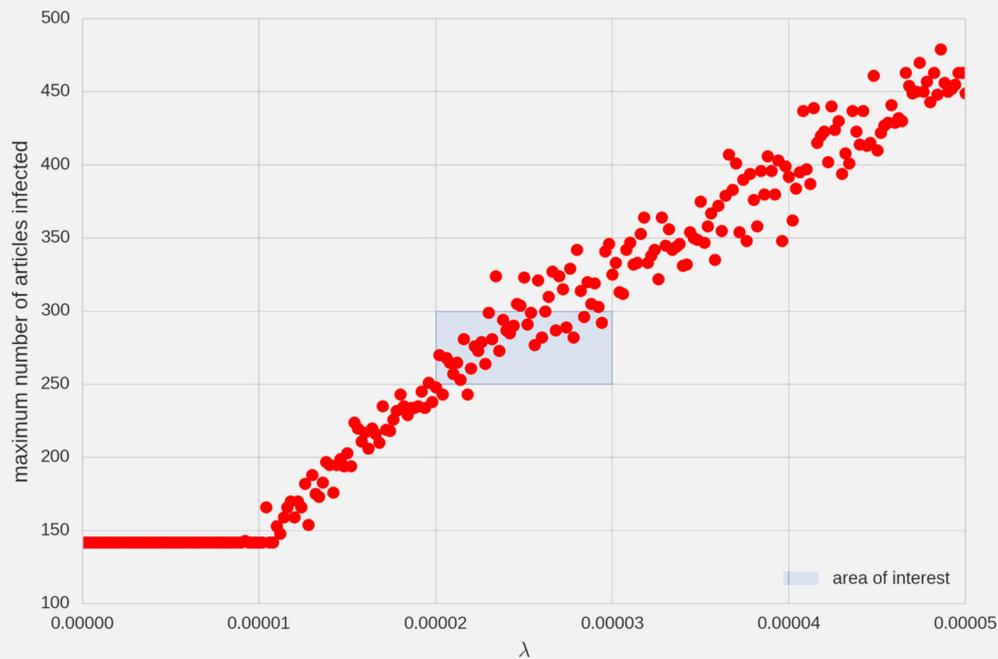


Figure 5: Total number of articles infected between $0 < \lambda < 0.00005$. The blue area is the area where the peak of the simulation is the same as the peak of the dataset distribution, therefore is the area where the λ values were tested for our simulation.

From the simulation (figure 6) we obtain the state matrix, which we use to compare the simulated infection distribution with the original data. Then we ran 10 thousand simulations to show that the model proposed matches the real world dynamics of news articles influences (figure 7).

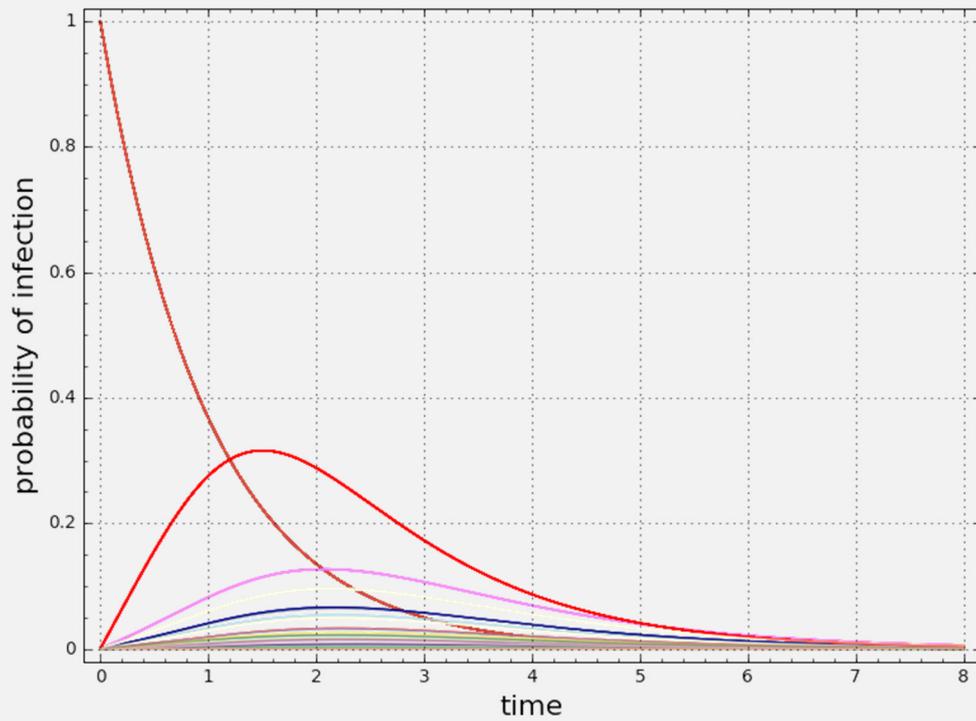


Figure 6: Simulation for $\lambda = 0.0000215$. Each curve represents the n the infectious state as a function of time, for every article. The time units are $1/\mu$.

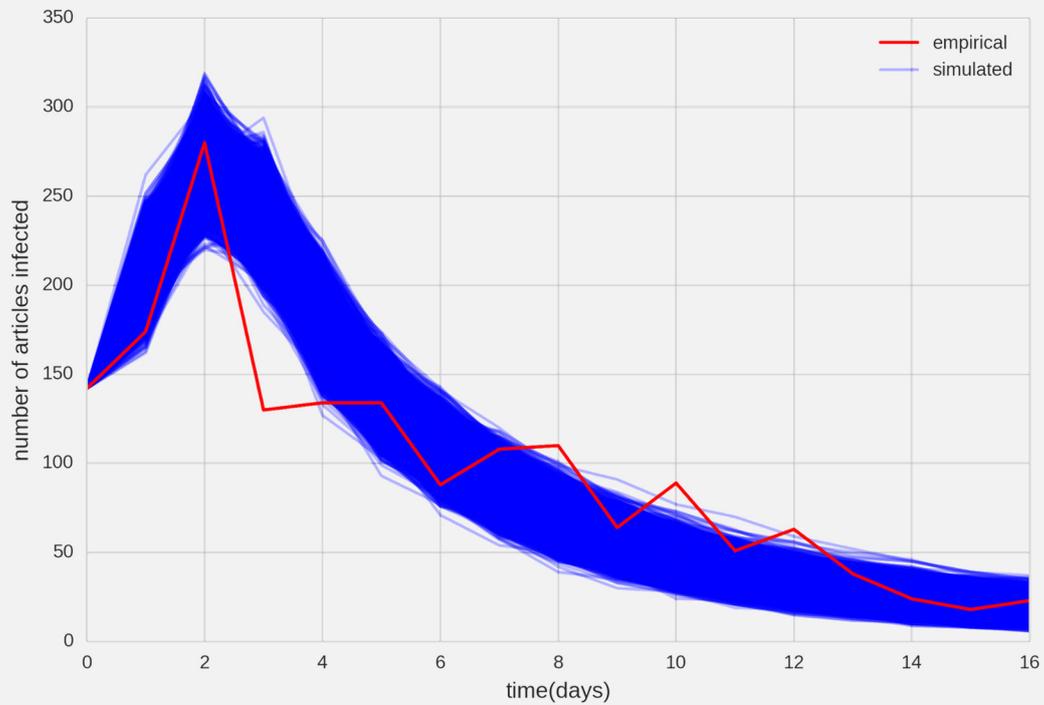


Figure 7: The blue curves are the 10.000 realizations of the state matrix. Notice that the simulated dynamics match the empirical curve.

CONCLUSIONS

In this paper, we presented a methodology for reconstructing the network representing the spread of news in digital media. The results proposed started from a well defined subset of articles with high semantic similarities. However, we believe the criteria of similarity used to reconstruct the network would work even on a random sample of articles, provided that it was large enough to contain a good portion of the putative spread network one is trying to characterize. In other words, the reconstruction algorithm can be used to detect contagious structures within any large enough collection of news articles. Although we have not tested, we believe that the article similarities could be obtained from a topic modeling algorithm less dependent on a large training corpus, such as the Latent Semantic Analysis model. We nevertheless still favor using the word embedding model presented here, as it is not that difficult to obtain reasonably large corpora for training such models. We remark that these corpus size requirements do not apply to the corpus of news whose network we want to reconstruct.

We also demonstrated that a classical SIR process over the network is driving the spread dynamics. This means that if one is able to observe the start of the spread, the overall reach and time of persistence in the media can be predicted from analytical results available for the SIR model, namely the early estimates of the basic reproductive number or R_0 of the model, which can be calculated very early on from the number of articles the first article influences.

The type of analysis introduced here could very easily be adapted to posts on social media, given that a reliable means to capture the full publication set is available. It is increasingly common to see mechanisms to boost social posts, by analyzing the spread dynamics of posts, could help discriminate boosted from unboosted posts. The dynamics of spread can also be used to measure the efficacy of said boosting mechanisms. More generally this methodology can be used for any kind of meme spread ranging from applications in marketing, to political propaganda. More research is needed to characterize the diversity of spreading patterns we can observe in digital media networks.

The news subject selected, "Charlie Hebdo attack", represents a very spontaneous media coverage given the great surprise with which it happened, but also due to the homogeneous response of the global media condemning the cowardly attack. We believe that deviations from the classical SIR dynamics shown here can hint at some form of media manipulation, but that hypothesis remains to be tested based on well defined cases, such as purchased media coverage during political campaigns, etc. With the current concerns about "fake" media pieces (Berkowitz, Schwartz (2016)), Perhaps the methods presented here can help to discriminate authentic media articles from fake ones based on their spread dynamics or influence patterns. We already began to see some attempts to automatically detect fake news (Jin et al. (2016), Rubin et al. (2015)), but they mostly rely on linguistic cues. We believe that qualitative and quantitative aspects of the spread networks can also be of use, as they reveal media manipulation mechanisms, such as "pay to publish" or editorial biases of mainstream outlets which can block or favour certain themes.

REFERENCES

- Berkowitz, D., & Schwartz, D. A. (2016). Miley, CNN and The Onion: When fake news becomes realer than real. *Journalism Practice*, 10(1), 1-17.
- Bettencourt, L. M., Cintrón-Arias, A., Kaiser, D. I., & Castillo-Chávez, C. (2006). The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. *Physica A: Statistical Mechanics and its Applications*, 364, 513-536.
- Hermida, A., Fletcher, F., Korell, D., & Logan, D. (2012). Share, like, recommend: Decoding the social media news consumer. *Journalism Studies*, 13(5-6), 815-824.
- Hiemstra, D. (2000). A probabilistic justification for using tf×idf term weighting in information retrieval. *International Journal on Digital Libraries*, 3(2), 131-139.
- Jin, Z., Cao, J., Zhang, Y., Zhou, J., & Tian, Q. (2016). Novel Visual and Statistical Image Features for Microblogs News Verification. *IEEE Transactions on Multimedia*.
- Mihalcea, R., Corley, C., & Strapparava, C. (2006, July). Corpus-based and knowledge-based measures of text semantic similarity. In *AAAI* (Vol. 6, pp. 775-780).
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
- Pastor-Satorras, R., Castellano, C., Van Mieghem, P., & Vespignani, A. (2015). Epidemic processes in complex networks. *Reviews of modern physics*, 87(3), 925.
- Romero, D. M., Galuba, W., Asur, S., & Huberman, B. A. (2011, September). Influence and passivity in social media. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (pp. 18-33). Springer Berlin Heidelberg.
- Rubin, V. L., Chen, Y., & Conroy, N. J. (2015). Deception detection for news: three types of fakes. *Proceedings of the Association for Information Science and Technology*, 52(1), 1-4.



A ÁREA DE DOMÍNIO NO ÂMBITO DA DISCIPLINA DE LINGUAGENS DOCUMENTÁRIAS

Heliomar Cavati Sobrinho¹, MariângelaSpottiLopesFujita², Natanna Santana de Moraes³

¹*Departamento de Ciências da Informação. Universidade Federal do Ceará, Brasil,
0000-0002-0447-0072, heliomarcavati@yahoo.com.br*

²*Departamento de Ciência da Informação, Universidade Estadual Paulista – Unesp, Brasil,
0000-0002-8239-7174, fujita@marilia.unesp.br*

³*Departamento de Ciências da Informação. Universidade Federal do Ceará, Brasil,
0000-0003-4355-4153, natanna.msantana@gmail.com*

RESUMO Investiga o problema da escolha do domínio e subdomínio no âmbito do ensino da disciplina de “Linguagens Documentárias Alfanuméricas”, por meio da metodologia de “análise de conteúdo” de Bardin (2011), uma técnica de investigação interdisciplinar e inovadora na área da Organização e Representação do Conhecimento. O problema caracteriza-se pela investigação de como a escolha dos domínios e subdomínios compromete ou não o desenvolvimento de uma Linguagem Documentária e, conseqüentemente, para a formação e atuação do profissional em organização do conhecimento. Seu objetivo principal é analisar a relevância da compreensão da teoria e a prática da “Escolha de Domínio” e “Subdomínio” na formação discente, por meio da análise do conteúdo dos minitesauros desenvolvidos no decorrer da disciplina de “Linguagens Documentárias Alfanuméricas”, que é ministrada no sétimo semestre, quando já estão quase se formando para atuar na sociedade e no mestrado acadêmico para a formação docente na linha de pesquisa de “Representação e Tecnologia da Informação”. Os principais resultados foram as análises das áreas de domínios e subdomínios de 46 minitesauros elaborados por 96 discentes ao longo de seis semestres. Conclui-se que a investigação do problema da escolha do domínio e subdomínio no âmbito do ensino da disciplina de “Linguagens Documentárias Alfanuméricas” foi alcançada, assim como a compreensão de que o conhecimento dos seus conceitos compromete a construção dessas linguagens e a formação e atuação do profissional em organização do conhecimento. Recomenda-se o aprofundamento da utilização da técnica de investigação de análise de conteúdo da Bardin (2011) como instrumento complementar as teorias e normas consolidadas de construção de Linguagens Documentárias, viabilizando a inovação e o desenvolvimento da área de representação do conhecimento.

PALAVRAS-CHAVE *Linguagem Documentária, Área de Domínio, Análise de Conteúdo, Formação e Atuação Profissional em Organização do Conhecimento.*

ABSTRACT It investigates the problem at domain and subdomain choice in the context of the discipline of “Alphanumeric Documentary Language” by means of the Content Analysis methodology by Bardin (2011), an innovative and interdisciplinary method in the Knowledge Organization and Representation field. The problem is characterized through the investigation of how the domain and subdomain choices compromise or not the Documentary Language development and the professional education and performance in Knowledge Organization field. Its main objective is analyze the relevance of the comprehension of the theory and practice of domain and subdomain choices in the students education through the Content Analysis of mini thesaurus made in the course of the discipline of “Alphanumeric Documentary Language” that is offered on seventh semester when the students are almost graduated and prepared to perform in society as well as in a Master Degree Program aimed at teacher education in “Information representation and technology” research line. The main results is the analysis of the domains and subdomains area of 46 mini thesaurus

made by 96 students during 6 semesters. As a result, the investigation of the problem at domain and subdomain choice in the context of the discipline of “Alphanumeric Documentary Language” performance was achieved as well as the comprehension that the knowledge about domain and subdomain’s meanings compromises these languages construction and the professional education and performance in Knowledge Organization field. It is suggests a deeper Content Analysis methodology using by Bardin (2011) as an additional instrument to the established theories and standards of Documentaries Languages Construction, enabling the innovation and the development the Knowledge Representation field.

KEYWORDS *Documentary Language, Domain Arera, Content Analysis, Professional education and performance in Konwledge Organization field.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Esta pesquisa investiga o problema da escolha do domínio e subdomínio no âmbito do ensino da disciplina de “Linguagens Documentárias Alfanuméricas”, de um Curso de Biblioteconomia de uma Universidade Federal, no Brasil, analisando os minitesauros de seis semestres, no período de 2014.2 a 2016.2, por meio da metodologia de “análise de conteúdo” de Bardin (2011), uma técnica de investigação interdisciplinar e inovadora na área da Organização e Representação do Conhecimento.

O problema caracteriza-se pela investigação de como a escolha dos domínios e subdomínios compromete ou não o desenvolvimento de uma Linguagem Documentária e, conseqüentemente, para a formação e atuação do profissional em organização do conhecimento.

Segundo Lancaster (1987, p. 6, **negrito do autor**), “um curso sobre elaboração de tesauros **deve** revestir-se de um cunho prático. De cada estudante, ou pequeno grupo de estudantes, deve-se exigir a compilação de um pequeno tesouro – de aproximadamente 200 termos – em alguma área de assunto”.

Neste sentido, justifica-se a execução desta pesquisa, pois a oportunidade de acompanhar o processo de ensino e aprendizagem desta disciplina nuclear, da área da Ciência da Informação, desde 2014, pretendendo-se dar continuidade, por pelo menos uma década, possibilitará desenvolver metodologias pedagógicas que agreguem valor ao discente e à sua formação, adequando-a as transformações tecnológicas informacionais da ciência e da sociedade.

Seu objetivo principal, portanto, é analisar a relevância da compreensão da teoria e a prática da “Escolha de Domínio” e “Subdomínio” na formação discente, por meio da análise do conteúdo dos minitesauros desenvolvidos no decorrer da disciplina de “Linguagens Documentárias Alfanuméricas”, que é ministrada no sétimo semestre, portanto quando já estão quase se formando para atuar na sociedade.

A Linguagem Documentária (LD’s) segundo Gardin (1968), citado por Cintra *et al.* (1994, p. 25), é “um conjunto de termos providos ou não de regras sintáticas, utilizado para representar conteúdos de documentos técnico-científicos, com fins de classificação ou busca retrospectiva de informações”. As LD’s, como instrumentos da representação documentária, são construídas para possibilitar a comunicação entre os conteúdos dos documentos e os usuários de um sistema de informação, por meio da utilização de termos e conceitos (Fujita, 2013; Lara, 1999).

A disciplina que auxilia as atividades de representação documentária é a Terminologia que, “ao proceder à sistematização dos conceitos de cada área do conhecimento, determina igualmente a condição de referência dos termos que irão compor instrumentos terminológicos para fins documentários” (CERVANTES, 2009, p. 16).

A Terminologia, seja qual for sua temática ou o contexto em que se produz, cumpre sistematicamente duas funções: de representar o **conhecimento especializado** e a de transmiti-lo. Evidenciamos a Terminologia que focaliza os estudos dos conceitos, definições e denominações. Tem sua atenção voltada, também, para a compilação, descrição, apresentação, criação e normalização de termos de **áreas especializadas** com a preocupação de promover a comunicação e o intercâmbio entre os especialistas e os profissionais. Nesse sentido, a Terminologia é uma disciplina que oferece, de forma sistemática, referencial do **vocabulário de uma determinada especialidade** e sua análise para a identificação de termos. Ao produzir novos conhecimentos, os especialistas em uma determinada matéria criam novos conceitos que necessitam ser identificados e compartilhados. (CERVANTES, 2009, p. 125, grifos nossos¹).

Para Cervantes (2009) a primeira etapa do seu Modelo Metodológico Integrado para Construção de Tesouro constitui-se no trabalho preliminar de: “escolha do domínio e da língua do tesouro; delimitação do subdomínio; estabelecimento dos limites da pesquisa terminológica temática; e a consulta a especialista do domínio/subdomínio” (p. 163). Neste sentido, faz-se necessário uma escolha atenta dos critérios de seleção do domínio, pois a partir desta será delimitado o subdomínio e os limites da pesquisa terminológica, ou seja, do sistema nocional de onde serão extraídos os termos que farão parte da construção da Linguagem Documentária.

É a partir do sistema nocional que as relações e as especificações da linguagem em uso podem ser analisadas e se constituírem numa linguagem documentária. Cintra et al. (1994) destaca a fase de delimitação nocional como parâmetro para a sistematização de um determinado campo do conhecimento ao permitir a identificação do sistema de noções e dos parâmetros de relações entre os termos.

Realizada a escolha do domínio é preciso delimitá-lo em um subdomínio, pois,

Recomenda-se não desenvolver uma pesquisa terminológica sobre um domínio completo: por um lado, devido à complexidade e amplitude que supõe uma tal tarefa; e, por outro lado, porque em grande parte do tempo, um domínio compreende não somente uma rede nocional que lhe é própria, mas também numerosas redes nocionais conexas. (Rondeau (1984) citado por Cervantes, 2009, pp. 147-149).

Após esta delimitação do subdomínio é imprescindível o estabelecimento dos seus limites terminológicos, de onde serão extraídos os termos para a construção do tesouro.

O limite da extensão da pesquisa terminológica, quanto ao número aproximado de termos, é estabelecido em função dos objetivos propostos, da disponibilidade de tempo e de meios financeiros. Desse modo, pode-se escolher um levantamento básico, compilando-se, em média, 300 termos, ou exaustivo, por volta de 2.500 termos. (Rondeau (1984) citado por Cervantes, 2009, p. 147-149).

¹ Grifamos o conceito de Terminologia para salientar que pode-se considerar o conceito de “domínio” como sinônimo de “conhecimento especializado”, “áreas especializadas”, e “vocabulário de uma determinada especialidade”.

Nesta pesquisa, utilizamos o conceito de *domínio* da área de Terminologia que, segundo a ISO 1087(2000), é o subconjunto de uma área determinado por um sistema de noções, sendo a *área* uma parte do saber cujos limites são definidos segundo um ponto de vista particular de uma ciência ou técnica.

A relevância da pesquisa de Cervantes (2009), portanto, ao elaborar um modelo metodológico integrado para construção de tesouro que, inicialmente, enfatiza como etapa inicial a delimitação de um domínio é ressaltada por Lara (1999) ao definir que:

O significado das palavras nos discursos de domínios e áreas de atividade é explicitado a partir e no universo desses grupos particulares, reproduzindo seus padrões valorativos e cognitivos, sua estrutura e processos. A terminologia constitui, assim, o lugar privilegiado para a convergência dos vários discursos correlatos, sejam eles contrários ou contraditórios. **A Terminologia, preocupada em “mapear” os conceitos de domínios, permite restaurar a referência a esses domínios, bem como oferece procedimentos orientadores para a identificação e/ou construção de vocabulários específicos.** (LARA, 1999, p. 17, grifo nosso).

Considera-se para finalidade deste trabalho, portanto, o domínio como a área do conhecimento da qual será delimitada um subdomínio que consistirá no sistema nocional de onde serão coletados os termos para construção da linguagem documentária.

Um exemplo é a escolha do domínio na área da Economia, sendo a delimitação do subdomínio, documentos específicos da área, como os Boletins de Conjuntura Econômica, que consistirão no sistema nocional de onde serão coletados os termos.

Assim, feita estas considerações conceituais, aplicou-se a metodologia da análise de conteúdo de Bardin (2011) para prospectar os domínios e subdomínios dos trabalhos de construção de tesouros das disciplinas já referidas, obtendo como principais resultados a análise das áreas de domínios e subdomínios de 46 minitesouros elaborados por 96 discentes ao longo de seis semestres.

METODOLOGIA

O método para revisão da literatura foi a pesquisa bibliográfica e para levantamento dos dados foi utilizada a técnica de investigação de análise de conteúdo de Bardin (2011). Este método é relevante porque permite fazer uma pré-análise das categorias e suas definições, a posteriori, possibilitando a inferência prática da resolução do problema, assim como dos resultados.

Segundo Bardin (2011) a análise de conteúdo apresenta métodos e técnicas de análise de conteúdo com exemplos práticos aplicáveis há um número expressivo de possibilidades, deixando claro que cada caso deve utilizar as técnicas que são peculiares ao objetivo da pesquisa, que neste caso são os minitesouros desenvolvidos na disciplina de Linguagens Documentárias Alfanuméricas, do Curso de Biblioteconomia de uma Universidade Federal no Brasil.

A sua organização, dependendo do tipo de análise (quantitativa ou categorial - que é o caso desta pesquisa) compreende três escolhas: o recorte; escolha das unidades; a enumeração: escolha das regras de contagem (se quantitativa); e a classificação e agregação: escolha das categorias (se qualitativa e categorial) (Bardin, 2011, p. 133).

Aqui é necessário esclarecer que a análise de conteúdo realizada nesta pesquisa é categorial temática, portanto, qualitativa, haja vista que o que a caracteriza "é o fato de a "inferência - sempre que é realizada - ser fundada na presença do índice (tema, palavra, personagem, etc), e não sobre a frequência da sua aparição, em cada comunicação individual." (Bardin, 2011, p. 146).

O índice ou tema é definido por Bardin (2011) como categoria, sendo a categorização

uma operação de classificação de elementos constitutivos de um conjunto por diferenciação e, em seguida, por reagrupamento segundo o gênero (analogia), com critérios previamente definidos. As categorias são rubricas ou classes, as quais reúnem um grupo de elementos (unidades de registro, no caso da análise de conteúdo) sob um título genérico, agrupamento esse efetuado em razão das características comuns destes elementos. (Bardin, 2011, p. 147).

O processo de categorização, por ser estruturalista, possui duas etapas: o *inventário*, que consiste em isolar os elementos; e a *classificação*, que consiste em "repartir os elementos e, portanto, procurar ou impor certa organização às mensagens", sendo, dessa forma, o ato de classificar os elementos a "investigação do que cada um deles tem em comum com outros. O que vai permitir o seu agrupamento é a parte comum existente entre eles." (Bardin, 2011, p. 148).

Por fim, segundo Bardin (2011), "um conjunto de categorias boas deve possuir as seguintes qualidades", das quais caracterizamos as principais relacionadas com esta pesquisa:

- a *exclusão mútua*: (...) cada elemento não pode existir em mais de uma divisão;
- a *homogeneidade*: o princípio de exclusão mútua depende da homogeneidade das categorias. Um único princípio de classificação deve governar a sua organização. Num mesmo conjunto categorial só se pode funcionar com um registro e com uma dimensão da análise. Diferentes níveis de análise devem ser separados em outras tantas análises sucessivas.
- A *pertinência*: uma categoria é considerada pertinente quando está adaptada ao material de análise escolhido, e quando pertence ao quadro teórico definido. Na pertinência (*pertinens*; que diz respeito a, relativo a...) há uma idéia de adequação ótima. O sistema de categorias deve refletir as intenções da investigação, as questões do analista e/ou corresponder às características das mensagens;
- A *objetividade* e a *fidelidade*: estes princípios, tidos como muito importantes no início da história da análise de conteúdo, continuam a ser válidos. As diferentes partes de um mesmo material, ao qual se aplica a mesma grade categorial, devem ser codificadas da mesma maneira, mesmo quando submetidas a várias análises.
- A *produtividade*: um conjunto de categorias é produtivo se fornece resultados férteis: em índices de inferências, em hipóteses novas e em dados exatos. (Bardin, 2011, p. 150).

Desta forma realizamos uma pré-análise dos minitesauros obtendo, inicialmente, as seguintes categorias: semestre; quantidade de tesaurus; área de domínio; subdomínio; quantidade de alunos; rede nocional; e fonte para pesquisa. A coleta dos dados foi realizada em uma planilha do *Excell*, sendo os dados e suas análises relatados a seguir.

RESULTADOS

A pesquisa foi aplicada com base na construção de Minitesauros nas turmas das disciplinas de Linguagens Documentárias Alfabéticas e de Linguagem Documentária, Representação e Comunicação da Informação do Curso de Biblioteconomia e do Mestrado em Ciência da Informação, respectivamente.

Inicialmente aplicamos os critérios de categorização estruturalista de Bardin (2011), *inventariando* as categorias detectadas, *a priori*, na fase de pré-análise, e, por meio da *classificação* percebemos o surgimento de novas categorias, permitindo um controle rigoroso, desde o início, dos termos, que denominaremos como uma pré-macroestrutura de uma linguagem documentária. Os subdomínios foram agrupados por áreas do conhecimento, como, por exemplo, na macroestrutura da área do Direito foram agrupadas as categorias “Informação Jurídica”, “Política de Segurança Pública” e “Jurídica Trabalhista”.

Em seguida, foram rigorosamente seguidas e mantidas as principais *qualidades* da categorização preconizadas por Bardin (2011), que são a *exclusão mútua* (os termos não se repetiram em nenhuma categoria); a *homogeneidade* (as categorias foram organizadas segundo um princípio rigoroso de classificação), a *pertinência* (o conteúdo das categorias caracterizam-se como uma garantia literária), a *objetividade* e a *fidelidade* (os procedimentos, fonte e categorias seguem parâmetros aplicáveis sob qualquer perspectiva de análise) e a *produtividade* (os índices qualitativos e quantitativos geraram inferências e dados inequívocos). Como exemplo, na macroestrutura de “Ciências Sociais Aplicadas” as categorias (subdomínios), como a de “Ciência da Informação” e “Biblioteconomia”, não se repetiam em outra macroestrutura, mantendo a *homogeneidade*, por meio de uma classificação rigorosa, o que permitiu a sua *pertinência*, *objetividade* e *produtividade* qualitativa (as categorias temáticas) e quantitativa (quantidade de categorias e termos), possibilitando as inferências que descrevemos, a seguir.

A partir destas categorias detectadas, *a priori*, foi realizada a análise dos conteúdos dos minitesauros que se concentraram nas seguintes categorias quantitativas e qualitativas: Semestre; Área de Domínio; Subdomínio; Quantidade de Minitesauros; Quantidade de Alunos; e Fontes de Pesquisa, conforme o seguinte exemplo: Semestre (2014.2)/ Área de Domínio (Direito)/ Subdomínio (Informação Jurídica)/ Quantidade de Minitesauros (2)/ Quantidade de Alunos (2)/ Fontes de Pesquisa (Artigos).

Durante os semestres de 2014.1 a 2016.2 foram analisados 46 minitesauros construídos segundo o “Modelo Metodológico Integrado para Construção de Tesouro” proposto por Cervantes (2009).

Nas aulas teórico-práticas, os 96 discentes foram divididos individualmente ou em grupos e desenvolveram as seguintes etapas do modelo (Cervantes, 2009): Delimitação do domínio e subdomínio; Estabelecimento dos limites da pesquisa terminológica; Coleta do *corpus* do trabalho terminológico; Etapa de Classificação, verificação e confirmação dos termos e Etapa de apresentação do Tesouro.

No início da fase de *delimitação do domínio* e do *subdomínio*, os estudantes pesquisaram seus temas em artigos científicos, base de dados, dicionários e outras fontes informacionais. Essas fontes informacionais subsidiaram o reconhecimento e a área de especialidade do tema que cada grupo escolheu inicialmente. Estas escolhas, do domínio e subdomínio, foram realizadas de acordo com o tema do projeto de monografia dos discentes, na sua maioria, ou de acordo com a sua área do estágio.

Desta forma, foram associadas à necessidade de atender duas qualidades de categorização da Bardin (2011), a *objetividade* e a *fidelidade*, extraíndo os termos de fontes confiáveis, permitindo a uniformidade dos procedimentos, à necessidade dos discentes se apropriarem do conhecimento que estavam familiarizados.

Como fontes de pesquisa para delimitação do domínio e subdomínio, além da coleta de termos, foram utilizados os seguintes tipos de documentos: Artigos (28 minitesauros); Livros (5 minitesauros); Glossários e Dicionários (4 minitesauros); Documentos Institucionais, como tabelas de temporalidade, boletins e resoluções (4 minitesauros); Vocabulários Controlados e Listas de Assuntos (2 minitesauro); Processos Trabalhistas (1 minitesauro); Vídeos-Aulas (1 minitesauro); e o Portal do IBGE (1 minitesauro). As pesquisas tiveram o intuito de contextualizar cada tema e identificar a terminologia.

Os domínios analisados foram: Direito; Ciências Sociais Aplicadas; Ciências da Saúde; Educação; Biblioteca; Vigilância Sanitária no Brasil; Educação à distância; Linguística; Responsabilidade Social; Ciência da Informação; Literatura; Restauração, Conservação e Preservação; Saúde; Linguagem; Computação; Geografia Estatística; Representação da Informação; Biblioteconomia; Botânica; Recursos Bibliográficos; e Curadoria Digital.

Dos 46 minitesauros (Quadro 1), 43 deles apresentaram subdomínio e 3 não foram descritos, verificando-se que 3 grupos apresentaram dificuldades de apontar qual o subdomínio determinante. Em comparativo a identificação do domínio, a quantidade de subdomínio foi mais expressiva. Assim, destacamos que é mais fácil delimitar o subdomínio do que o domínio e sua divisão enquanto área do conhecimento, considerando análise e inferência das informações desta pesquisa.

Quadro 1. Subdomínio dos Minitesauros

Subdomínio	
Informação Jurídica	Ortografia Lusofônica
Ciência da Informação	A Inclusão Social por Meio da Biblioteconomia
Política de Segurança Pública	Legislação Trabalhista
Teologia Reformada	Gestão da Informação.
Biblioteconomia	Contos de Fada
Recursos Informacionais na Educação	Restauração e documentação bibliográfica
Saúde Pública	Video Aulas do Núcleo de Tec. e Ens. à Distância da UFC
Ballet	Linguagem Politicamente Correta
Saúde Ocupacional	Jogos Eletrônicos
Cinema	Biblioteca do Instituto do Câncer
Jurídica Trabalhista	Canal Temático IBGE Teen
Arquivos da Companhia de Água e Esgoto	Indexação
Castanhas (Frutífera)	Períódio Lamparina: Rev. de Ens. em Teatro EBA/UFMG
Arquitetura	Biblioteca Escolar
Desigualdade Social no Brasil	Biblioterapia
Segurança Pública	Plantas Medicinais
Expressões de linguagem típicas cearenses	Estudos de Usuários
Resoluções da Anvisa sobre os Alimentos Geneticamente Modificados – Transgênicos	Mediação e Uso da Informação
Arquivos Permanentes	Metadados para descrição de recursos bibliográficos
Biblioteca Escolar	Ciência da Informação
Jogos e Brincadeiras	Memória
Estudos em Patrimônio e Memória Social	

Fonte: Elaborado pelos autores.

Um fator influenciador é que parte dos minitesauros apresentaram domínios com base na temática escolhida, ao invés de entender a amplitude do conceito de domínio em comparação com as demais áreas do conhecimento e suas especialidades, parâmetros esses fundamentais para a sua definição e delimitação do subdomínio.

Os domínios mais frequentes entre os trabalhos constituíram-se da área de Ciências Sociais Aplicadas, apresentando subdomínios como, por exemplo: Ciência da Informação, Biblioteconomia e Arquivos Permanentes.

No domínio da Ciência da Informação, propriamente dito, foram prospectados assuntos sobre: Mediação e Uso da informação; Indexação; Estudos de Usuários; e Metadados de Descrição de Recursos Bibliográficos, sendo possível visualizar subcampos científicos de estudos realizados pelo domínio em questão.

O Direito, também, se destaca com 5minitesauros compostos pelos subdomínios de: Informação Jurídica; Política de Segurança Pública e Legislação Trabalhista.

Ciências da Saúde e Vigilância Sanitária no Brasil foram os domínios escolhidos para desenvolver 6 dos minitesauros.

Já na Linguística e Literatura foram desenvolvidos trabalhos com temas voltados para Contos, Linguagem e Memória.

Destacam-se os subdomínios dos minitesauros de Castanhas, Plantas Medicinais, no domínio da Botânica.

Os demais subdomínios foram: Jogos Eletrônicos (no domínio da Computação); Estudos em Patrimônio e Memória Social (no domínio de Patrimônio e Memória); Teologia Reformada (no domínio da Teologia); e Arquitetura no Conselho de Engenharia e Arquitetura (no domínio de Arquitetura).

CONCLUSÕES

Conclui-se que a investigação do problema da escolha do domínio e subdomínio no âmbito do ensino da disciplina de “Linguagens Documentárias Alfanuméricas” foi alcançada, assim como a compreensão de que o conhecimento dos seus conceitos compromete a construção de LD’s. Nessa perspectiva, o domínio e o subdomínio indicados permitiram conhecer o sistema de noções que deram continuidade nas etapas seguintes da construção dessas linguagens documentárias.

Pode-se inferir dos resultados que a maioria dos discentes souberam escolher os Domínios e Subdomínios e que estes contemplaram várias áreas do conhecimento, permitindo deduzir que a teoria e metodologia empregada nesta pesquisa pode ser aplicada no desenvolvimento de LD’s de qualquer área, independente do suporte, assim como no ensino e aprendizagem das disciplinas da área de Organização e Representação do Conhecimento.

Esta pesquisa terá continuidade, no mínimo por dez anos, para que mais informações possam ser inseridas, tornando-a mais relevante e fidedigna as suas mensurações. Neste semestre, 2017.2, estamos desenvolvendo um questionário para aplicar no início e no final da disciplina, visando melhorar a percepção dos discentes em relação à metodologia aplicada.

Também pretende-se complementá-la com a utilização dos softwares “Tematres” e o “Protegé”, assim como o laboratório de informática estiver disponível. Com esta medida espera-se que o aprendizado seja enriquecido para que os discentes agreguem mais conhecimento ao seu aprendizado e *expertise* para atender as necessidades da sociedade.

Percebemos que a técnica de investigação de análise de conteúdo pode ser utilizada em todas as etapas da construção de uma linguagem documentária, principalmente na etapa final de relacionamentos entre os termos, seja por ordenação, subordinação e coordenação; do genérico para o específico, da parte para o todo; hierárquico; vertical e horizontalmente, dentre outras regras das normas, nacionais e internacionais, de construção de linguagens documentárias.

Neste sentido recomenda-se o aprofundamento da utilização da técnica de investigação de análise de conteúdo da Bardin (2011) como instrumento complementar as teorias e normas consolidadas de construção de Linguagens Documentárias, viabilizando a inovação e o desenvolvimento da área de Representação do Conhecimento.

Recomenda-se, ainda, a continuação dos estudos teóricos e práticos nesta área, assim como a replicação desta pesquisa em outras disciplinas afins com as LD's, para contribuir para com a excelência na formação e atuação do profissional em Organização e Representação do conhecimento.

REFERÊNCIAS BIBLIOGRÁFICAS

Bardin, L.(2011). *Análise de conteúdo*. (1a ed., rev. e ampl.). São Paulo: Edições 70.

Cervantes, B. M. N. (2009). *A construção de tesouros com a integração de procedimentos terminográficos*. (Tese de Doutorado em Ciência da Informação). Universidade Estadual Paulista, Marília.

Cintra, A. M. M., Tálamo, M. F. G.M., Lara, M. L. G.,&Kobashi, N. Y. (1994). *Para entender as linguagens documentárias*. (2rd. ed.). São Paulo: Polis.

Fujita, M. S. L. (2013). Representação documentária no processo de indexação com o modelo de leitura documentária para textos científicos e livros: uma abordagem cognitiva com protocolo verbal. *PontodeAcesso*, 7(1), 42-66.

Gardin, J-C.et al. (1968). *L'automatisation des recherchesdocumentaires: unmodèlegénérl “Le SYNTOL*. (2rd. ed. Revue el augmentée). Paris: Gauthier-Villars.

International Organization For Standardization. (1996). ISO 1087 – Terminologia – Vocabulário.Proposta de versão brasileira. [Rio de Janeiro: ABNT, 1996].

Lancaster, F. W. (1987). *Construção e uso de tesouro: curso condensado*. Trad. de César Almeida de Meneses Silva; Rev. de Odilon Pereira da Silva. Brasília: IBICT.

Lara, M. L. G. (1999). *Representação e linguagens documentárias: bases teórico-metodológicas*.(Tese de Doutorado em Ciências da Comunicação). Universidade de São Paulo, Escola de Comunicações e Artes, São Paulo.

Rondeau, G. (1984). *Introduction à la terminologie*. (2rd.ed.). Québec, Canadá: GaëtanMorin.



PRÁTICA DA CATALOGAÇÃO DE MATERIAIS ESPECIAIS COMO PARTE DA FORMAÇÃO DO PROFISSIONAL DA INFORMAÇÃO

Luciana Moreira¹, Monica Marques Carvalho Gallotti²

¹Professora Associada I da Universidade Federal do Rio Grande do Norte, 0000-0001-7265-3164,
lucianamoreiraufrn@gmail.com

²Professora Adjunta IV da Universidade Federal do Rio Grande do Norte, 0000-0002-3044-2032,
monica_mcg@hotmail.com

RESUMO A Representação Descritiva, a Catalogação de itens de informação se constitui desde há bastante tempo em uma prática importante e necessária que garante o registro, processamento e a posterior difusão da informação disponível, contribuindo para a perpetuação da memória. Vivemos em um contexto em que os suportes informacionais não se limitam a um formato único, físico, palpável. Nesse sentido, refletir sobre como se dá o ensino e os desafios presentes na representação da informação de materiais especiais é instigante e fundamental para enriquecer a formação do futuro bibliotecário. Este trabalho, em forma de pôster, visa detalhar ações de ensino realizadas no âmbito da disciplina Representação Descritiva II, ofertada ao curso de graduação em Biblioteconomia da Universidade Federal do Rio Grande do Norte (UFRN) no Brasil. É uma discussão que está ligada à linha de pesquisa “Organização e Tratamento da Informação”, integrante do grupo Informação e Conhecimento na Sociedade Contemporânea, cadastrado no Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq no Brasil. Metodologicamente o trabalho se desdobrou em duas vertentes, uma teórica por meio de uma pesquisa bibliográfica com a fundamentação teórica expressos em conceitos relativos à organização da informação e representação descritiva. Outra vertente foi a de relato de experiências ancorado nos preceitos da pesquisa qualitativa conforme Minayo (1994) e Connelly & Clandinin (1995). Como resultados espera-se refletir sobre ações de ensino de catalogação de materiais especiais, seus desafios e especificidades. Outrossim, visa potencialmente agregar valor à base empírica existente desta área científica bem como induzir à abertura de novos diálogos com profissionais que atuam na formação do profissional bibliotecário fortalecendo assim a área de organização da informação como um todo.

PALAVRAS-CHAVE *Representação Descritiva. Coleções Especiais. Organização da Informação. Representação da Informação. Profissional da Informação.*

ABSTRACT Descriptive Representation and cataloguing of Information has been an important and necessary practice that guarantees registration, processing and subsequent dissemination of the available information, contributing to the perpetuation of memory. Currently we live in a context in which informational means are not limited to a single, physical, palpable format. In this sense, reflecting on the teaching and the challenges present in information representation regarding special materials is important and fundamental in order to contribute to the formation of the future librarian. This paper, in the form of a poster, aims to detail teaching actions carried out within the discipline Descriptive Representation II, offered to the undergraduate degree in Librarianship of the Federal University of Rio Grande do Norte (UFRN) in Brazil. It is a discussion that is linked to the research area "Information Organization and Treatment", member of the research group "Information and Knowledge in Contemporary Society", registered in the National Council of Scientific and Technological Development - CNPq in Brazil. Methodologically the work had two dimensions a theoretical one through a bibliographical research that considered concepts related to the organization of

information, descriptive representation on one hand and the other the work dealt with descriptions of experiences according to the area of qualitative research as seen in Minayo (1994) and Connelly & Clandinin (1995). As results, it is expected to reflect on teaching experiences related to cataloging of special materials, their challenges and specificities. In addition, it aims at potentially adding value to the existing empirical basis of this scientific area as well as induce new dialogues with professionals who work in the formation of the professional librarian thus strengthening the area of information organization as a whole.

KEYWORDS *Descriptive Representation. Special Collections. Information Organization. Information Representation. Information Professional.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A Representação Descritiva, a Catalogação de itens de informação se constitui desde há bastante tempo em uma prática importante e necessária que garante o registro, processamento e a posterior difusão da informação disponível, contribuindo para a perpetuação da memória. Ao longo do tempo, evoluem as técnicas associadas a estes processos, surgem novas tecnologias e diferentes suportes informacionais que não se limitam a um formato único, físico e “palpável”, apresentando-se numa “ecologia” variada e complexa. Por outro lado, estas mudanças, trazem em si desafios e novas questões a serem solucionadas.

Nesse sentido, é necessária uma reflexão constante sobre todos os elementos envolvidos nestes processos e em específico, como está a ocorrer o ensino de temas relacionados a Catalogação em espaços acadêmicos tais como nos Cursos de Biblioteconomia e na formação de futuros profissionais da informação. Outrossim, é necessário abordar casos específicos tais como os relacionados à representação da informação de materiais especiais, prática instigante e desafiadora no contexto atual.

Diante dessas premissas, este trabalho, em forma de pôster, visa detalhar ações de ensino realizadas no âmbito da disciplina Representação Descritiva II, ofertada ao curso de graduação em Biblioteconomia da Universidade Federal do Rio Grande do Norte (UFRN) no Brasil. Tal medida servirá aos propósitos de se partilhar em forma relato, as práticas e estratégias de ensino adotadas, potencialmente agregando valor à base empírica existente desta área científica. O trabalho está dividido em duas facetas. Inicialmente no campo teórico que se assenta em conceitos relacionados à Informação bem como a Representação Descritiva no geral e em específico a catalogação de Materiais Especiais. Empiricamente a pesquisa se ancora na descrição das práticas de ensino adotadas na referida instituição.

Perceber que, para além do objeto em si, um material especial contém informação passível de ser transmitida, de formar um corpo de conhecimento e de transformar uma realidade, é a base para se compreender a importância dos materiais especiais na formação de um acervo, e conseqüentemente, a valorização dos itens informacionais como parte deste.

Assim, o conceito de informação que clarifica de forma basilar a diversidade de suportes informacionais é o apresentado por Silva (2006, p.150, grifo nosso) em que o objeto científico é identificado como sendo um:

Conjunto estruturado de representações mentais e emocionais codificadas (signos e símbolos) e modeladas com/pela interação social, **passíveis de serem registadas num qualquer suporte**

material (papel, filme, banda magnética, disco compacto, etc.) e, portanto, comunicadas de forma assíncrona e multi-direccionada.

Partindo do entendimento que informação é uma ação humana e social, e aplicando a esta noção o delineamento do objeto científico como vimos na definição acima, podemos caminhar bem mais a vontade para tratar os materiais não-livro como ricas fontes de informação. O segundo tema trabalhado, que está diretamente ligado ao conceito de informação é o de documento. Visto a princípio como termos sinônimos (informação, fontes de informação e documentos), direcionamos para uma outra perspectiva, em que a noção de documento, de acordo com Rodriguez Bravo apud Silva (2006, p. 45), é “entendido como o suporte que contém e dá acesso a uma mensagem potencialmente informativa”. Na sequência, a autora diferencia documento, de fonte de informação e de recurso informativo, colocando-se da seguinte forma:

No primeiro [documento], há a vontade do emissor de informar ou de deixar registo de um facto para a posteridade, sendo a informação potencial; e, os outros dois conceitos, ultrapassam o sentido estrito de documento, na medida em que abarcam tudo o que proporcione a um utilizador concreto *información, documentos, objetos, instituciones y personas*.

Isto posto faz-se necessário perceber melhor os conceitos de representação de informação. De forma geral a representação da informação é uma prática de informação que visa a identificação de elementos complexos presentes em determinado suporte que possam ser identificados e representados num novo formato, agora de forma mais condensada. Esse pensamento é corroborado por Novellino (1996, p.38) quando enfatiza que “A principal característica do processo de representação da informação é a substituição de uma entidade lingüística longa e complexa - o texto do documento - por sua descrição abreviada”. Além destes aspectos esta atividade visa a “representação daquilo que se pensa ou como substituição, descrição é representar como ato de reproduzir, descrever, tornar algo presente, interpretar” (Ferreira, 1990).

Tal ação tem como objetivo favorecer a comunicação da informação com fins de sua recuperação. A informação pode ser representada sob o ponto de vista temático, ou seja, relacionado ao assunto do documento ou sob o ponto de vista descritivo, interesse do enfoque deste trabalho. A representação descritiva :

[...] refere-se aos aspectos da descrição formal dos documentos, o que inclui a descrição física e a descrição dos elementos para identificação dos mesmos; a atividade de representação descritiva é também chamada de catalogação (ou, mais especificamente, catalogação descritiva) em especial entre a comunidade de bibliotecas, e de descrição bibliográfica entre a comunidade dos serviços de informação científica (Ortega, 2011, p.45).

A catalogação descritiva prevê a execução de etapas de identificação de “pontos de acesso”, ou seja, entidades representativas de informação com vistas a confecção de registros que serão por sua vez inseridos num sistema de forma ordenada visando a sua recuperação. Para Arilla (1996, p.40) “catalogar um documento é realizar um processo global que obriga a executar uma série de operações, umas identificativas, outras analíticas e outras de ordenação e localização documental, terminando na confecção do produto: o registo bibliográfico”. Todos estes procedimentos objetivam ao final, proporcionar aos usuários a localização da informação de que precisam por meio de um sistema de recuperação da informação como um catálogo.

No entanto, o principal desafio da catalogação é o trabalho com uma grande quantidade e diferentes tipos de materiais, suportes, o que demanda constante atualização e reformulação de seus pressupostos. Esse quadro se agrava ainda mais devido a evolução rápida das tecnologias de informação e comunicação.

Como foi exposto anteriormente a catalogação vem sofrendo mudanças significativas que trazem novos desafios e modelos no tratamento destes recursos informacionais e um desafio pontual é posto nesse panorama e relaciona-se ao ensino da representação descritiva nos cursos de biblioteconomia. Torna-se necessário um estar atualizado à respeito destas mudanças que ocorrem neste cenário a fim de promover uma educação moderna que atenda às necessidades do mercado de trabalho. Além disto, é fundamental refletir sobre casos específicos para que se registrem e se tenham dados empíricos a fim de partilhar experiências e promover mais discussão sobre o assunto.

A PRÁTICA DA REPRESENTAÇÃO DESCRITIVA DE MATERIAIS ESPECIAIS NO CONTEXTO DA UFRN

O curso de Biblioteconomia da UFRN foi criado em 29 de outubro de 1992, e a primeira turma do curso ingressou em 1997, sendo em 2000 formada a primeira turma. E de forma análoga a outros cursos da área, são ofertadas disciplinas voltadas a Representação Descritiva. Na Estrutura curricular atual estão previstas três disciplinas nesta esfera, a Representação Descritiva I que enfatiza os condicionantes históricos e a evolução da catalogação bem como aspectos relacionados a catalogação de livros e sua descrição bibliográfica. A disciplina Representação Descritiva II, que traz noções gerais sobre a organização e tratamento de materiais especiais e a Representação Descritiva III que enfoca a catalogação de recursos informacionais em ambientes digitais.

Nosso enfoque será na disciplina Representação Descritiva II oferecida semestralmente aos alunos do 3º período de formação. O relato de experiências centrou-se em detalhar o ensino da catalogação de itens informacionais como partituras, gravações de som (CD), mapas e artefactos tridimensionais e realias. A escolha dessas categorias de análise se deu devido ao acesso ampliado a acervos organizados que compõem, em grande parte, o Sistema de Bibliotecas da UFRN. Outros motivos se prendem a instigar docentes e discentes a refletir sobre a teoria e a prática, especificamente da catalogação de materiais especiais em um contexto repleto de possibilidades de novas formas de mediação da informação.

METODOLOGIA

A abordagem metodológica usada para a construção dessa comunicação, se baseou na pesquisa bibliográfica conforme proposto por Lakatos e Marconi (2009) e Gil (2008) com a fundamentação teórica relativa à área da Representação Descritiva e conceitos relacionados neste âmbito. A estratégia da pesquisa bibliográfica serviu para melhor fundamentar as práticas empíricas propostas por este trabalho bem como se constituir de elementos de reflexão de como vêm ocorrendo os fenômenos relacionados a Representação Descritiva como um todo e especificamente a catalogação de materiais especiais. A outra vertente da pesquisa se relaciona ao relato de experiências e neste íterim o relato de experiências pedagógicas na área de ensino de catalogação de materiais especiais. O embasamento desta segunda faceta da pesquisa foi feito de acordo com os pressupostos preconizados na área da Pesquisa

Qualitativa e relatos de experiências conforme Minayo (1994) e Connelly & Clandinin (1995) cuja concepção geral prevê que registrar, observar e relatar experiências pedagógicas permite a construção e reconstrução de estórias sociais, métodos de aprendizagem a fim de que as narrativas sejam partilhadas gerando significado e agregando valor à base empírica existente. Outra possibilidade é a busca de reflexão por parte de docentes e discentes sobre a teoria e a prática da catalogação de materiais especiais. A motivação se deu pela prática profissional das autoras em ministrar a referida disciplina, e perceber através do cotidiano, o potencial campo de discussão que está envolto nesta temática de representação de materiais especiais. Como forma de delimitar o escopo de apresentação, dentre uma variedade de materiais especiais, optamos por investigar os seguintes itens informacionais: partituras, gravações de som (CD), mapas e artefatos tridimensionais e realia. A escolha destes, se dá devido ao acesso ampliado a acervos organizados que compõem, em grande parte, o Sistema de Bibliotecas da UFRN.

Partiremos agora para a apresentação da disciplina Representação Descritiva II, que tem como objetivos:

- 1) Oferecer uma visão da catalogação de materiais especiais, familiarizando o aluno com os conceitos básicos da disciplina;
- 2) Desenvolver no aluno a valorização dos materiais especiais como importantes itens informacionais dentro de um sistema de informação;
- 3) Conhecer a estrutura e manusear o AACR2-R;
- 4) Refletir sobre os problemas que envolvem os materiais especiais.

O livro básico utilizado nas aulas é uma adaptação do Código de Catalogação Anglo-Americano (AACR2R), intitulado Catalogação de Recursos Bibliográficos: Aacr2 de autoria de Antonia de C. M.de C. Memória Ribeiro (2015). Nele a autora explica as normas e ilustra com exemplos completos a catalogação dos diversos itens informacionais, possibilitando uma melhor assimilação do conteúdo.

Nesse sentido, Santos e Corrêa (2009, p.69) enfatizam que os conteúdos relacionados à catalogação devem ser repassados de modo que seja “uma disciplina estruturada teoricamente que ser constantemente discutida e aprimorada, de maneira a tornar visível a técnica que permeia o processo de construção de formas de representação e apresentação dos recursos informacionais nos mais diversos ambientes.”

Então corroboramos com o pensamento dos autores na medida em que o ensino da disciplina supracitada requer contante avaliação das metodologias utilizadas a fim de que o processo de ensino/aprendizagem seja realizado da melhor forma possível fomentando as competências adequadas para a formação do futuro profissional de informação.

DESCRIÇÃO E RESULTADOS DAS EXPERIÊNCIAS DE ENSINO

A partir do que foi exposto anteriormente nos deteremos agora a abordar as experiências no cotidiano de processo de ensino/aprendizagem da disciplina Representação Descritiva II, no âmbito da UFRN, para que assim possamos partilhar experiências positivas na formação de profissionais que tem pela frente, múltiplas possibilidades de atuação no mercado da informação.

Um dos desafios enfrentados no ensino da catalogação é tornar a disciplina e as práticas relacionadas à ela mais atraentes. Uma das estratégias adotadas para contornar essa situação foi a de aproveitar as vantagens trazidas pela dinamicidade de formatos de materiais utilizados para a prática da catalogação

na referida disciplina. Nesse sentido, antes de abordarmos os procedimentos técnicos dos materiais buscamos familiarizar os alunos com o tipo de formato, suas características históricas e sociais, o contexto onde estão subscritos. Por exemplo, ao se catalogar obras de arte fazemos visitas de campo em galerias de arte a fim de mostrar a diversidade de tipos de materiais, técnicas utilizadas na sua confecção. Convidamos os alunos a refletirem sobre os possíveis pontos de acesso que representam a informação naquele suporte não convencional, dentre outras ações realizadas com o intuito de despertar o interesse do aluno para a riqueza informacional do material.

Os alunos são, assim conduzidos a vivenciar situações próximas da realidade com os objetos informacionais. Na figura a seguir, a atividade foi com materiais cartográficos, especificamente com mapas, para que fosse feita a leitura técnica e posterior registro de acordo com o AACR2.



Figura 1 – Aula de catalogação de mapas

A dinâmica usada em sala de aula inicia-se sempre com a abordagem teórica sobre o item informacional, seguida de exercícios de aprendizagem e posteriormente a visita a espaços que tem como parte do acervo, materiais especiais.

Um desses espaços é o Museu Câmara Cascudo - MCC, pertencente à UFRN. O acervo é composto por fósseis (realias), objetos de arte, objetos indígenas, exposições temporárias e permanentes (artefatos tridimensionais)¹, dentre outros.

No caso do Museu, os alunos tinham o objetivo de comparar as áreas estabelecidas na catalogação com base no AACR2, com a forma de representação das informações museológicas. Unir a função de memória natural aos museus à representação de cunho informativo é um desafio instigante para perceber as oportunidades de atuação do profissional da informação.

¹ Fundado em 1973, sua principal função é educativa, através das ações de ensino, pesquisa e extensão. Busca “através da interdisciplinaridade, intercâmbio junto aos Departamentos que integram as unidades de ensino da Universidade, tendo como público alvo, a comunidade universitária, e a partir dela, a sociedade em geral.” Museu Câmara Cascudo. Disponível em: < <http://mcc.ufrn.br/historico/>>. Acesso em: 27 jun. 2017.

A figura abaixo (Figura 2) mostra a representação de um artefato pertencente ao Museu Câmara Cascudo, sendo observado pelos alunos da disciplina Representação Descritiva II, do semestre 2017.1. Se trata de uma canoa indígena encontrada na lagoa de Extremoz, no Rio Grande do Norte, Brasil e que data provavelmente do século XVI. De acordo com Lins Júnior (2014, p.109) é “até o presente momento, o artefato náutico mais antigo do Brasil”. Além dos artefatos manufaturados pelo homem, os alunos tiveram a oportunidade de conhecer os objetos da natureza, a exemplo de esqueletos de animais, como o exemplo da imagem a seguir (Figura 3).



Figura 2- Canoa datada do séc.XVI-Museu Câmara Cascudo

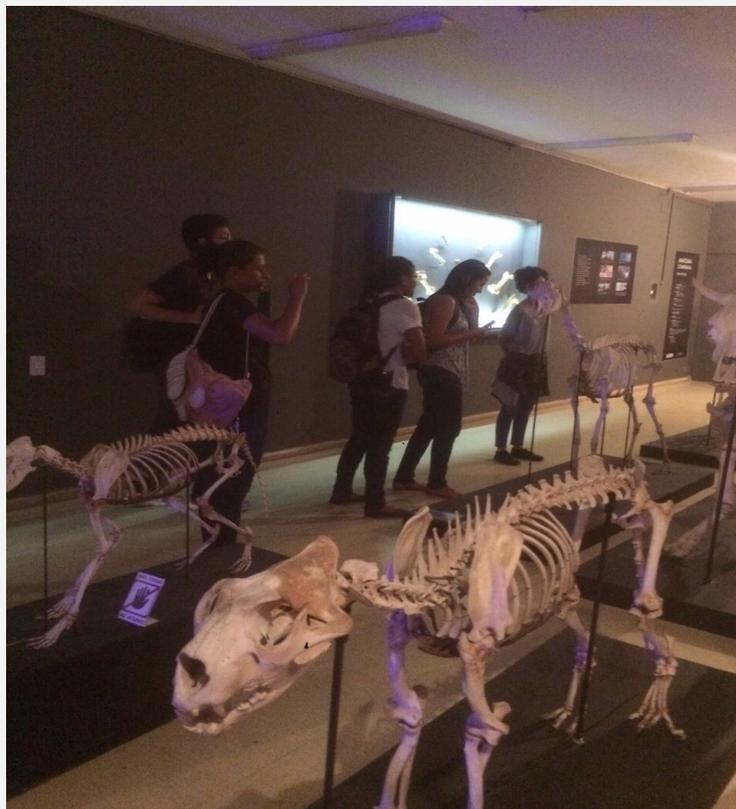


Figura 3 – Alunos observando acervo de realias – Museu Câmara Cascudo

Essa experiência de sair da sala de aula e visitar outros espaços de informação tem se mostrado muito salutar para a absorção de conceitos extremamente pertinentes como contexto informacional. A apresentação dos suportes informacionais está atrelada, portanto, a noção de que em contextos específicos os objetos não-livro são igualmente fontes de informação, seja pelo seu formato, ou pelo seu conteúdo, (desde que haja a intenção de aquisição de informação através desses objetos).

De acordo com Silva (2006, p.144) contexto, de uma forma geral é a “inter-relação de circunstâncias que acompanham um facto ou uma situação”. No entanto, buscando alinhar o conceito de contexto com a área de Ciência da Informação, numa perspectiva de um fenômeno info-comunicacional, Silva (2006, p.144) o define como

Uma unidade agregadora de elementos materiais (um edifício, um ou mais aposentos quaisquer que constitui cenário para a acção info-comunicacional), tecnológicos (mobiliário, material de escritório, computadores com ou sem ligação à Internet, etc.) e simbólicos (o estatuto e os papéis desempenhados pelas pessoas ou actores sociais) que envolvem o(s) sujeito (s) de acção info-comunicacional através de momentos circunstanciais delimitados cronologicamente (situação).

A partir dessa definição, percebe-se que o contexto, ou a situação irão influenciar diretamente no uso do objeto como recurso informacional. Dessa forma, a relevância em abordar na sala de aula o conceito de contexto, e aliar a este o complemento informacional, sendo portanto, o contexto informacional, reforça a ideia de ser uma unidade agregadora, como vimos na definição de Silva (2006), já que há uma variedade de materiais especiais vistos durante a condução da disciplina Representação Descritiva II. Tendo em vista as múltiplas possibilidades de agregar valor ao ambiente informacional através desses materiais especiais, a noção de contexto informacional irá motivar o nível de aprofundamento nas regras que regem a disciplina em tela, sendo mais ou menos profunda, adequando-se ao contexto.

Nesse sentido, podemos considerar que mesmo não havendo inicialmente a intenção de informar, como há nos documentos, as fontes de informação, ou objetos informacionais não-livro são artefatos culturais. (Martelete & Saldanha, 2016, p.80). Sem, contudo, aprofundar as questões referentes ao processo informacional como ação cultural (Smiraglia apud Martelete e Saldanha, 2016, p.80), o caminho discursivo que compõe a trajetória da disciplina, busca usar a noção de contexto para dar sentido aos objetos informacionais que não são livro (artefatos, realias, cartas, fotografias, partituras etc.).

Para tratar de partituras e recursos sonoros, temos um espaço de excelência, a Escola de Música da UFRN. Esse espaço abriga uma biblioteca especializada na temática (Biblioteca Pe. Jaime Diniz) e que semestralmente proporciona uma aula interativa com os alunos do curso de Biblioteconomia, sempre ministrada por bibliotecários. É um momento de aprendizado e encantamento com a profissão tendo em vista que a música “é formada por signos e símbolos que tornam seus conceitos teóricos, composições e partituras, representações informacionais e de conhecimento em linguagem de especialidade.” (Hipólito, Silva, 2012, p.311).



Figura 4 – Prática de catalogação de partituras Biblioteca Pe. Jaime Diniz-UFRN

Os resultados obtidos a partir da experiência relatada permitem afirmar que as estratégias adotadas em termos de ensino da catalogação de materiais especiais no âmbito da UFRN têm contribuído positivamente para despertar o interesse na disciplina. A metodologia adotada na disciplina tem sido apontada como inovadora e isso tem sido objeto de reconhecimento por meio dos relatos dos alunos expressos nas avaliações institucionais realizadas de forma periódica na UFRN. A forma do ensino adotada, ou seja a partir de uma forte ênfase num componente prático por meio das visitas de campo torna o processo mais dinâmico e atraente. Outros resultados deste trabalho se relacionam ao aspecto de que o trabalho pode servir de contributo para uma maior e necessária reflexão em torno do ensino da catalogação nos cursos de Biblioteconomia e para a formação e preparo do profissional da informação para atuar num mercado de trabalho em constante mutação.

CONCLUSÕES

O trabalho em tela visou propor um cotejo entre a teoria relacionada à área da Representação descritiva e a prática de sala de aula em forma de relato de experiências. Os resultados obtidos a partir da experiência relatada permitem afirmar que as estratégias adotadas em termos de ensino da catalogação de materiais especiais no âmbito da UFRN têm contribuído positivamente para despertar o interesse na disciplina. A metodologia adotada na disciplina tem sido apontada como inovadora e isso tem sido objeto de reconhecimento por meio dos relatos dos alunos expressos nas avaliações institucionais realizadas de forma periódica na UFRN. A forma do ensino adotada, ou seja a partir de uma forte ênfase num componente prático por meio das visitas de campo torna o processo mais dinâmico e atraente. Outros resultados deste trabalho se relacionam ao aspecto de que o trabalho pode servir de contributo para uma maior e necessária reflexão em torno do ensino da catalogação nos cursos de Biblioteconomia e para a formação e preparo do profissional da informação para atuar num mercado de trabalho em constante mutação.

Além destes aspectos o cotejo entre teoria e prática induz à abertura de novos diálogos com profissionais de áreas que manuseiam ou detêm material informacional, fortalecendo assim a área de organização do conhecimento como um todo. Os relatos de experiências na área do ensino de catalogação de materiais especiais se constituem em um importante recurso para a socialização de informação e metodologias adotadas nesta área que podem ser potencialmente aproveitadas em outros contextos com necessidades semelhantes.

REFERÊNCIAS BIBLIOGRÁFICAS

- Arilla, M.R.G. (1996). *Teoria y história de la Catalogación de documentos*, Madrid:editorial síntesis, 190p.
- Connelly, M & Clandinin, J. (1995). *Relatos de experiencia e investigacion narrativa*. In: Larrosa, J. Déjame que te cuente. Barcelona: Editorial Laertes, 1995.
- Ferreira, A. B. D. H. (1990). Novo dicionário da língua portuguesa. Nova fronteira.
- Gil, A. C. (2008). Métodos e técnicas de pesquisa social. In: *Métodos e técnicas de pesquisa social*. São Paulo: Atlas.
- Hipólito, Marta M. de Brito & Silva, Luceni Caetano da. (2012). Música: um arquivo documental e de memória auditiva. In: Albuquerque, M^a Elizabeth et AL (Org.). Representação da informação: um universo facetado. João Pessoa: Editora da UFPB. p.309-329.
- Lakatos, E. M. M., & Maria, E. (2009) Metodologia Científica. São Paulo, Atlas.
- Lins Júnior, H. M. M. (2014). *Arqueologia marítima: a evolução da canoa monóxila em Pernambuco, Brasil (séc. XVI-XX)*. Recife . 151f. Dissertação de mestrado, Programa de Pós-Graduação em Arqueologia – Universidade Federal de Pernambuco. Disponível em: <<http://repositorio.ufpe.br/handle/123456789/15320>> Acesso em: 27 jun. 2017.
- Marteleteo, R. & Saldanha, G. (2016). *Informação: qual o estatuto epistemológico?* In: Morigi, V.; Jacks, N & Golin, Cida (Org.). Epistemologias, comunicação e informação. Porto Alegre: Sulina, p. 69-90.
- Minayo, M.C. (1994). *O desafio do conhecimento*. São Paulo: Hucitec-Abrasco, 269p.
- Museu Câmara Cascudo. Disponível em: < <http://mcc.ufrn.br/>>. Acesso em: 27 jun. 2017.
- Novellino, M. S. F. (1996). Instrumentos e metodologias de representação da informação. *Informação & Informação*, 1(2), 37-45.
- Ortega, C.D. (2011). Do princípio monográfico à unidade documentária: exploração dos fundamentos da catalogação. *Liinc em Revista*, v.7, n.1, p.43-60, 2011.
- Ribeiro, Antonia de Castro M.M. (2015). Catalogação de recursos bibliográficos: pelo AACR2R em marc21. 6.ed. Brasília:Edição do autor.

Ribeiro, F. (2005). Organizar e representar informação: apenas um meio para viabilizar o acesso. *Revista da Faculdade de Letras. Ciências e Técnicas do Património*. Porto: Faculdade de Letras da Universidade do Porto, I Série, 4, 83-100.

Santos, P. L. V. A. C., & Corrêa, R. M. R. (2009). *Catálogo: trajetória para um código internacional*. Niterói: Intertexto.



COMPETÊNCIAS E COMPORTAMENTO INFORMACIONAL NA FORMAÇÃO CONTINUADA DE BIBLIOTECÁRIOS DO SISTEMA UNIVERSITÁRIO DE BIBLIOTECAS DA UNIVERSIDADE FEDERAL DA BAHIA

José Carlos Sales dos Santos¹, Sinara Abreu Alves²

¹Universidade Federal da Bahia, 0000-0003-1758-3639, jsalles@ufba.br,

²Universidade Federal da Bahia, naraabreu31@hotmail.com

RESUMO Objetivou analisar as opiniões de bibliotecários do Sistema Universitário de Bibliotecas da Universidade Federal da Bahia (SIBI/UFBA), Brasil, relativas aos incentivos institucionais de formação continuada promovidos pela Coordenação de Desenvolvimento Humano (CDH). Os cursos corresponderam às práticas de representação e organização do conhecimento de materiais bibliográficos da universidade, assim como atualizações de usos de recursos tecnológicos. Para cumprir o objetivo da investigação, delinearam-se os procedimentos metodológicos, com enfoque ao método de procedimento monográfico de caso ilustrativo (cursos de aperfeiçoamento), ao nível de pesquisa descritiva e ao questionário eletrônico como instrumento de coleta de dados, com enfoques às necessidades informacionais/tecnológicas dos sujeitos investigados ao desenvolvimento de competências básicas ao cumprimento das atividades do SIBI. Os resultados e as conclusões apontaram para a relevância dos cursos promovidos pela CDH/UFBA, orientados aos bibliotecários do Sistema, mas com ressalvas às adaptações de conteúdos dos cursos às especificidades informacionais dos profissionais envolvidos em labores inerentes à representação e organização da informação.

PALAVRAS-CHAVE *Competências profissionais, Comportamento informacional, Sistemas universitários de bibliotecas.*

ABSTRACT Aimed to analyze the views of librarians in the Sistema Universitário de Bibliotecas da Universidade Federal da Bahia (SIBI/UFBA), Brazil, concerning institutional incentives continued training promoted for Coordenação de Desenvolvimento Humano (CDH). The courses correspond to the practices of knowledge organization and representation of bibliographic materials from the University, as well as updates of uses of technological resources. To fulfill the purpose of the research, outlined the methodological procedures, focusing to the monographic illustrative case procedure (courses), to the level of descriptive research and electronic questionnaire as a tool for data collection, with informational needs/tecnológicas approaches of the subjects investigated the development of basic skills for the performance of activities of SIBI. The results and conclusions pointed to the relevance of courses promoted by the CDH/UFBA library system-oriented, but with caveats to adaptations of course content informational specificities of professionals involved in tasks relating to representation and organization of information.

KEYWORDS *Professional skills, informational Behavior, University Systems libraries.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A formação continuada procura a auxiliar os profissionais no processo de equacionar elementos práticos e teóricos em espaços de atuação. A partir desta perspectiva, pode-se dizer que a formação continuada confere acesso a novos conhecimentos, habilidades e atitudes profissionais sintonizadas com os novos paradigmas da sociedade contemporânea, provocando uma reflexão sobre os processos de trabalhos, que deve ser feita não de forma arbitrária, mas com autocrítica.

A história da educação continuada, de acordo com Destro (1995), surgiu da preocupação do capital com a formação de mão-de-obra qualificada, que atendesse as necessidades de sua base técnica em constante transformação face às novas tecnologias que a ela vem sendo incorporadas desde o século XIX. Um exemplo plausível desse processo é a ampliação e criação da rede pública de escolas na Europa e nos Estados Unidos nesse período com o intuito de adequar os trabalhadores às novas exigências da produção continuada.

Os recentes perfis dos profissionais determinados na contemporaneidade, de acordo com Oliveira et al (2000), enaltecem a criatividade, interatividade, flexibilidade e aprendizado contínuo. As habilidades e conhecimentos que são adquiridos ao longo da vida não superam a necessidade de uma educação permanente, voltada às mudanças reais ocorridas em seu ambiente profissional.

Neste sentido, Cunha (2000) entende que as categorias de profissionais que atuam com a informação estão conseqüentemente sob dois fatores primordiais: a) diversificação do mercado e funções relacionadas com a informação, já que muitos serviços de tratamento da informação se efetivaram com maior eficácia e precisão graças à introdução da informática e/ou uso do computador. b) a utilização das tecnologias informacionais nas tarefas diárias do profissional da informação reforça a idéia de inovação contínua das habilidades e conceitos já formados na graduação e nas experiências anteriores, trazendo alterações tanto nas rotinas quanto no seu perfil.

A sociedade exige dos profissionais a aprendizagem contínua, seja através da educação formal (cursos de pós-graduação, cursos de aperfeiçoamento, realizados fora do sistema de ensino), seja pela informal, em cursos particulares de tecnologias digitais, pesquisa autônoma na internet, acesso e manejo de aparelhos digitais como: smartphones (*download* e armazenagem de arquivos de mídias diversas). Tudo isso, para estar preparado para as novas demandas que a profissão exige.

As questões vivenciadas no cotidiano das bibliotecas e centros de documentação têm levado esses profissionais a buscarem o caminho da pesquisa, objetivando encontrar não só a melhor resposta para seus problemas, mas entendê-los, refletindo sobre eles, desmistificando-os e contextualizando às necessidades de suas bibliotecas, sabe-se, que se tem uma ação pensada para um SRI (Sistema de Recuperação da Informação) não se aplica de forma gratuita a outro, bastando apenas pequenas adequações.

A realidade não tem mostrado que as adequações guardam limites e contradições às vezes intransponíveis. A escolha e a compra de um novo software, assim como a definição das políticas de informação de um SRI deve ser pensada coletivamente pelos usuários do sistema, levando-se em conta aspectos como: os atributos da nova tecnologia (software) e o nível de descrição e análise da informação por ela suportada, caso contrário toda a filosofia de trabalho da unidade de informação fica comprometida. O profissional bibliotecário tem que retomar seu papel como intelectual, repensando

sua práxis e a função social de suas ações nas diferentes áreas de trabalho, em especial na área da representação descritiva dos documentos.

A educação continuada está inserida nas profissões de forma a implementar os conhecimentos atualizados na carreira dos profissionais, na Biblioteconomia não é diferente. Castro (2002) salienta que na reestruturação do curso de Biblioteconomia da Biblioteca Nacional em 1944, já havia uma preocupação em estabelecer a Educação Continuada através das modalidades nominadas como Curso fundamental de Biblioteconomia; Curso Superior de Biblioteconomia (CSB); Cursos Avulsos (CA).

Os Cursos Avulsos tinham a “finalidade de atualizar os conhecimentos dos bibliotecários e bibliotecário-auxiliares, divulgar conhecimentos sobre biblioteconomia e promover a homogeneidade básica dos serviços de bibliotecas” (Neves apud Castro, 2002, p.30). O que o autor considera ser o início da sistematização da educação continuada formal do bibliotecário. A necessidade de Educação Continuada foi aos poucos adquirindo nova realidade, sendo visualizada como um caminho para acompanhar as necessidades do mercado.

Entender a relevância da educação continuada é primordial para que a atuação do graduado torne-se integrada na vida do profissional, no momento em que esta passa a ser exercida, irá complementar a carreira do indivíduo, tão exigida na sociedade atual. Essa prática busca contribuir com possíveis distorções geradas pela formação inicial, como também aproximar-se do aprendizado inovador, alterado pelas transformações sociais, refazendo a forma de pensar, sentir e agir das novas gerações. Ruchinski (2009, p. 28) a define educação continuada como: “O processo contínuo de atualização, aperfeiçoamento, treinamento e aprimoramento das qualificações e habilitações individuais de cada profissional”. “A educação continuada pode ser definida como atividades educacionais que têm por objetivo atualizar e desenvolver o conhecimento e as habilidades profissionais, de forma a permitir ao profissional um melhor desempenho de sua função” (Crespo et al.,2006, p.3). Nesta pesquisa entende-se como Educação Continuada todo aprendizado adquirido nas diversas formas, realizado após a educação formal possibilitando o conhecimento atualizado do profissional.

Em relação aos formatos da Educação Continuada evidenciam-se, na literatura científica, diversas tipologias que podem ser desenvolvidos por iniciativas pessoais e/ou institucionais, e que evoluíram ao longo dos anos. Como formas de Educação Continuada, a leitura de livros e periódicos profissionais; cursos oferecidos em reuniões profissionais; estudos domiciliares ou individuais; pesquisa em Biblioteconomia e Documentação e visitas técnicas e estágios.

O bibliotecário enquanto profissional responsável por exercer essa atividade deve estar atento para acompanhar as atualizações dos estudos de representação e organização de informações e a Educação Continuada, na modalidade de eventos, pesquisas, reuniões associativo, ou cursos de atualização, dentre outras, é uma boa alternativa, pois possibilita ao profissional manter-se atualizado e melhorar a qualidade na realização da representação descritiva da informação. Diante do exposto o presente artigo retrata a importância da educação continuada do bibliotecário na área da Catalogação

No presente contexto, o profissional bibliotecário, responsável pela representação e organização de informações, ação cultural, serviços de referência, orientação de usuários e preservação da memória cultural e institucionai (documentação), constitui o mediador entre informação e usuário, passou a ser conhecido também como Profissional da Informação na Classificação Brasileira de Ocupações (CBO). Assim, os bibliotecários trabalham com o tratamento amplo da informação, e são responsáveis pela

mediação de conhecimento entre a sociedade e os conteúdos registrados, contribuindo para o desenvolvimento social, cultural, científico e tecnológico.

O profissional bibliotecário é também considerado atualmente como o Gestor da Informação, compreendido, conforme Gomes & Mota (2007, p. 3), como o “conjunto de processos capaz de garantir o uso efetivo da informação”. Na contemporaneidade, as técnicas bibliotecárias avançaram para cumprir as demandas da sociedade, favorecendo a segmentações da atuação profissional. Os profissionais apresentam “novos” desafios com o advento das tecnologias de informação e comunicação inscritas em unidades de informação, sempre orientados à organização, ao controle e disseminação e recuperação da informação. Esses profissionais vêm através dos tempos aprimorando técnicas de controle de registros do conhecimento e recuperação da informação.

A partir das considerações prévias, acerca da educação continuada e as funções administrativas e operacionais dos profissionais da informação, a presente comunicação procurou analisar as competências e comportamento informacional de bibliotecários vinculados ao Sistema Universitário de Bibliotecas da Universidade Federal da Bahia (SIBI/UFBA), Brasil, a partir dos cursos de extensão da Coordenação de Desenvolvimento Humano (CDH), como ‘Estratégias de Pesquisa em Bancos e Bases de Dados Científicos’, ‘Avaliação de metadados do Repositório Institucional da UFBA’ e ‘O SIBI nas trilhas do conhecimento e dos saberes’, orientados aos profissionais da informação. Para cumprir as análises, estruturam-se um comparativo da literatura relativa à temática ‘competências’ e ‘comportamento informacional’ com a realidade empírica investigada, em relação à organização do conhecimento.

A seção metodológica assumiu o método de procedimento monográfico (estudo de caso) e o nível da pesquisa descritivo; a técnica e o instrumento repusaram no questionário eletrônico focado aos bibliotecários do Sistema, com o intuito de delinear algumas percepções pertinentes aos cursos promovidos pelo CDH. Os resultados e as conclusões apontaram para a relevância dos referidos cursos, com ressalvas para o processo de adaptação dos conteúdos; os profissionais investigados admitem que carecem de informações específicas para responder às demandas e aos desafios de instituições e organizações, a exemplo do sistema universitário de bibliotecas analisado.

A presente comunicação compreende somente um aspecto da realidade pesquisada, inscrita em tempo e espaço específicos. Carece, portanto, de aprofundamento e ampliação do instrumento de pesquisa, servindo como insumo a novas pesquisas interessadas na referida temática.

METODOLOGIA

A comunicação assumiu o *método monográfico* (estudo de caso) devido à análise da percepção dos profissionais da informação (bibliotecários) lotados no SIBI/UFBA, considerando as características da população inscritas em tempo e espaço específicos de análise. Martins (2006, p.9) endossa o conceito de estudo de caso como uma possibilidade de inserção na realidade social, autorizada pelo levantamento amostral e pela avaliação restritamente quantitativa. As análises pertinentes ao caso ilustrativo precederam ao planejamento e à apropriação do referencial teórico, assim como a peculiaridades dos sujeitos investigados, inserindo-se na elaboração do ‘protocolo de aproximação’, com as principais atividades pertinentes ao estudo (introdução, desenvolvimento e conclusão). O referido protocolo

“[constituiu-se] em um conjunto de códigos, menções e procedimentos suficientes para se replicar o estudo, ou aplicá-lo em outro caso que mantém características semelhantes ao Estudo de Caso original.”.

O estudo de caso compromete-se com investigações de campo correspondentes a fenômenos contemporâneos da realidade. Entretanto, ao estabelecer os contornos de uma investigação científica, observa-se uma expressiva recorrência, na literatura de metodologia da pesquisa, categorias de estudos de ‘caso único’ e ‘casos múltiplos’. A primeira categoria corresponde a etapas particulares da unidade de análise, com contornos especiais para o processo do avanço da pesquisa. Há pesquisas que reservam diversos casos para análise, denominadas de ‘casos múltiplos’; categoria assumida na presente investigação. Os projetos de pesquisa que engendram as categorias de ‘caso único’ ou ‘casos múltiplos’ conferem a variantes inseridas na estrutura metodológica, e a maioria dos estudos que anuncia as categorias ‘únicas’ ou ‘múltiplas’ de análise não estabelece distinções relevantes.

O *nível de pesquisa* correspondeu ao descritivo. As pesquisas descritivas objetivam a circunscrição de determinadas populações ou fenômenos (bibliotecários do SIBI/UFBA), ou o estabelecimento de relações entre variáveis. Para Koche (2006, p.112), as variáveis constituem “[...] aspectos, propriedades, características individuais ou fatores, mesuráveis ou potencialmente mensuráveis, através dos diferentes valores que assumem, discerníveis em um objeto de estudo, para testar a relação enunciada em uma proposição.”.

As técnicas e os instrumentos de pesquisa procuraram coletar as opiniões dos profissionais da informação concernentes aos cursos de extensão do CDH, com o lastro do questionário eletrônico direcionadas aos sujeitos da observação. No entanto, com o avanço da pesquisa empírica, e as naturais redefinições das estratégias de coleta de dados, regularam-se os instrumentos da investigação para adequá-los à realidade observada, ajustando o questionário ao mecanismo de coleta de dados e informações.

A porcentagem de retorno constitui, aproximadamente, 41% dos bibliotecários lotados no SIBI, e os resultados geometrizados em gráficos. A coleta de dados ocorreu entre maio e junho de 2016, com aplicação do questionário estruturado em enunciados pertinentes às competências e aos comportamentos informacionais dos profissionais. A atividade posterior à coleta dos dados, os autores transcreveram e analisaram as respostas, buscando compreender e interpretá-las a partir da revisão de literatura.

RESULTADOS

Os cursos de extensão enfocados na formação continuada dos bibliotecários do SIBI/UFBA, segundo os respondentes, referem-se à orientação estritamente técnica relativa à complementação pedagógica suplementar, necessitando de adequações à concepção humanística e ética. Em contrapartida, evidenciaram-se que 44% dos entrevistados admitem que os citados cursos de extensão poderiam coadunar com as necessidades dos bibliotecários do Sistema, como uma pesquisa de sondagem/interesse elaborada pela superintendência.

A expressiva consideração pautada na avaliação dos cursos de aperfeiçoamento correspondeu aos 29% dos bibliotecários analisados, que aventam como ‘razoável’ os conteúdos abordados nos citados cursos para a execução das atividades e, assim, elucidar o comportamento informacional destes profissionais.

Apenas 3% dos profissionais investigados não souberam responder acerca da relevância das atividades de extensão no labor diário.

Cumprir às especificidades dos bibliotecários corresponde, na contemporaneidade, aos imperativos da formação continuada, associada às adequações das necessidades informacionais. Com o intuito de assegurar a formação qualificada de profissionais da informação, novas investigações deverão estruturar as análises ao comportamento informacional de usuários de conteúdos específicos. A seguir, o gráfico relativo à avaliação dos cursos de aperfeiçoamento do CDH:



Gráfico 1 – Avaliação cursos de aperfeiçoamento orientados aos seus profissionais da informação do SIBI/UFBA.

Os resultados apontam para expressivos percentuais concernentes à intenção dos profissionais em participar dos cursos ofertados pelo CDH, com 73% das respostas. Percebe-se um interesse contínuo em continuar nas atividades de aperfeiçoamento orientadas aos servidores do SIBI. Parece, no entanto, que a participação de bibliotecários nos referidos cursos estará condicionada aos ajustes de conteúdos, prerrogativas essenciais aos trabalhos desenvolvidos no SIBI, como observado no percentual de 21%.

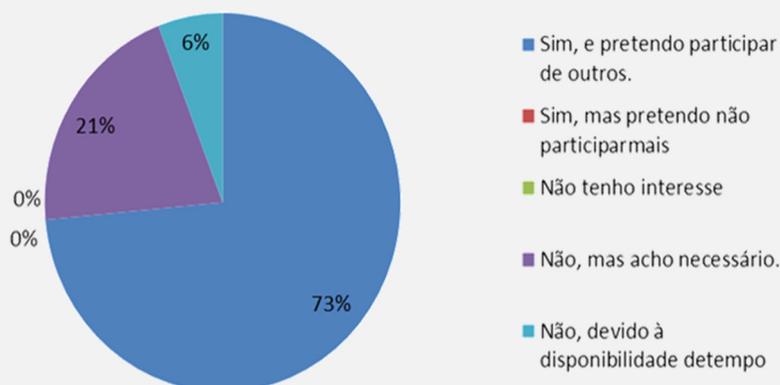


Gráfico 2 – Participação dos Bibliotecários nos cursos de aperfeiçoamento nos últimos três anos.

Assim, partir da análise integrativa dos dados obtidos na investigação, observou-se que a formação continuada orientada aos bibliotecários do SIBI/UFBA ainda carece de ajustes, apesar do contínuo interesse. Sociedade contemporânea exige dos profissionais informações e conhecimento atualizados para responder as demandas de instituições e organizações, como observadas nos gráficos supra apresentados.

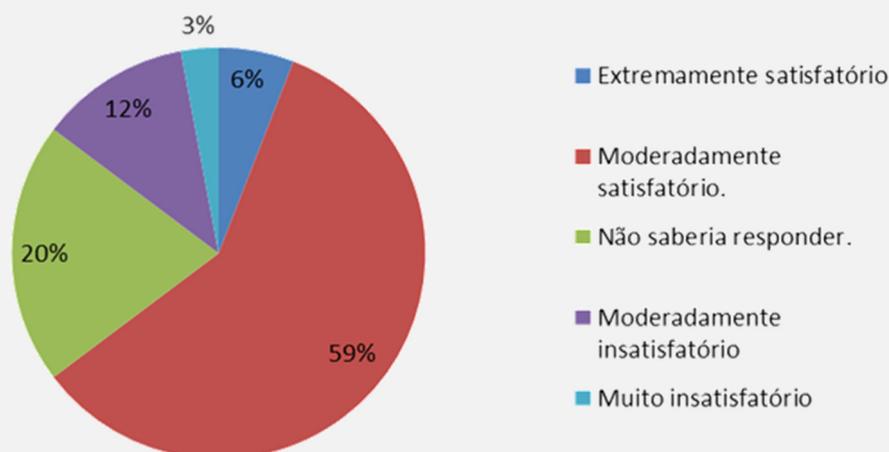


Gráfico 3 – A expectativa dos cursos oferecidos pela Coordenação de Desenvolvimento Humano (CDH/UFBA).

No Gráfico 3 não analisaremos com detalhes as questões respondidas, mas colocaremos os resultados em blocos, que *a priori*, dar-nos-iam uma possibilidade analítica mais prudente quanto aos reais objetivos de um curso de graduação, o grau de satisfação dos formandos junto aos objetivos e metas que uma graduação deve atender.

Portanto, o somatório do bloco fora dos padrões mínimos de satisfação para com o curso, considerando o “não saber responder” como possível não compromisso com todo o processo, temos 35% sem atingir as metas de formação que o mercado de trabalho exige. Isso sem falar que a grande maioria respondeu como ‘moderadamente satisfatório’. Respaldaado em avanços tecnológicos, nos efeitos da globalização e nas exigências profissionais, tornam-se relevante repensar o currículo dos cursos de formação técnica e intelectual para reduzir os riscos de obsolescência formativa.

Pelo que constatamos na investigação, as novas tecnologias da informação e comunicação causaram (e causarão) impactos diretos nas atividades bibliotecárias, como no processo de selecionar, organizar, representar, descrever, preservar, disseminar ou compartilhar a informação. Neste âmbito, Tarapanoff (1999, p. 32) anuncia que “no mundo virtual ou digital, elas terão outras aplicações além da tradicional atividade dentro de uma biblioteca, centro de documentação ou informação.”

A referida autora ainda complementa que, em um contexto atual,

[...] o profissional da informação deve buscar a sua identidade no novo mercado, sem perder de vista a sua característica mais intrínseca de responsável pelo ciclo documentário e informacional. Deve apossar-se de novos perfis, novas descrições de emprego, que sejam baseados nesta sua característica única. (Tarapanoff, 1999, p.11)

Ao que podemos contextualizar como uma sinalização direta que, a despeito das expectativas dos profissionais da informação, perante os cursos e ou disciplinas ofertados durante o percurso acadêmico,

estamos diante de cumprimento de necessidades relativas ao aprendizado de novos conceitos e à apropriação do letramento digital. O público inserido na realidade digital exigirá cada vez mais conhecimento, domínio e rapidez às suas demandas, buscando a autonomia para o traquejo dos dispositivos que os sistemas modernos oferecem.

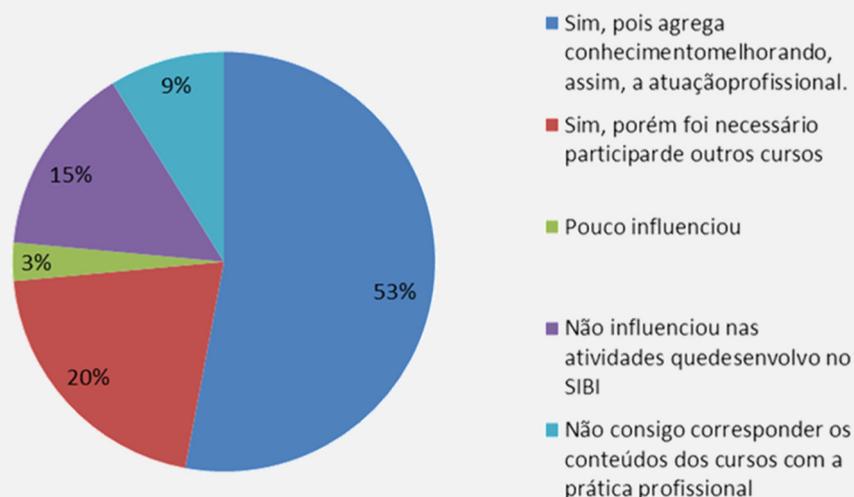


Gráfico 4 – A relação entre a participação de curso e a atuação profissional no SIBI/UFBA.

Os últimos resultados da pesquisa apontam a fundamental importância que os cursos voltados para o domínio e desenvolvimento das novas tecnologias agregam ao profissional da informação. Significativamente, 53% sinalizam a importância dos cursos de aperfeiçoamento e as consequências profissionais advindas dessa formação. Em contrapartida, 20% afirmam a necessidade de transcender os cursos oferecidos pelo Sistema, no intuito de apropriar-se de informações pertinentes às atividades desenvolvidas no pleito profissional. Assim, a busca da formação continuada independente constitui a tônica do interesse dos bibliotecários lotados no sistema universitário de bibliotecas analisado.

A partir da análise integrativa dos dados obtidos na investigação, perceberemos que a formação continuada disponibilizada pelo SIBI/UFBA ainda carece de ajustes. Sociedade contemporânea requer dos profissionais da informação conhecimento atualizados para responder as demandas de instituições e organizações. A mencionada sociedade tem se concentrado muito no indivíduo, no sujeito – compreendido como ator social –, que constrói, modifica, adapta e reproduz novos conhecimentos. E para isso ele agrega valores, apropria-se e se beneficia das tecnologias de informação.

CONCLUSÕES

Considerando a literatura a respeito da formação continuada enfocada aos profissionais da informação, contraposta à amostra empírica da investigação, evidenciaram-se que a necessidade de adaptações de conteúdos orientadas aos bibliotecários constitui a tônica de cursos de extensão estruturados pelo CDH. Os resultados da pesquisa apontaram para a fundamental relevância dos cursos enfocados ao domínio e ao desenvolvimento de competências profissionais. Com significado evidente, os profissionais sinalizam para essencialidade dos cursos de aperfeiçoamento na formação continuada, no comportamento direcionado à procura e à busca de informações que engendrem o aperfeiçoamento das atividades inscritas no sistema universitário de bibliotecas investigado.

A pesquisa também evidenciou uma relativa insatisfação com a modalidade de preparação técnica e intelectual, denominando-a incompleta e insuficiente ao profissional envolvido em cursos de aperfeiçoamento. As transformações decorrentes do surgimento das “novas” tecnologias da informação e comunicação, e da centralização paradigmática na produção extremamente capitalista, orientado ao lucro imediato em detrimento à sistematização do saber, parecem afetar significativamente as atividades de bibliotecários no aspecto de representação e organização de informações em sistemas computacionais.

Os resultados da pesquisa evidenciaram ainda necessidades e anseios dos profissionais em adquirir, ainda na base formativa (durante a graduação), habilidades e competências satisfatórias ao exercício de atividades profissionais. Urge, contudo, uma reconfiguração curricular e extensiva que possibilite aos profissionais conhecimentos como: estratégias de pesquisa e recuperação da informação em bancos e bases de dados, analista de informação da *web*, avaliador de fontes de informação digitais, serviços de referência especializados e estratégias organizacionais no âmbito informacional.

Os fatores limitantes corresponderam a pouca adesão dos profissionais (sujeitos da investigação) em responder ao questionário, ou a demora em responder ao solicitado. Mas a persistência dos autores garantiram respostas fidedignas às solicitações, autorizando a execução da pesquisa. Contudo, com as análises integrativas dos resultados logrados na pesquisa, inferem-se que a formação continuada disponibilizada ao SIBI/UFBA carece de ajustes.

Em relação às limitações da presente investigação, consideramos que a pesquisa deveria comportar a totalidade (censo) dos bibliotecários lotados no sistema universitário de bibliotecas da UFBA para evidenciar resultados mais pertinentes à realidade analisada, assim como a efetivação de conhecimento logrado em cursos oferecidos pelo CDH/UFBA. Desta forma, propomos também uma agenda de pesquisa que comporte entrevistas com servidores-chave lotados na referida coordenação com o intuito de analisar a percepção que eles têm em relação aos cursos de aperfeiçoamento propostos pela superintendência do SIBI e servidores envolvidos nas atividades basilares das bibliotecas universitárias e setoriais.

Assim, os profissionais precisam de informações específicas para responder às demandas e aos desafios de instituições e organizações, desenvolvendo competências e habilidades em organização do conhecimento. A sociedade contemporânea exige dos sujeitos comportamento proativo e interesse em construir novos conhecimentos. É fundamental analisar e reconfigurar, quando necessário, a estrutura de conteúdos inscritos nos cursos de extensão promovidos pelo CDH.

REFERÊNCIAS BIBLIOGRÁFICAS

- Castro, C. A. (2002). Histórico e evolução curricular na área de Biblioteconomia no Brasil In. VALENTIM, Marta Lígia. *Formação do profissional da informação*. São Paulo: Polis.
- Crespo, I. M., RODRIGUES, A. V. F. & MIRANDA, C. L. (2006). Educação continuada para bibliotecários: características e perspectivas em um cenário de mudanças. *Biblios*, v.7, n.25-26, p.1-14, jul.
- Cunha, M. B. (1984). O desenvolvimento profissional e a educação continuada. *Revista de Biblioteconomia de Brasília*, Brasília, v.12, n. 2, p. 145-156, jul/dez.

Destro, M. R. P. (1995). *Educação Continuada: visão histórica e tentativa de conceitualização*. Campinas: CEDES/PAPIRUS, Cadernos CEDES, n. 36, p. 37-45.

Gomes, M. A. & MOTA, F. R. L. (2007). *Gestão da informação no contexto organizacional*. Maceió.

Köche, J. C. (2006). *Fundamentos de metodologia científica*. 23.ed. Petrópolis: Vozes.

Martins, G. A. (2006). *Estudo de caso: uma estratégia de pesquisa*. São Paulo: Atlas.

Oliveira, M.A., GOMES FILHO, A.C., HONESKO, A., & PEREIRA, E.C. (2000). Gerenciamento do capital humano em bibliotecas ou centros de informação: desafio imposto pela sociedade do conhecimento. *Transinformação*, v.12, n.2.

Ruchinski, A. L. (2009). *Capacitação e atualização do bibliotecário: estudo de caso da Biblioteca Demonstrativa de Brasília*. 2009.55f. Monografia (Bacharelado em Biblioteconomia) – Universidade de Brasília, Brasília.

Tarapanoff, K. (1999). (Org.) *Inteligência Organizacional e Competitiva*, Ed. Unb Brasília.



GESTÃO DO CONHECIMENTO: FERRAMENTAS PARA RETENÇÃO E DISSEMINAÇÃO DO CONHECIMENTO ORGANIZACIONAL NO TRIBUNAL REGIONAL FEDERAL DA PRIMEIRA REGIÃO

Marcia Mazo Santos¹, Marília de Souza de Mello², Márcio da Silva Albuquerque³

¹Diretora da Biblioteca do Tribunal Regional Federal da Primeira Região,
0000-0003-0561-0783, marciamazo@gmail.com

²Supervisora da Seção de Atendimento ao Usuário da Biblioteca do Tribunal Regional Federal
da Primeira Região, 0000-0002-9559-2277, marilia.mello@trf1.jus.br

³Diretor do Centro de Desenvolvimento e Aperfeiçoamento de Servidores do Tribunal Regional Federal
da Primeira Região, 0000-0001-5476-2834, marcio.albuquerque@trf1.jus.br

RESUMO Objetivos: O Tribunal Regional Federal da Primeira Região, no Brasil, abrange 14 unidades da federação, 13 Estados (Acre, Amapá, Amazonas, Bahia, Goiás, Minas Gerias, Mato Grosso, Maranhão, Pará, Piauí, Rondônia, Roraima, Tocantins) e o Distrito Federal e está diante da saída de grande parte da sua força de trabalho em função da proximidade de aposentadorias de servidores. Isso representa expressiva perda de conhecimento tácito: experiências e habilidades que fomentam o desempenho da instituição. Com as experiências adquiridas no exercício de suas atividades profissionais, os servidores passaram a deter todo o conhecimento e a inteligência do negócio, tornando-se parte da própria identidade e cultura organizacionais. Destaca-se, ainda, que o gerenciamento desses ativos é fundamental para a sustentabilidade da instituição. **Método:** O modelo SECI de conversão do conhecimento de Nonaka e Takeuchi foi utilizado na conversão do conhecimento organizacional assegurando aos servidores a utilização de ferramentas específicas para que pudessem transmitir, formalmente e de maneira sistematizada, suas habilidades e conhecimentos para outros servidores. De forma integrada ao SECI, foram utilizadas três ferramentas de gestão do conhecimento: o mapeamento do conhecimento crítico dos processos de trabalho; a *storytelling*; e o registro de boas práticas. **Resultados:** O mapeamento do conhecimento crítico identificou lacunas nos processos de trabalho; foi desenvolvida uma base de dados de boas práticas utilizando a plataforma *Dspace*; e as primeiras narrativas estão sendo gravadas com servidores em vias de aposentadoria. **Conclusão:** As medidas de registro de boas práticas e identificação do conhecimento crítico inerente aos processos de trabalho, disponíveis de maneira organizada, podem trazer contribuições para criação de conhecimento organizacional novo, seu registro, compartilhamento, e disseminação da memória organizacional.

PALAVRAS-CHAVE *Gestão do conhecimento, Conhecimento organizacional, Mapeamento do conhecimento crítico, Storytelling, Boas práticas.*

ABSTRACT Objectives: The Federal Regional Court of the First Region, in Brazil, covers 14 federal units, 13 states (Acre, Amapá, Amazonas, Bahia, Goiás, Minas Gerias, Mato Grosso, Maranhão, Pará, Piauí, Rondônia, Roraima, Tocantins) and the Federal District. It now faces the departure of a large number of its workforce due to their retirement application. This represents an expressive loss of tacit knowledge: experiences and skills that foster the institution's performance. The servants hold all the knowledge and the intelligence of the business with the experiences acquired in their professional activities, becoming part of the organizational identity and culture. It should also be pointed out that the management of these assets is

essential for the sustainability of the institution. **Method:** Nonaka and Takeuchi knowledge conversion (SECI model) was used in the conversion of organizational knowledge, assuring to servants the use of specific tools so that they could share, in a systematic way, their skills and knowledge to other servers. Integrated to SECI, three knowledge management tools were used: critical knowledge of work processes mapping; storytelling; and the best practices recording. **Results:** The critical knowledge mapping identified gaps in work processes; DSpace platform was developed using the best practices databasis; and the first narratives are being recorded with servers in process of retirement. **Conclusion:** Measures to register best practices and identify the critical knowledge inherent to the work processes, available in an organized way, can contribute to the creation of new organizational knowledge, its registration, sharing, and dissemination of organizational memory.

KEYWORDS *Knowledge management, Organizational knowledge, Critical knowledge, Storytelling, Best practices.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

1. INTRODUÇÃO

A estrutura do Poder Judiciário no Brasil é constituída de diversos órgãos que funcionam no âmbito da União e dos estados, incluindo o Distrito Federal. No campo da União, o Poder Judiciário conta com a Justiça Federal (comum) e as Justiças Especializadas (Justiça do Trabalho, a Justiça Eleitoral e a Justiça Militar). Apresenta, ainda, os tribunais superiores (Superior Tribunal de Justiça, Tribunal Superior do Trabalho, Superior Tribunal Militar e Tribunal Superior Eleitoral) e acima de todos esses está o Supremo Tribunal Federal, Corte guardiã da Constituição. O Tribunal Regional Federal da Primeira Região (TRF1) compõe a Justiça Federal, instituição que é o objeto de estudo de caso deste trabalho, que abrange os seguintes Estados: Acre, Amapá, Amazonas, Bahia, Goiás, Maranhão, Mato Grosso, Minas Gerais, Pará, Piauí, Rondônia, Roraima, Tocantins e o Distrito Federal.

O Conselho Nacional de Justiça (CNJ), órgão que visa aperfeiçoar o trabalho do sistema judiciário brasileiro, instituiu em 2014, a Estratégia Nacional do Poder Judiciário 2015-2020 (<http://www.cnj.jus.br/gestao-e-planejamento/2015-2020>), da qual constam os macrodesafios.

No macrodesafio *Instituição da governança judiciária*, está estabelecido o seguinte:

“Formulação, implantação e monitoramento de estratégias flexíveis e aderentes às especificidades regionais e próprias de cada segmento de justiça, produzidas de forma colaborativa pelos órgãos da justiça e pela sociedade. Visa à eficiência operacional, à transparência institucional, ao fortalecimento da autonomia administrativa e financeira do Poder Judiciário e à adoção das melhores práticas de comunicação da estratégia, de gestão documental, da informação, de processos de trabalho e de projetos.”

Tais proposições contribuíram para a reflexão sobre algumas questões relacionadas à gestão e seus possíveis desdobramentos.

Nos últimos anos, o Tribunal tem enfrentado o processo da saída de grande parte da sua força de trabalho em função das aposentadorias de servidores e da nova reforma da previdência, objeto de discussão, atualmente, no Brasil. Isso representa expressiva perda de conhecimentos, experiências e habilidades que fomentam o desempenho da instituição.

É necessário observar, também, que o envelhecimento da força de trabalho, supõe outra ação: o planejamento da sucessão, ou seja, a formação das futuras gerações. Com as experiências adquiridas no exercício de suas atividades profissionais, os servidores passaram a deter todo o conhecimento e a inteligência do negócio, tornando parte da própria identidade e cultura organizacionais.

Para Moresi (2001) a gestão do conhecimento é um conjunto de atividades que busca desenvolver e controlar todo tipo de conhecimento em uma organização visando à utilização na consecução de seus objetivos. Como ferramenta de apoio à gestão, ainda segundo Moresi (2001) “é preciso estabelecer políticas, procedimentos e tecnologias que sejam capazes de coletar, distribuir e utilizar efetivamente o conhecimento, representando fator de mudança no comportamento organizacional”.

O compartilhamento do conhecimento da organização é um processo dinâmico que depende do relacionamento das pessoas, quer para a sua criação, quer para a sua manutenção e desenvolvimento. (Nonaka & Takeuchi, 1997).

Considerando, portanto, que a gestão do conhecimento traz transparência institucional; que o TRF1, não havia, até o ano de 2016, estabelecido um programa para registrar o conhecimento que está com as pessoas no desempenho de seus processos de trabalho; que uma grande parte de servidores deste Tribunal está em vias de aposentadoria; que uma base de dados de boas práticas contribui, para a instituição reduzir tempo em seus processos, e, conseqüentemente cortar custos; que nos macrodesafios constantes no planejamento estratégico do Poder Judiciário para 2015-2020 está contemplada a adoção de melhores práticas, foi elaborada exposição de motivos para o Presidente do Tribunal, encaminhando três propostas para utilização de ferramentas de gestão do conhecimento: (1) a criação de uma base de dados de boas práticas, (2) a utilização da *storytelling* para registrar o conhecimento dos servidores em vias de aposentadoria e (3) o mapeamento do conhecimento crítico dos processos de trabalho.

2. O TRIBUNAL REGIONAL FEDERAL DA PRIMEIRA REGIÃO

A atual Constituição, promulgada em 5 de outubro de 1988, criou cinco Tribunais Regionais Federais – TRFs.

Atendendo aos preceitos constitucionais, o Tribunal Federal de Recursos editou a Resolução n. 1, de 6 de outubro de 1988, fixando as sedes e as jurisdições dos cinco TRFs. Assim, ficou estabelecido que o TRF- 1ª Região teria sede em Brasília e jurisdição nos Estados do Acre, Amapá, Amazonas, Bahia, Goiás, Maranhão, Mato Grosso, Minas Gerais, Pará, Piauí, Rondônia, Roraima, Tocantins e no Distrito Federal, a 2ª Região no Rio de Janeiro (RJ), a 3ª Região em São Paulo (SP), a 4ª Região em Porto Alegre (RS), e a 5ª Região em Recife (PE).

A Justiça Federal é regulamentada pela Lei n.5. 010 de 30 de maio de 1966.

3. REFERENCIAL TEÓRICO

3.1 GESTÃO DO CONHECIMENTO ORGANIZACIONAL

As organizações são organismos vivos, feitas de elementos com funções variadas que contribuem para o todo e para a coletividade. Segundo Allee (c2003) é um “sistema social complexo e adaptativo onde as pessoas cooperam sistematicamente para alcançar um objetivo comum”.

Para Moresi (2001) a gestão do conhecimento é um conjunto de atividades que busca desenvolver e controlar todo tipo de conhecimento em uma organização visando à utilização na consecução de seus objetivos. Como ferramenta de apoio à gestão, ainda segundo Moresi (2001) “é preciso estabelecer políticas, procedimentos e tecnologias que sejam capazes de coletar, distribuir e utilizar efetivamente o conhecimento, representando fator de mudança no comportamento organizacional”.

O compartilhamento do conhecimento da organização é um processo dinâmico que depende do relacionamento das pessoas, quer para a sua criação, quer para a sua manutenção e desenvolvimento. (Nonaka & Takeuchi, 1997).

Para eles, existem quatro modos de conversão do conhecimento, o conhecido modelo SECI:

Socialização: de conhecimento tácito em conhecimento tácito. É o processo de compartilhamento das experiências, tais como modelos mentais ou habilidades técnicas compartilhadas. O aprendizado dos alunos, não por meio da linguagem de seus mestres, mas sim por meio da observação, imitação e prática, constitui-se numa forma de socialização. O segredo para a aquisição do conhecimento tácito é transmissão da experiência.

Externalização: de conhecimento tácito em conhecimento explícito. É um processo de articulação do conhecimento tácito em conceitos explícitos. É tido como o processo perfeito de criação do conhecimento, considerando-se que a forma explícita é expressa por metáforas, analogias, conceitos, hipóteses ou modelos. A escrita é uma forma de converter o conhecimento tácito em conhecimento articulável. A externalização da conversão do conhecimento é provocada pelo diálogo ou pela reflexão coletiva.

Combinação: de conhecimento explícito em conhecimento explícito. Esse modo de conversão, como os anteriores, utiliza-se do diálogo e envolve a combinação de conjuntos diferentes de conhecimento explícito, sendo realizado por meio da troca ou combinação de conhecimento como documentos, reuniões, conversas ao telefone ou redes de comunicação computadorizadas.

Internalização: de conhecimento explícito em conhecimento tácito. O processo de incorporação do conhecimento explícito em tácito consiste em aprender fazendo.

Pode-se observar que a socialização pode gerar o conhecimento compartilhado; a externalização gera conhecimento conceitual; a combinação dá origem ao conhecimento sistêmico; a internalização produz conhecimento operacional.

3.2 BOAS PRÁTICAS

Muita informação e conhecimento importantes ficam dispersos nas organizações, em documentos impressos, em bancos de dados ou mesmo na mente das pessoas, dificultando o acesso e a disseminação em momentos oportunos e, conseqüentemente, a melhoria de processos. O valor deste conhecimento aumenta à medida que ele se torna mais acessível. É muito comum resolver um problema várias vezes, partindo do zero, desperdiçando esforços pelo não compartilhamento de soluções já criadas.

O estudo da transferência de boas práticas tem sido um tópico na pesquisa nos últimos anos (Prochno, 2004). Isso ocorreu em função de fenômenos como o aumento da importância das redes organizacionais, a expansão geográfica das organizações, e o uso disseminado do benchmarking como uma fonte de melhoria do desempenho.

Vários estudos (Huysman & Wit, 2004; Jen-te Yang, 2007; Pereira, Silva, Freitas & Oliveira, 2016; Sabbadini, Chamovitz & Desirério, 2016; Ribeiro, Soares, Jurza, Ziviani & Neves 2017) indicam que compartilhar conhecimento resulta na melhoria da eficiência organizacional.

A essência em identificar e compartilhar as boas práticas está em aprender com os outros e reusar o conhecimento evitando o desperdício. O compartilhamento efetivo das boas práticas pode ajudar as organizações a: identificar e substituir uma prática pouco produtiva; melhorar o desempenho de trabalhadores colocando-os ao lado de trabalhadores com bom desempenho; evitar que a roda seja reinventada; minimizar o retrabalho; economizar custos por meio de uma melhor produtividade e eficiência; e melhorar os serviços (Ramos & Teixeira, 2016).

Uma base de dados de boas práticas pode contribuir para a economia de custos por meio do compartilhamento de boas práticas; melhoria na tomada de decisão a partir do aconselhamento obtido com colegas de outros setores; aumento da receita através da partilha de conhecimentos e produtos; inovação através da combinação e polinização ideias; reforço da capacidade de ação coletiva que envolve unidades dispersas.

3.3 STORYTELLING

Storytelling significa a capacidade de contar histórias e narrativas relevantes e que expressem não só as vivências, mas também o processo evolutivo de uma organização.

É o uso de narrativas de histórias para compartilhamento de conhecimentos de uma forma mais significativa que estimula a curiosidade de quem recebe a informação. É uma forma de trocar informação e gerar conhecimento. Em síntese, torna o novo mais tangível, clarifica as ideias e engaja as pessoas no projeto em que está atuando (Jenkins, 2006).

Os métodos mais comuns para este processo são a utilização de palavras ou recursos audiovisuais para transmitir a história. Neste quesito e objetivando o registro formal, as histórias não devem ser contadas de improviso, devem seguir uma metodologia minimamente capaz de estabelecer critérios inteligíveis e também que retratem efetivamente o que se quer contar. É uma ferramenta importante para o processo de aprendizagem, repasse de expertise, registro de processos inovadores, transmissão de elementos culturais, regras e valores éticos, dentre outros aspectos.

O *storytelling* é uma forma inovadora de registro das atividades e da própria história da organização, podendo utilizar ferramentas disponíveis no mercado ou mesmo por meio de desenvolvimento interno. Destaca-se seu papel na motivação dos colaboradores internos e na divulgação da imagem institucional ao público externo, impulsionando a espiral do conhecimento (Menezes, Johann, Valentim & Scott 2017).

Para Nassar e Goco (2013) a ideia é que a narrativa organizacional ganha atratividade e legitimidade com a contação de histórias ou *storytelling*.

3.4 MAPEAMENTO DO CONHECIMENTO CRÍTICO

O conhecimento crítico é aquele considerado necessário e fundamental para a tomada de decisões e resolução de problemas e que deve ser capitalizado, ou seja, transformado em conhecimento institucional.

De acordo com Fukunaga (2015), para elaborar um mapa é necessário percorrer quatro passos: identificação, priorização, caracterização e o plano de ação.

O primeiro passo, a identificação, é a fase onde será observado qual o tipo de conhecimento, se explícito ou tácito, como está armazenado e, se for o caso, quem são os detentores do conhecimento. Essa fase serve para colher onde e como os colaboradores fazem para adquirir conhecimento necessário para realizar suas atividades.

A priorização, que vem em seguida, vai selecionar alguns conhecimentos identificados. Para isso, devem ser criados critérios baseados na criticidade do conhecimento para a organização, a fim de dar sequência ao mapeamento.

O terceiro passo faz a convergência do conhecimento mapeado com a metodologia da Gestão do Conhecimento. A caracterização é o registro das ações em torno do conhecimento gerado. É necessário fazer uma entrevista com os desenvolvedores para detalhar o conhecimento em conjunto com a área responsável pelo registro. Também, nessa fase, será feita a análise da situação atual para a situação desejada.

Na figura 1, adaptada pelos autores, Fukunaga (2015) resume a importância de se mapear o conhecimento crítico de uma organização.

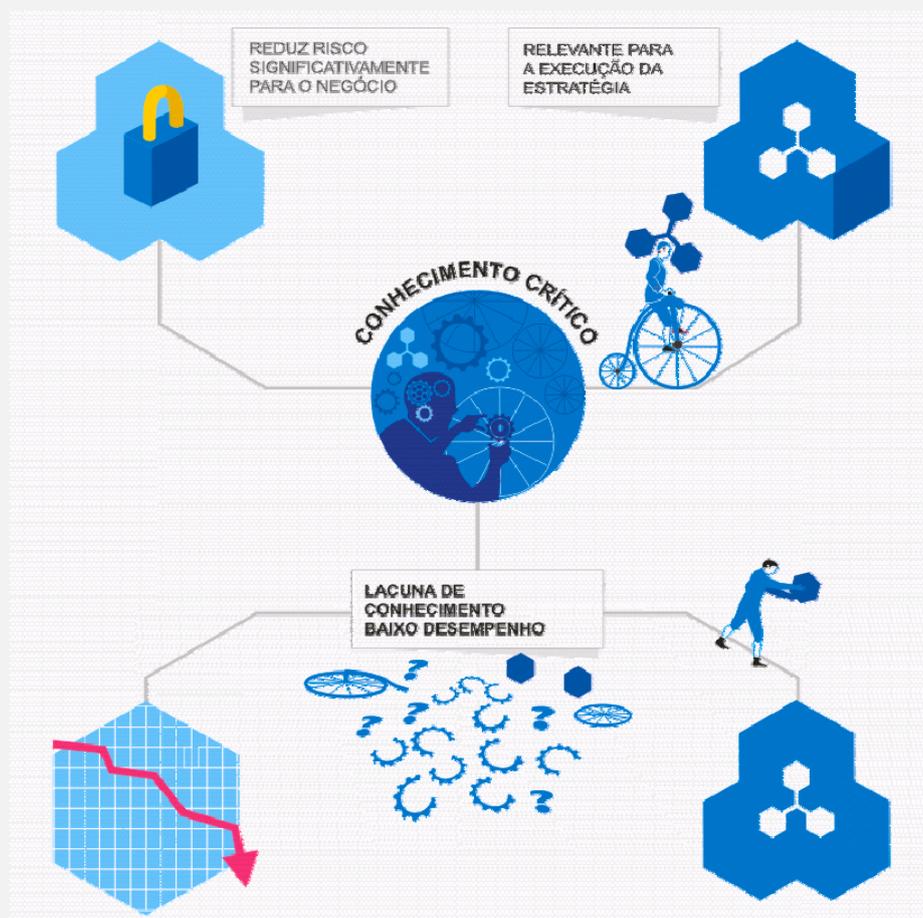


Figura 1. Conhecimento Crítico

Fonte: Fukunaga (2015) Adaptada pelos autores

A visão macro e o diagnóstico do conhecimento crítico vão direcionar a criação do plano de ação que irá nortear o desempenho esperado de acordo com a metodologia. O conhecimento crítico deve estar alinhado à estratégia para o melhor desempenho organizacional, pois lacuna de conhecimento impacta em desempenho organizacional.

4. METODOLOGIA

O referencial teórico utilizado visando cumprir as atividades propostas, a saber: mapeamento do conhecimento crítico, retenção do conhecimento por meio de *storytelling*, criação de base de dados de boas práticas, foi o modelo SECI de conversão do conhecimento de Nonaka e Takeuchi (1997) (Figura 2).

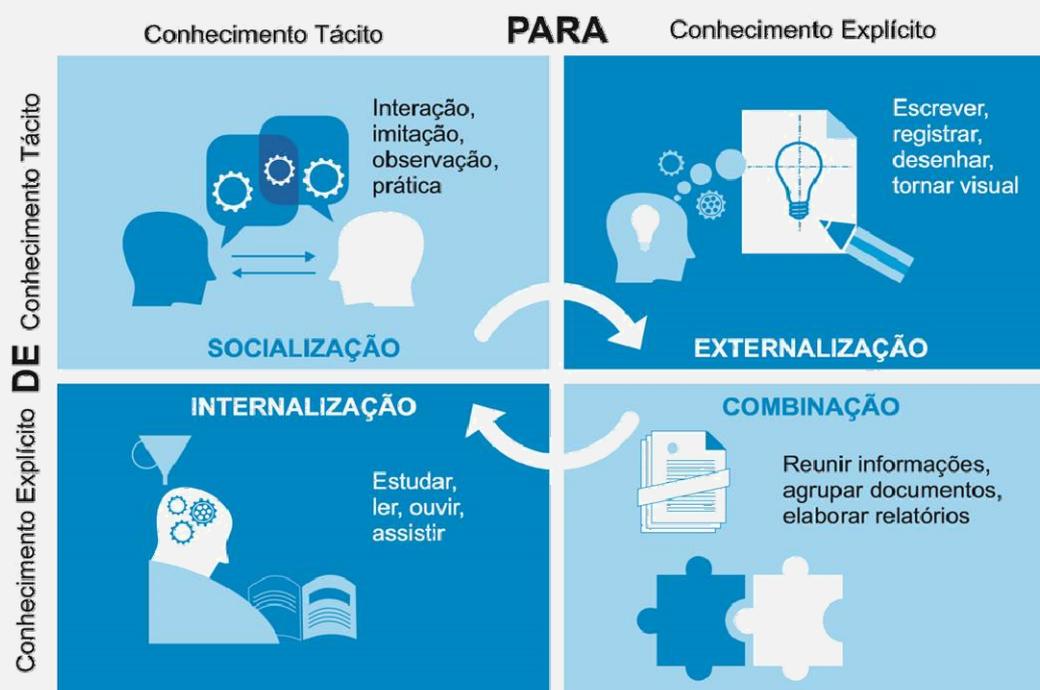


Figura 2. Modelo SECI

Fonte: Nonaka e Takeuchi (1997) Adaptada pelos autores.

Na abordagem do compartilhamento do conhecimento, buscou-se respaldo no suporte teórico dado pela Gestão do Conhecimento que consiste no “processo sistemático de identificação, criação, renovação e aplicação dos conhecimentos que são estratégicos na vida de uma organização” (Santos, Pacheco, Pereira & Bastos Jr., 2001).

Segundo o modelo SECI os conhecimentos identificados que devem suprir essas lacunas devem ser socializados (S), e o método encontrado para isso foi o Café com conhecimento abordado no item 7.

O evento Café com conhecimento, composto por 4 (quatro) painéis, divididos em 4 (quatro) sextas-feiras, uma por mês, de agosto a novembro do ano de 2016, não só permitiu a divulgação dos aspectos práticos da Gestão do Conhecimento Organizacional no âmbito da Justiça Federal da 1ª Região, Tribunal, seções e subseções judiciárias, como também a troca de conhecimento em ambiente informal e descontraído propício para essa interação. As seções e subseções judiciárias puderam acompanhar o evento por meio de videoconferência em tempo real.

Especificamente, observou-se que as condições favoráveis ao contexto compartilhado de conhecimento em uma organização estão condizentes com o modelo SECI (Socialização, Externalização, Combinação e Internalização).

No mapeamento do conhecimento crítico foram encontradas algumas lacunas de conhecimento, tanto nos processos de trabalho, como também a identificação de servidores em vias de aposentadoria.

O quadrante de explicitação (E) do Nonaka inspirou para que o conhecimento explicitado no Café se transformasse em *storytelling* – registrando as experiências dos servidores em vias de aposentadoria, e combinados (C) com o registro dos processos de trabalho em uma base de dados de boas práticas.

Esse conjunto de conhecimento permite que servidores e magistrados acessem e internalizem-no (I) propiciando a geração de novos momentos de socialização fazendo a espiral sair do pessoal, para o grupo, para as unidades, e conseqüentemente para a organização.

5. O USO DA INFORMAÇÃO COMO CONHECIMENTO ORGANIZACIONAL

Diversos fatores impedem uma maior produtividade nos processos de trabalho: o conhecimento retido na mente das pessoas, volumes imensos de documentos existentes no Tribunal e que escondem informações importantes, erros que são cometidos por não se considerarem experiências passadas, e gestores que gastam tempo em busca de informação para execução do trabalho.

Foi identificada a necessidade de mapear o **conhecimento crítico** dos processos de trabalho. O foco deste procedimento é a compreensão e entendimento sobre quais conhecimentos possuem relevância estratégica e complexidade de aquisição e retenção no âmbito da organização. Uma prerrogativa foi entender e estudar o planejamento estratégico, ressaltando quais conhecimentos eram relevantes para cada etapa. Como o mapeamento é atividade que depende intrinsecamente das áreas a serem entrevistadas, fez-se necessária a sensibilização por meio de contato direto com os dirigentes das áreas a serem trabalhadas. Analisado o processo de trabalho e sua correspondente cadeia de valor, passou-se a enumerar e avaliar os resultados, entregas, riscos e criticidades.

No mapeamento do conhecimento crítico identificamos, além das lacunas nos processos de trabalho, servidores em vias de aposentadoria, cujos conhecimentos se perderiam se não fossem registrados, e a ferramenta escolhida foi *storytelling*.

Foi identificado que uma **base de dados de boas práticas** contribui para a retenção do conhecimento e pode reduzir tempo em seus processos de trabalho, incorporar conhecimentos, transformar o aprendizado individual em organizacional e, conseqüentemente, cortar custos. Funciona como um catalisador para a aprendizagem e ajuda na construção da memória organizacional. Esse instrumento preenche a lacuna na gerência da informação, visto que nele é possível registrar, organizar e compartilhar tais conhecimentos institucionais.

6. RESULTADOS

Todos os registros estão no repositório institucional na biblioteca digital do Tribunal que utiliza o *DSpace*, software gratuito desenvolvido pelo *Massachusetts Institute of Technology* – MIT <http://www.trf1.jus.br/dspace>.

Assim, objetivando corrigir este *gap* na gestão do conhecimento do TRF1 foi proposto:

- 1) Criação do grupo de trabalho multidisciplinar para definição das estratégias de implementação da gestão do conhecimento e da informação no âmbito do TRF1 com a seguinte composição: representantes das áreas de biblioteca, capacitação, gestão estratégica e inovação, e tecnologia da informação;
- 2) Resultados esperados: criação da Base de Boas Práticas disponibilizado no portal do TRF1, e definição de critérios para inserção das práticas no repositório;

- 3) Cronograma de viabilização: até 30 de agosto de 2016, realização de reunião presencial; até 30 de setembro de 2016, proposta de metodologia, plano de capacitação e de reuniões; até 30 de outubro de 2016, definição de metodologia e plano de ação; e até o final de dezembro de 2016, divulgação do cronograma para execução em 2017.
- 4) Unidades envolvidas: a alta administração, área de biblioteca, área de capacitação e desenvolvimento, área de informática e representante de uma seccional.

Vários atos administrativos foram editados, objetivando a garantia do comprometimento de servidores e magistrados, bem como, a assertividade na consecução dos objetivos traçados.

7. CONCLUSÃO

O cronograma estabelecido pela alta administração para o desenvolvimento das atividades de mapeamento do conhecimento crítico, identificação dos servidores em vias de aposentadoria, criação de uma base de dados de boas práticas, e criação do evento Café com conhecimento foi cumprido para o ano de 2016. Foram identificadas lacunas de conhecimento no desenrolar do mapeamento do conhecimento crítico, foi desenvolvida a base de dados de boas práticas utilizando o *DSpace*, além de realizadas as primeiras gravações de narrativas com servidores em vias de aposentadoria - *storytelling*.

Para disseminar o conhecimento, criar novos, estimular e integrar magistrados e servidores a compartilharem e registrarem seus conhecimentos foi idealizado o evento denominado Café com conhecimento, com 4 quatro edições. Foram estabelecidos assuntos de interesse para a gestão do conhecimento organizacional e, em cada encontro, houve a participação de um *expert* no assunto, um moderador e um case de sucesso, com ênfase em casos corporativos, a saber:

- Painel 1 - Gestão do conhecimento na administração pública;
- Painel 2 - A inovação no Poder Judiciário como diferencial em tempos de crise;
- Painel 3 - Gestão do conhecimento na área judiciária: desafios e perspectivas;
- Painel 4 - Como os líderes podem abrir as portas do conhecimento para as suas equipes?

Tanto no Tribunal, quanto nas seccionais a ele vinculadas, o resultado da disseminação da importância da implantação da gestão do conhecimento como forma de potencializar o alcance dos objetivos estratégicos e preservar o conhecimento gerado já pode ser percebido por meio do engajamento dos servidores participando efetivamente na coleta das boas práticas em suas unidades e promovendo o Café com conhecimento.

A gestão do conhecimento contribuiu para a organização conhecer, e repensar procedimentos e ações, e se relaciona diretamente segundo Dalkir (2005):

[...] “coordenação sistêmica e voluntária das pessoas, tecnologia, processos e estrutura organizacional com o objetivo de agregar valor por meio do aproveitamento e inovação. Essa gestão é realizada por meio da criação, compartilhamento e aplicação de conhecimento, além de suprir a organização com lições valiosas e melhores práticas para a memória institucional com vistas à melhoria e aprendizagem contínuas” (p. 3, tradução nossa).

Assim, o TRF1, diante do cenário de perda de seu potencial humano e do aprendizado gerado, não poderia escapar da busca de uma estratégia que possibilitasse a permanência, do aprendizado, sob pena de baixar o seu desempenho e empobrecer a qualidade da prestação dos serviços à sociedade.

REFERÊNCIAS BIBLIOGRÁFICAS

- Allee, Verna. (c2003). *The future of knowledge: increasing prosperity through value networks*. Amsterdam, Nederland: Butterworth-Heinemann.
- Dalkir, K. (2013). Knowledge management in theory and practice. Recuperado de <https://dianabarbosa.files.wordpress.com/2009/03/knowledge-management-kimiz-dalkir.pdf>.
- Fukunaga, F. (2017). Vamos aprender mais sobre práticas de GC? Parte 1 - Mapeamento de Conhecimento Crítico. 2015. Recuperado de <http://www.sbgc.org.br/blog/category/mapeamento-de-conhecimento-critico>.
- Huysman, M., & Wit, D. (2004). Practices of managing knowledge sharing: towards a second wave of knowledge management. *Knowledge and Process Management*, 11(2), 81-92.
- Jen-te Yang. (2007). The impact of knowledge sharing in organizational learning and effectiveness. *Journal of Knowledge Management*, 11(2), 83-90.
- Jenkins, H. (2006). *Converge culture: where old and new media collide*. New York, United States: New York University Press.
- Menezes, K. C. de, Johann, J., Valentim, P. P. y Scott, P. (2017) . Gestão do conhecimento nas organizações: uma aprendizagem em rede colaborativa. *Perspectivas em Gestão & Conhecimento*, 7, 145-159. Recuperado de <http://www.biblionline.ufpb.br/ojs2/index.php/pgc/article/view/33294>.
- Moresi, E. A. D. (2001). Inteligência organizacional: um referencial integrado. *Ciência da Informação*, 30 (2), 35-46.
- Nassar, Paulo, y Cogo, R. S. (2011). Comunicação e memória organizacional: o poder da contação de histórias. *Redes Sociais, comunicação, organizações*, 5, 1-17. Recuperado de http://www.abrapcorp.org.br/anais2011/trabalhos/trabalho_rodrigo_paulo.pdf.
- Nonaka, I., y Takeuchi, H. (1997). *Criação do conhecimento na empresa*. 9.ed. Rio de Janeiro, Brasil : Campus.
- Pereira, M. O. F., Silva, H. de F. N., Freitas, M. do C. D., y Oliveira, A. A. (2016). Memória organizacional e as suas contribuições para o fundo setorial–CT–INFRA–UFPR. *Perspectivas em Gestão & Conhecimento*, 6 (1), 128-140. Recuperado de <http://www.okara.ufpb.br/ojs2/index.php/pgc/article/view/23587> .
- Prochno, P. (2004). Transferindo práticas: construindo conhecimento arquitetural localmente. *Revista de Administração de Empresas*, 44 (1), 70-81.
- Ramos, D. N., y Teixeira, C. S. (2016). Redes de inovação: alinhamento conceitual e o panorama atual da temática no território brasileiro. *REAVI: Revista Eletrônica do Alto Vale do Itajaí*, 5 (8). Recuperado de <http://revistas.udesc.br/index.php/reavi/article/view/8942>.
- Ribeiro, J. S. de A. N., Soares, M. A. C., Jurza, P. H., Ziviani, F., y Neves, J. T. R. (2017). Gestão do conhecimento e desempenho organizacional: integração dinâmica entre competências e recursos.

Perspectivas em Gestão & Conhecimento, 7, 4-17. Recuperado de <http://www.okara.ufpb.br/ojs/index.php/pgc/article/view/32936/17289>.

Sabbadini, F. S., Chamovitz, I., Ferreira Filho, Silva, C. F. da, y Desidério, M. (2016). Teoria da criação do conhecimento organizacional: a consolidação de dois estudos de casos. *Revista de Administração da Fatea*, 12 (12), 7-153. Recuperado de <http://publicacoes.fatea.br/index.php/raf/article/viewArticle/1636>.

Saliba, R. (2016). Retenha conhecimento com a espiral do conhecimento. Recuperado de <http://ucj.com.br/retenha-conhecimento-usando-espiral-do-conhecimento/>.

Santos, A. R. dos, Pacheco, F. F., Pereira, H. J., y Bastos Jr. P. A. (2001). Gestão do conhecimento como modelo empresarial. In: *Gestão do conhecimento: uma experiência para o sucesso empresarial* (pp. 11-48). Curitiba, Brasil: Champagnat.



CONCEITOS EM ORGANIZAÇÃO DO CONHECIMENTO A PARTIR DO TEMA ÉTICA: DISCIPLINA DE USO DO SISTEMA CDU

Benildes Coura Moreira dos Santos Maculan

Universidade Federal de Minas Gerais (UFMG), Brasil, 0000-0003-4303-9071, benildes@gmail.com

RESUMO Apresenta um experimento para a apreensão de conceitos em organização do conhecimento a partir da exploração e análise do tema “ética”, em suas diversas dimensões. Tem como ambiente de aplicação o ensino do sistema de classificação bibliográfico Classificação Decimal Universal (CDU) para alunos do curso de Biblioteconomia da Escola de Ciência da Informação da Universidade Federal de Minas Gerais (ECI/UFMG). Para alcançar o objetivo, empregou os procedimentos da estratégia da Pedagogia de Projetos para a classificação de recurso documental de um acervo sobre ética. Considera-se que o sistema CDU é um tipo de sistema de organização de conhecimento (SOC) que foi desenvolvido como instrumento de codificação para a organização de assuntos de uma área do conhecimento em sistemas de recuperação de informação. Ao utilizar esse instrumento, o aluno reverte o processo ao fazer uma análise no documento para determinar os assuntos que são tratados, que é o conhecimento que está registrado nele, e determinar um lugar para ele ocupar no acervo, levando em consideração o conhecimento representado pela coleção como um todo. Os resultados indicam que os alunos são capazes de perceber os conceitos fundamentais da organização do conhecimento e que os elementos do instrumento são mecanismos importantes para a organização do conhecimento em diferentes contextos, sobretudo para propósitos específicos.

PALAVRAS-CHAVE *Organização do conhecimento, Classificação Decimal Universal, Representação do conhecimento, Ensino em organização do conhecimento.*

ABSTRACT This paper presents an experiment for the seizure of concepts in organization of the knowledge from the exploitation and analysis of the theme "ethics" in its various dimensions. It was applied in teaching the bibliographic classification system, Universal Decimal Classification (UDC), for students of the course of library science, school of information science at the Federal University of Minas Gerais (ECI/UFMG). It used the procedures of the Project Pedagogy strategy to classify documentary resource of a collection on ethics. The UCD is considered as a type of knowledge organization system (KOS) that was developed as an instrument of codification for the organization of subjects of a knowledge area in information retrieval systems. When using this tool, the student reverses the process to make an analysis in the document to determine the matters that are dealt with, which is knowledge that is registered. The student determine a place for the document in the collection, taking into account the knowledge represented by the collection as a whole. The results indicate that students are able to perceive the fundamental concepts of knowledge organization and that the elements of the instrument are important mechanisms for the organization of knowledge in different contexts, especially for specific purposes.

KEYWORDS *Knowledge organization, Universal Decimal Classification, Knowledge representation, Teaching in knowledge organization.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

O ensino-aprendizagem exige a superação do atual cenário de dinamicidade do conhecimento, inovações e de multifacetadas relações e transformações, impondo o desafio de não oferecer apenas concepções lineares e fragmentadas nesse processo. Morin (2004, p. 105) afirma que “a disciplina é uma categoria organizadora do conhecimento científico [...] [que tende] pela delimitação de suas fronteiras, pela linguagem que ela constitui, [...] pelas teorias que lhe são próprias”. Entretanto, é preciso avançar e não se isolar na disciplina e fomentar a cooperação entre as disciplinas e o trabalho coletivo e compartilhado. Assim, é necessário significar e ressignificar o conhecimento, buscando conexões que coadunam em novas formas de pensar e fazer a educação.

O campo dos estudos da Organização do Conhecimento engloba a ordenação, estruturação e sistematização dos conceitos, de acordo com suas características (elementos de herança do objeto), e a aplicação desses conceitos em classes ordenadas por seus valores (conteúdos dos objetos ou assuntos) (Dahlberg, 2006). Nesse sentido, a formação no campo da Organização do Conhecimento compreende entender o tratamento e a recuperação da informação, usando teorias para a construção, manutenção, uso e avaliação de instrumentos lógico-linguísticos (linguagens documentárias). Essa afirmação é corroborada por Smiraglia (2002, p. 331) quando declara que “a Organização do Conhecimento, no âmbito da Biblioteconomia e Ciência da Informação, é o ramo do conhecimento da construção de ferramentas para o armazenamento e recuperação de entidades documentárias”. Para Hjørland (2003), a Organização do Conhecimento possui conceito amplo, envolvendo a organização cognitiva do conhecimento (sistemas conceituais e teorias) e a organização social do conhecimento, que se refere à representação do conhecimento disciplinar, tal como ocorre no sistema CDU.

Os instrumentos lógico-linguísticos visam orientar os processos de representação, classificação, ordenação e armazenamento dos recursos informacionais (documentos), com o objetivo de sua recuperação e comunicação (transferência) (Esteban Navarro, 1996). Nesse sentido, a organização do conhecimento tem como finalidade permitir a recuperação da informação nos mais diversificados contextos de produção e uso, suscitando esquemas para a construção de linguagens documentárias que acolham os diferentes propósitos para contextos específicos.

O ensino dos sistemas de classificação bibliográficos, e, no caso específico do ensino da Classificação Decimal Universal (CDU), exige, além de aprender a usar os mecanismos disponíveis para a formação de uma notação, empregando a sintaxe do esquema (tabelas) de forma adequada, permitir que o aluno apreenda a teoria por trás do sistema de classificação, cuja filosofia permite organizar o conhecimento representado pelo conteúdo dos documentos.

Em um sentido pragmático, Vickery (2008, s.p.) afirma que entre as tarefas dos profissionais da informação inclui “tornar o conhecimento [...] [disponível] para aqueles que a procuram, para armazená-la de forma acessível, e fornecer ferramentas e procedimentos que tornam mais fácil para as pessoas encontrarem o que procuram nesses armazenamentos”. Nessa afirmativa pode-se considerar três processos de Organização do Conhecimento: “o processo de produção do conhecimento (registrado); a sua organização (mediante sistemas de classificação, tesouros etc.) e, por conseguinte, a possibilidade de seu acesso, que reverte no uso” (Rabello; Guimarães, 2006, p. 20). Assim, ao classificar um documento (conhecimento registrado) utilizando o sistema CDU, o aluno está realizando um diagnóstico sobre o assunto que é tratado nele e determinando um lugar funcional para ele ocupar no acervo (Thompson e Carr, 1990), levando em consideração o conhecimento representado pela coleção

como um todo. Para Mann (1962), o objetivo primordial da atividade da classificação é a organização do conhecimento de forma sistemática. Dessa maneira, o usuário poderá, facilmente, acessar ao documento ou grupo de documentos que tratam sobre um mesmo assunto, pois, na organização da coleção os documentos são arranjados por seus assuntos, um em relação ao outro, logicamente dispostos. Assim, a Organização do Conhecimento

[...] é a expressão mais abrangente para designar a função da biblioteca desempenhada pela classificação. Indica a habilidade não apenas para identificar itens de informação específicos e definidos de forma precisa, mas também para demonstrar a completa gama de assuntos disponíveis na biblioteca e suas relações entre si. Essa é uma função educacional muito importante das bibliotecas e tem sido um pouco negligenciada nos modernos estudos de classificação (Langridge, 2006, p. 19)

Tendo isso em vista, como estratégia para o ensino-aprendizagem de conceitos em Organização do Conhecimento foram realizadas atividades de classificação de um acervo contendo documentos que versam sobre o tema “ética”, durante a disciplina “Sistemas de Classificação: CDU”, aplicadas a alunos do curso de Biblioteconomia da Escola de Ciência da Informação da Universidade Federal de Minas Gerais (ECI/UFMG), cujos procedimentos e resultados são apresentados neste artigo. Destaca-se que a escolha em trabalhar com o tema “ética” com os alunos do curso se justifica pela inexistência desse tópico como disciplina, o que é considerado uma deficiência da grade curricular, estando já em estado avançado de estudo a reformulação e inclusão da disciplina na nova grade.

Para cumprir o objetivo de alcançar o ensino-aprendizagem de conceitos em Organização do Conhecimento, foi utilizada a Pedagogia de Projetos, e, nessa perspectiva, o projeto envolveu trabalhar a classificação de um acervo digital sobre o tema “ética” em suas diversas dimensões, o que permitiu aos alunos perceber os conceitos que envolvem a organização do conhecimento do tema em questão a partir de diferentes dimensões, demandas de necessidades e contextos diversos.

PEDAGOGIA DE PROJETOS

Fernando Hernández (1998) esclarece que a Pedagogia de Projetos é uma concepção de ensino e não um método, uma vez que a ideia não é apresentar uma sequência regular e pré-determinados de ações. O objetivo da Pedagogia de Projetos é possibilitar a construção de conteúdos orientada pela determinação prévia de metas (intenção do projeto), em acordo celebrado entre alunos e professor, sendo o percurso flexível e reformulado em razão dos problemas que vão surgindo no caminho. O autor afirma que

A função do projeto é favorecer a criação de estratégias de organização dos conhecimentos escolares em relação a: 1) o tratamento da informação e 2) a relação entre os diferentes conteúdos em torno de problemas ou hipóteses que facilitem aos alunos a construção de conhecimentos, a transformação da informação procedente dos diferentes saberes disciplinares em conhecimento próprio (Hernández, 1998, p. 61).

Segundo Hernández (1998), essa visão não é nova, pois teve origem com J. Dewey (1859-1952) e Willian H. Kilpatrick (1871-1965), no século XX, durante o movimento educacional progressista nos Estados Unidos, amparada em uma educação aberta, na qual o aluno é agente na própria aprendizagem. Assim, o aluno aprende pelas próprias experiências, tomando atitudes diante dos problemas, ao visar alcançar as metas propostas (Leite, 2000). Segundo Schön (2000, p. 79), o “paradoxo de aprender uma

competência realmente nova é este: um estudante não pode inicialmente entender o que precisa aprender; ele pode aprendê-lo somente educando a si mesmo e só pode educar-se começando a fazer o que ainda não entende”. Dessa forma, o aluno se torna atuante no processo de ensino-aprendizagem, pois tem que buscar soluções ao problema apresentado, articulando a nova questão com os seus conhecimentos prévios. A Pedagogia de Projetos propicia, assim, diferentes formas de aprender.

Junto à estratégia da Pedagogia de Projetos é interessante utilizar a técnica dos Mapas Conceituais, criada em meados da década de 1970, por Joseph Novak e alguns de seus colaboradores (Novak e Gowin, 1996), nos Estados Unidos, e que pode ser compreendida a partir da teoria cognitiva de aprendizagem de David Ausubel (Ausubel, 2003), que propõe uma aprendizagem com base na solução de problemas. Os Mapas Conceituais são diagramas compostos por rótulos hierarquizados e interligados por relações significativas (expressões verbais), conforme mostra a Figura 1.

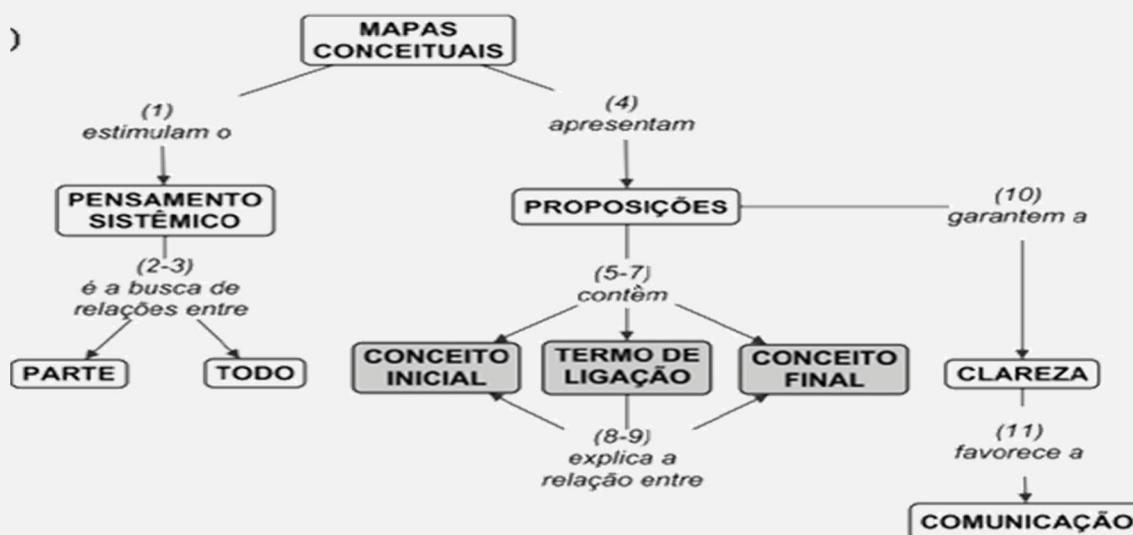


Figura 3 – Construção de Mapas Conceituais

Fonte: Correia, 2016, p. 43.

Pela análise da Figura 1 é possível perceber que os mapas conceituais estimulam o (1) pensamento sistêmico, que é a busca de relações entre o todo e suas partes (2-3), e apresentam proposições (4) compostas por um conceito inicial e um conceito final (5-7) conectados por um termo de ligação que explica a relação entre esses conceitos (8-9), garantindo a clareza (10) e favorecendo a comunicação (11) do conhecimento mapeado. Dessa forma, ao elaborar um mapa conceitual o aluno concretiza a sua aprendizagem na medida em que uma informação nova (conceito, proposição) adquire significado para ele ao se juntar à estrutura cognitiva preexistente evidenciando a sua aprendizagem (Ausubel, 2003).

Sobre os procedimentos aplicáveis, Hernández (1998) apresenta a Pedagogia de Projetos em quatro etapas: (1) problematização: apresenta-se a intenção do projeto, quando os alunos podem apresentar suas percepções sobre o problema em questão, expressando suas dúvidas e conhecimentos sobre ela, sendo incentivados pelo professor; (2) desenvolvimento: são estabelecidos grupos de trabalho e estratégias para a solução do problema, a partir de propostas e de socialização do conhecimento (compartilhamento de informações), criando um ambiente propício para contrapor pontos de vista e para refletir os conteúdos a serem apreendidos; (3) aplicação: promoção de novos conhecimentos a partir da execução do projeto; (4) avaliação: estima-se o grau de desenvolvimento dos alunos na

construção do próprio conhecimento, atribuindo valor na medida em que se considera que o processo está alinhado aos objetivos propostos e aos resultados obtidos.

Com esses procedimentos, o aluno aprende a partir do levantamento de dúvidas, pesquisas e na criação de relações com os prévios e os novos conhecimentos. Nesse processo, o professor é um mediador, pois “... no desenvolvimento do projeto o professor pode trabalhar com diferentes tipos de conhecimentos que estão imbricados e representados em termos de três construções: procedimentos e estratégias de resolução de problemas, conceitos disciplinares e estratégias e conceitos sobre aprender (Valente, 2000, p. 4).

Assim, do ponto de vista de aprendizagem, o aluno tem a oportunidade de ressignificar os conceitos prévios, recontextualizando conhecimentos já apreendidos e, a partir das estratégias utilizadas na solução do problema proposto pelo Projeto de Ensino, ampliar o seu universo de aprendizagem (Prado, 2001). Com isso, é possível apreender os fundamentos da teoria do sistema CDU, juntamente com a prática do uso do instrumento, conforme metodologia descrita a seguir.

METODOLOGIA E PROCEDIMENTOS APLICADOS

A estratégia da Pedagogia de Projetos vem sendo aplicada na disciplina “Sistemas de Classificação: CDU” desde o segundo semestre de 2015, ou seja, já foi aplicado em seis diferentes turmas, sendo duas turmas no segundo semestre de 2015, três turmas nos dois semestres de 2016 e uma turma no primeiro semestre de 2017.

Para cumprir a proposta do projeto, primeiramente, a turma de alunos é dividida em pequenos grupos de três a cinco integrantes cada, em geral, totalizando cinco grupos. A estratégia da Pedagogia de Projetos é, então, aplicada nas quatro etapas descritas na seção anterior, a saber: problematização, desenvolvimento, aplicação e avaliação.

Na etapa da problematização, é apresentada a intenção do projeto, com a proposta de objetivos e metas. Neste momento, é solicitado aos grupos que façam uma busca na literatura sobre o tema “ética” e que tragam os conteúdos recuperados para discussão em sala de aula. Após as discussões, em geral, com raras adaptações, são selecionadas cinco dimensões da ética a serem trabalhadas pelos grupos (uma dimensão para cada grupo): Ética Moral, Ética Profissional (com foco no bibliotecário), Ética na Política, Ética na Saúde e Ética Ambiental. Ao final desta etapa, cada grupo elabora uma definição para a dimensão da ética trabalhada que deve nortear a classificação que será realizada.

Também a partir do levantamento na literatura, é compilado um acervo, com cerca de trinta e cinco documentos sobre o tema “ética”, composto por artigos, relatórios, resenhas de livros e teses/dissertações, nacionais e internacionais, selecionados de forma que sempre há pelo menos dez documentos com conteúdo específico para cada uma das cinco dimensões trabalhadas sobre ética.

Na etapa do desenvolvimento, com os grupos já estabelecidos, inicia-se a preparação das propostas de uma situação-problema e de solução do problema, sendo incentivada a troca e compartilhamento de informações entre os grupos, uma vez que o acervo é o mesmo, mas cada grupo irá focar a classificação dentro da dimensão que lhe foi atribuída. Neste momento, em conjunto, a turma elabora as referências bibliográficas dos documentos do acervo, utilizando as normas da Associação Brasileira de Normas Técnicas (ABNT).

Na etapa de aplicação, os alunos trabalham na construção de uma situação-problema para justificar a classificação do acervo, com uma representação bastante específica do assunto e utilizando o sistema CDU, para uma biblioteca especializada, criando uma demanda para usuários específicos, que são o público-alvo. A situação-problema, com a organização das ideias e do conhecimento, é montada a partir das seguintes perguntas: O que? Como? Por quê? Para que? Onde? Quando? Quem?. Nesta etapa é ainda necessário criar um mapa conceitual (ver Figura 1) para representar a situação-problema estabelecida, de forma que seja um modelo da realidade, onde o cenário do problema é representado em todos os seus elementos, com base no estabelecimento de relações de causa e de efeito entre as diversas variáveis do problema. Por fim, nesta etapa, os grupos realizam a classificação do acervo seguindo os critérios: (1) definição de uma ordem de citação: estabelecimento de uma ordem de citação adequada ao público-alvo; (2) classificação mais específica: seleção de quinze documentos, sendo pelo menos três itens estrangeiros; (3) classificação mais geral: os documentos restantes ou seleção de vinte documentos.

Na etapa de avaliação, o Projeto é apresentado na seguinte ordem: (1) apresentação, pelo grupo, da situação-problema e da solução encontrada, determinando a biblioteca, o público e a necessidade de informação; (2) apresentação individual: cada aluno faz a apresentação da classificação específica realizada para um documento diferente, com explicações detalhadas dos elementos utilizados, não sendo aceito a classificação de documentos iguais.

RESULTADOS

Na etapa da problematização, a partir do material recuperado na literatura, os alunos são incentivados a declarar as suas percepções sobre o tema, que, ao final, sempre giram em torno das virtudes éticas que advêm da instrução, segundo afirma Aristóteles, tais como sabedoria, inteligência e discernimento, e, outras, advindas da força do hábito, tais como coragem, moderação, liberalidade, amabilidade, imparcialidade, honestidade, veracidade, laboriosidade, recato e indignação justa (Pereira, 2010). Os alunos são levados a uma reflexão crítica sobre as diferenças entre os relacionamentos profissionais e pessoais e sobre a importância e influência desses relacionamentos nas próprias decisões éticas e como outros indivíduos podem ser afetados por elas. Ao final do Projeto, a partir do *feedback* dos alunos, evidencia-se que eles percebem que a ética envolve todo o comportamento humano (princípios, valores, moral) e visa o bem-estar da vida em sociedade.

Ainda na etapa da problematização, ao estabelecer a tarefa da busca bibliográfica sobre o tema, os alunos já precisam praticar a organização do conhecimento sobre ética, pois montar a estratégia de busca exige compor um conjunto de regras que possam responder à questão formulada, de modo a estabelecer as suas diversas dimensões, pois isso ocorre no processo de “OC [que] produz representações de conhecimento utilizadas na OI para padronizar as representações de informação (Bräscher & Café, 2008, p. 12).

Dessa forma, no planejamento da estratégia de busca de informações sobre o tema “ética”, os grupos têm que seguir algumas ações de organização, tais como: compreensão da questão proposta; escolhas das bases para a busca do tema; definição das palavras-chave para a busca; uso de operadores booleanos; verificação da consistência e relevância dos documentos recuperados. Com os resultados das discussões realizadas nesta etapa, definem-se as cinco dimensões que serão trabalhadas (citadas na seção de metodologia). Ao final desta etapa, os grupos elaboram uma definição para cada uma das cinco

dimensões sobre ética, o que exige estabelecer semelhanças e diferenças, agrupando os conceitos, assim como as relações entre esses conceitos.

Na etapa do desenvolvimento, os grupos verticalizam o conhecimento acerca do assunto sobre a dimensão trabalhada, construindo a proposta de uma situação-problema e suas possíveis soluções, os grupos discutem entre os integrantes e listam elementos tais como: (a) de ideias e associações iniciais, buscando informações adicionais sobre a dimensão da ética trabalhada, de forma a permitir a teorização e a construção de novos conhecimentos e significados para melhor enfrentar o problema observado; (b) de percepções, sentimentos e valores acerca do problema verificado; (c) das possíveis explicações para o problema observado, a partir de fundamentos estudados durante a formação no curso de Biblioteconomia; (d) levantamento de hipóteses ou pressupostos, formulando explicações que serão a base para a elaboração da proposta de classificação.

Na etapa de aplicação, os alunos trabalham na definição da situação-problema e são incentivados e desafiados a encontrar uma solução, o que implica em tomar decisões e fazer recortes no domínio trabalhado. Neste momento, os alunos são orientados no sentido de aplicar os conceitos envolvidos na Organização do Conhecimento, uma vez que têm que realizar diferentes operações para permitir a apropriação do conhecimento sobre o tema, tais como trabalhar com diferentes hipóteses ou pressupostos para definir a unidade de informação específica e quais seriam as necessidades de informação da comunidade de usuários. Nessa atividade, os grupos levantam diversas questões e dúvidas que vão sendo sanadas durante as orientações com o professor, enriquecendo o aprendizado.

Em seguida, os grupos elaboram o Mapa Conceitual visando representar, graficamente, a solução para a situação-problema (tema focal). A elaboração dos mapas tem relação estreita com a Organização do Conhecimento, uma vez que aos alunos é exigido: listar conceitos-chave do conteúdo (conceitos mais gerais e mais específicos), organizar as informações relevantes (sintetizar e hierarquizar um todo coerente), estabelecer relações conceituais (causa e efeito, etc.), descrever os eventos que envolvem o conteúdo, responder a uma questão articulando proposições, sistematizar um projeto. Dessa forma, o mapa permitiu a organização e a visualização do conhecimento sobre a questão, de maneira sistemática e dinâmica, uma vez que o mapa é criado pelo grupo e não individualmente, o que permitiu o compartilhamento de impressões e uma interação cognitiva entre os conhecimentos prévios e os novos conhecimentos de todos os envolvidos na tarefa.

Na etapa de avaliação, quando o Projeto é apresentado como um todo pelo grupo, os integrantes são incentivados a fornecer *feedback* sobre os procedimentos realizados. Em geral, todos sempre declaram que a elaboração do Mapa Conceitual auxiliou tanto na construção da situação-problema quanto na definição da solução encontrada, pois permitiu enxergar falhas na representação e *insights* importantes para compreender melhor o problema e mais claramente perceber elementos que os levaram a alcançar uma resolução adequada. Ademais, os alunos revelam que o fato de terem que classificar os itens do acervo de forma bem detalhada, permitiu esclarecer algumas dúvidas acerca dos elementos do sistema CDU, solidificando o aprendizado adquirido e evidenciando novas questões que foram se agregando ao conhecimento acumulado sobre o uso do instrumento, tanto para os alunos quanto para o professor. Ao trabalharem uma dimensão específica dentro do tema “ética”, possibilitou aos alunos verticalizar o conhecimento acerca do assunto, para que a notação definida para o item ficasse adequada ao contexto temático trabalhado.

Ao final, mais que a formação teórica, intelectual e procedimental, com a apreensão de conceitos em Organização do Conhecimento e do aprendizado sobre o uso dos elementos do sistema CDU, ficou

evidente que houve a assimilação do tema trabalhado, agregando valor aos conteúdos da disciplina, pois os alunos têm que sistematizar e organizar o conhecimento sobre o tema. Na atividade do profissional da informação está incluída uma função social, e, como agente informacional, o ato de classificar para representar o conteúdo de um documento pressupõe atitudes éticas no sentido de que ele deve se preocupar com a coletividade que é atendida pela unidade de informação, comportando-se de forma responsável no atendimento às necessidades informacionais de todos, sem distinção.

CONCLUSÕES

Este artigo apresentou uma forma de apreensão de conceitos em Organização do Conhecimento a partir da classificação de um acervo sobre o tema “ética”, em suas diversas dimensões, utilizando a disciplina “Sistemas de Classificação: CDU” e empregando a estratégia da Pedagogia de Projetos. Essa atividade, além de concretizar o aprendizado do uso do instrumento, permitiu ao aluno apreender importantes conceitos em organização do conhecimento, uma vez que exigiu pensar a classificação do acervo a partir de suas dimensões específicas, dentro de distintos contextos de uso e de necessidades informacionais particulares.

Partiu-se do pressuposto de que o ensino-aprendizagem exige a ação do aluno sobre o objeto a ser conhecido de forma significativa. Nessa perspectiva, é necessário desafiar o aluno a resolver problemas, através de interação entre colegas e professor, considerando as questões e reflexões levantadas. Isso foi alcançado com o emprego da estratégia da Pedagogia de Projetos, que possibilitou unir o aprendizado da teoria por trás do sistema de classificação, que envolve conceitos da Organização do Conhecimento, com a prática da atividade de classificação dos saberes registrados, com a formação de notações específicas dos assuntos dos documentos do acervo, que também envolve lidar com conceitos da Organização do Conhecimento. Nas atividades do Projeto, foi imperativo aplicar os diversos mecanismos do sistema CDU (tabelas gerais e auxiliares, assim como os sinais relacionais) para representar os assuntos de forma detalhada e relacionada.

O uso do mapa conceitual permitiu aos alunos investigar a natureza do conhecimento sobre a temática tratada e sobre a organização do acervo para contextos e públicos diversos, apreendidos durante a dinâmica da aplicação prática e em conjunto da representação da construção da situação-problema e da solução. Nessa atividade, os alunos trabalharam o processo cognitivo por meio da captação das percepções individuais dos integrantes do grupo, com o objetivo de construção do conhecimento e de assimilação do pensamento compartilhado pelo grupo. Dessa forma, em sala de aula, é preciso criar situações nas quais os conceitos possam ser apreendidos de forma mais profunda, adquirindo significados e gerando aprendizado, com os alunos assumindo papel ativo no ensino-aprendizagem. Ao longo do tempo que a elaboração dos mapas vem sendo empregada na disciplina, às turmas do curso de Biblioteconomia, percebe-se a autonomia do discente na construção do seu próprio conhecimento, característica que o aluno levará, no futuro próximo, para a prática profissional.

A construção da situação-problema desencadeou um processo de problematização, com a articulação entre os conhecimentos prévios e aqueles adquiridos na construção de uma solução viável para o problema. Considera-se que a problematização é o ponto central para o desenvolvimento da ciência e da tecnologia, que sempre buscam por soluções para problemas da sociedade, e, assim, promovendo o desenvolvimento humano.

AGRADECIMENTOS

Agradecimento às agências de fomento Fundação de Amparo à Pesquisa em Minas Gerais (FAPEMIG), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Pró-Reitoria de Pesquisa (PRPq) da Universidade Federal de Minas Gerais pelo apoio financeiro com bolsas de Iniciação Científica e Monitoria, que permitiu desenvolver atividades de inovação na disciplina em questão, e pela ajuda de custo para este evento.

REFERÊNCIAS BIBLIOGRÁFICAS

- Ausubel, D. P. (2003). *Aquisição e retenção de conhecimentos: uma perspectiva cognitiva*. Lisboa: Plátano Edições Técnicas.
- Bräscher, M., & Café, L. (2008). *Organização da Informação ou Organização do Conhecimento? Anais do XIX Encontro Nacional de Pesquisa Em Ciência da Informação*. São Paulo: ANCIB.
- Correia, P. M. (jul. de 2016). Por que vale a pena usar mapas conceituais no ensino superior? . *Rev. Grad. USP*, v. 1(n. 1), pp. p. 41-51. Acesso em 28 de maio de 2017, disponível em <http://www.revistas.usp.br/gradmais/article/view/117724/115384>
- Dahlberg, I. (2006). *Knowledge Organization*. Copenhagen, Denmark. Acesso em 2017, disponível em http://www.db.dk/bh/lifeboat_ko/CONCEPTS/knowledge_organization_Dahlberg.htm
- Esteban Navarro, M. A. (1996). El marco disciplinar de los lenguajes documentales: la Organización del Conocimiento y las ciencias sociales. *Scire*, 2(1), pp. 93-107.
- Fujita, M. S. (2001). *Organização do Conhecimento: algumas considerações para o tratamento temático da informação, Organização do Conhecimento*. Em K. (. Carrara, *Educação, universidade e pesquisa* (pp. 29-34). Marília: Unesp Marília Publicações.
- Hernández, F. (1998). *Transgressão e mudança na educação: Projetos de trabalho*. Porto Alegre: ArtMed.
- Hjorland, B. (2003). Fundaments of Knowledge Organization. *Know. Org.*, 30(2), pp. 87-111.
- Lamgridge, D. (2006). *Classificação: abordagem para estudantes de biblioteconomia*. Rio de Janeiro: Interciência.
- Leite, L. H. (jan./jun. de 2000). Os Projetos de trabalho: um espaço para viver a diversidade e a democracia na escola. *Revista de Educação*, 3(4), pp. 25-29.
- Mann, M. (1962). *Catálogo e classificação de livros*. Rio de Janeiro: Editora Fundo de Cultura.
- Novak, J., & Gowin, D.B. (1996). *Aprender a aprender*. Lisboa: Plátano Edições Técnicas.
- Pereira, E. (10 de out. de 2010). Ética na política: uma reflexão a respeito da imediata aplicabilidade da Lei da Ficha Limpa. *Revista Jus Navigandi* , 15(2676), pp. 1-2. Acesso em 29 de maio de 2017, disponível em <https://jus.com.br/artigos/17738>

Prado, M. (2001). *Articulando saberes e transformando a prática*. Brasília: Secretaria de Educação a Distância, SEED; Ministério da Educação.

Smiraglia, R. (2002, Winter). The progress of theory in knowledge organization. *Library Trends*, 50(3), pp. 330-349. Acesso em 28 maio de 2017, disponível em https://www.ideals.illinois.edu/bitstream/handle/2142/8414/librarytrendsv50i3d_opt.pdf?sequence=1

Thompson, J., & Carr, R. (1990). *La biblioteca universitária: introducción a su gestión*. Salamanca, Madrid: Fundación Germán Sánchez Ruipérez.

Valente, J. (2002). *Repensando as situações de aprendizagem: o fazer e o compreender*. Brasília, DF: Secretaria de Educação a Distância, SEED; Ministério da Educação.

Vickery, B. (2008). On 'knowledge organisation'. *NTERNET Archive, Wayback Machine; Brian Vickery at home: on information, knowledge, and ideas*. (I. Archive, Ed.) Oxford, England, UK. Acesso em 28 maio de 2017, disponível em <http://web.archive.org/web/20080505050546/http://www>



GESTÃO DA INFORMAÇÃO OU DO CONHECIMENTO? UM ESTUDO DE CASO NA DIREÇÃO-GERAL DOS ESTABELECIMENTOS ESCOLARES-DIREÇÃO DE SERVIÇOS DO CENTRO

Ana Isabel Coelho¹, Maria Cristina Vieira de Freitas²

¹Faculdade de Letras da Universidade de Coimbra, 1996021922@student.uc.pt

²Faculdade de Letras da Universidade de Coimbra 0000-0002-8849-8792, cristina.freitas@fl.uc.pt

RESUMO Este estudo desenvolve-se numa instituição pública portuguesa, regional e desconcentrada, detentora de um sistema de gestão documental que modela o seu sistema de arquivo. Examina-se se o mesmo contribui para a partilha colaborativa de conhecimento e para as tomadas de decisão ou se a informação, apesar de gerida, não se transforma em conhecimento. Desenvolve-se por meio de um estudo de caso indutivo e observacional, em duas vertentes: i) verificação do Sistema de Gestão Documental, observando se, existindo os instrumentos indispensáveis, estes estão a ser corretamente utilizados e postos ao serviço do fim para o qual foram concebidos; ii) verificação do Sistema de Gestão da Informação, observando se a informação está sendo corretamente gerida, consubstanciando-se em conhecimento convertido (Nonaka & Takeuchi, 1995) e partilhado em diferentes vias. Considera-se o período crítico que se inicia em 2007-2008, quando foi implementado o SGD/SGI, até à presente data. Os resultados indicam que, mesmo detendo a instituição um sistema de gestão documental e de informação concebido e a funcionar, a informação, em certos casos, permanece fechada em “nichos”, não havendo uma partilha formalmente instituída. Conclui-se que a existência de um sistema desta natureza poderá ser condição necessária, mas não suficiente, para que surja um Sistema de Gestão do Conhecimento (SGC). Apontam-se como fatores inibidores da geração de processos de criação e gestão do conhecimento, a continuidade da instabilidade estrutural e funcional da organização, bem como o modelo de gestão verticalizado tradicionalmente adotado na Administração Pública Portuguesa. Sugerem-se como medidas corretivas, a curto e médio prazo, a promoção de formação continuada e o fomento de comunicação baseada em plataformas horizontais do tipo Web 2.0.

PALAVRAS-CHAVE *Gestão da Informação, Gestão do Conhecimento, Sistemas de Gestão Documental, Administração Pública, Portugal.*

ABSTRACT This study is carried out in a regional and deconcentrated Institution that takes part of the Portuguese Public Administration, which owns a Record Management System (RMS) that models its recordkeeping processes. The main purpose is to analyze whether the RMS contributes, or not, to the collaborative process of sharing of knowledge and decision making. Also, whether the information managed by the System, contributes, or not, to the process of knowledge creation. The methods and techniques used to collect and analyze data were developed under an inductive and observational case study research, highlighting two specific aspects: verification of the RMS main processes and tasks, examining if the technology and tools used nowadays are correct for the service that they were designed; verification of the SGI main processes and tasks, examining if the information that is inserted in the System is well managed to contribute to the process of conversion of knowledge (Nonaka & Takeuchi, 1995) and to the process of sharing knowledge in a collaborative way, among the different hierarchical organizational levels. The timeline

considered was the critical period that starts with the implementation of the RMS/IMS (2007-2008) to nowadays. The main results suggest that, although the Institution has a RMS/IMS that is operational and working, the information seems to remain closed in "niches", and there are no formal and systematic collaborative strategies to share knowledge and enhance the process. The main conclusion, points out that the existence of the System may be a necessary condition to exist a Knowledge Management System (KMS), but it not sufficient. In a holist approach, people, and not only technologies and process, are the core part of the System. Also, it is pointed out that the continuity of the structural and functional instability of the Institution over time, as well as the vertical management model traditionally adopted in the Portuguese Public Administration, are factors that strongly inhibit the emergence of a spontaneous and/or formal process of knowledge creation and sharing in the analyzed Institution. In the short and medium term, we suggest corrective measures to promote continuing training to the staff, and to develop communication strategies based on horizontal platforms (e.g., Web 2.0).

KEYWORDS *Information Management, Information Knowledge, Record Management Systems, Public Administration, Portugal.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

O advento das Tecnologias de Informação e Comunicação (TIC) gerou alguma imprecisão entre o uso dos termos informação e conhecimento, que habitualmente tendem a se mesclar. Esta confusão tem sido abusivamente perpetuada, mormente, consequência da desmaterialização e da rapidez do acesso e uso da informação. Dado, informação e conhecimento formam um sistema hierárquico, numa cadeia de agregação de valor, sendo as fronteiras que delimitam estes conceitos esbatidas e variáveis, em conformidade com o interlocutor, importando a sua clarificação.

Nas últimas décadas, tem sido recorrente, na literatura científica, a ideia de que os dados são a matéria-prima informacional em bruto e, portanto, descontextualizada. E que, ao atribuir significado aos dados, obtém-se informação pronta a ser utilizada (Davenport, & Prusak, 1998; Cong, & Pandya, 2003). Segundo Angeloni (2003), o uso estratégico, tático, operacional e controlado dessa informação constitui objeto da Gestão da Informação (GI). Contudo, para que “surja” o conhecimento, é necessário que a informação seja processada pelas pessoas. O conhecimento é, portanto, algo que não pode dissociar-se dos indivíduos e dos seus esquemas mentais; por isso, é variável, mas também mais desejável e mais complexo e difícil de gerir e de transmitir. Num tal sentido, prossegue, a Gestão do Conhecimento (GC) será o culminar de todo este processo, em que a informação flui, é processada e partilhada, gerando múltiplas combinações, tornando-se numa vantagem competitiva.

Quanto ao processo de GC, segundo Nonaka e Takeuchi (1995) e Nonaka e Konno (1998), ele implica a existência de estratégias de conversão, num ciclo ininterrupto e em espiral contínua, potenciadoras de quatro tipos de conhecimento nos meios organizacionais: socialização, externalização, combinação e internalização. Esses tipos de conhecimento são desenvolvidos em ambientes ricos e propícios à sua criação, utilização transformação e reutilização, e que são chamados, pelos mesmos autores, *ba* organizacionais (Nonaka & Konno, 1998).

Partindo desses pressupostos, depreende-se que: i) a existência de um Sistema de Gestão de Informação (SGI) poderá ser um começo, mas não bastará para alcançar padrões elevados de eficácia e de eficiência,

geradores de poupança de recursos e inibidores do desperdício, conceitos em voga nos dias que correm; e que ii) para que exista uma verdadeira Gestão do Conhecimento (GC), mesmo partindo da Gestão da Informação (GI), sendo esta fundada na boa estruturação e organização dos dados informacionais, é necessário algo mais.

Diversos estudos produzidos na senda das ideias reveladas, nos anos 90 e seguintes, por Nonaka e Takeuchi (1995), têm vindo a corroborar essas ideias, alguns incidindo sobre essas necessárias diferenciações entre os conceitos (*e.g.* Bustelo Ruesta & Amarilla Iglesias, 2001), outros, menos frequentes, focalizando na Administração Pública Portuguesa (APP) (*e.g.*, Vasques, Borges, & Freitas, 2013), demonstrando que, neste último caso, apesar de terem sido feitos largos investimentos em TIC, os retornos em termos de construção de um Sistema de Gestão do Conhecimento (SGC), que vise servir à missão para a qual os serviços foram criados, têm ficado aquém do potencial e dos resultados esperados.

Ainda, refira-se que, de acordo com Cong e Pandya (2003) cit. por Vasques, Borges e Freitas (2013), as componentes-chave para o sucesso dos processos de GC, nas organizações, residem numa tríplice aliança, capaz de proporcionar altos níveis de produtividade e, conseqüentemente, reduzir custos e promover a melhoria dos serviços (figura 1):

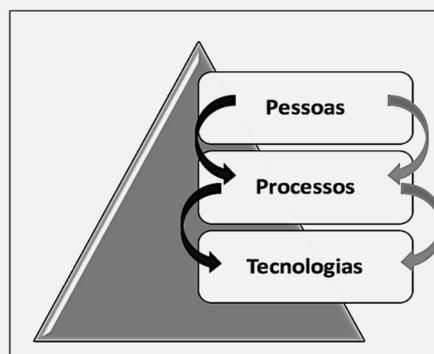


Figura 1. Componentes-chave do processo de conhecimento organizacional.

Fonte: Elaboração nossa, com base em Cong e Pandya (2003).

Para que essa tríplice aliança realmente funcione, as pessoas precisam ser encorajadas a criar, usar e partilhar conhecimento, devendo, para isso ser identificados, implementados e usados os processos e as tecnologias adequadas (figura 1).

Em Portugal, na Administração Pública, as chefias estão habituadas a uma gestão no curto prazo, coincidente com o ciclo político. As mudanças estruturais são aquelas que se prolongam para além desse limiar, não interessando no imediato aos gestores de topo, mais preocupados e por vezes assoberbados por aquilo que são os problemas urgentes. Não obstante, os arquivos organizacionais, enquanto instituições que salvaguardam e que disseminam a memória coletiva organizacional, nem sempre se enquadram nesse ciclo curto e nem naquilo que é visível numa vista superficial. Igualmente, em muitos casos, ainda subsiste a ideia do arquivo como uma entidade “morta”, com tudo de pó, de imobilidade e de isolamento que isso significa. Isso leva também as chefias, habitualmente com um conhecimento muito superficial dessas matérias, normalmente a encararem os seus arquivos com base no senso comum, levando-as a desvalorizar essa vertente e a empurrar os funcionários que consideram

mais inaptos para este setor, tornando-o ainda menos apetecível, num ciclo vicioso de perda e de desorganização.

No que toca à Educação, desde que, em 1881, o conceito de proximidade em relação à população escolar e, ao mesmo tempo, uniformidade de procedimentos, foi pela primeira vez implementado no terreno, com a criação das Circunscrições Escolares, se verificou que o conceito de conhecimento, apesar de perene na sua substância, tem sofrido numerosas alterações que se refletem na estrutura formal do(s) organismo(s), umas vezes, unindo-se num só serviço, outras vezes, pulverizando-se, depois voltando-se a unir. Muitas dessas alterações não se justificam em termos de mudanças de paradigmas sociais, mas sim de estratégias políticas (cf. Diário do Governo, 1881-1975 e Diário da República, 1976-2011).

Para toda a confusão reinante muito contribui a gestão de “não proximidade” e de “não partilha” e de “não uso”, em relação ao conhecimento que os agentes mais próximos dos problemas naturalmente dispõem, o que é o mesmo que dizer que a gestão de topo segue um ciclo político e tende a não se interessar por este saber adquirido, fruto da informação processada e que muito aproveitaria no processo de tomada de decisão.

O organismo da APP, desde 2012, designado Direção-Geral dos Estabelecimentos Escolares-Direção de Serviços do Centro (DGEstE-DSRC), é um corolário deste percurso de 136 anos de mudanças e é o objeto deste estudo. Para além de se contextualizar a sua evolução histórica, procura-se aferir se o funcionamento do atual Sistema de Gestão Documental e Sistema de Gestão da Informação (SGD/SGI), implementado em 2007-2008 e que modela todo o seu sistema de arquivo, potenciando a organização, a recuperação e a comunicação da informação produzida e acumulada, tem contribuído de forma ativa para a criação e a partilha do conhecimento, em diferentes vias, e também para as tomadas de decisão (Angeloni, 2003), ou se, pelo contrário, a informação gerida pelo organismo não se transforma em conhecimento, como seria desejável (Vasques, Borges, & Freitas, 2013), apresentando, nessas circunstâncias, soluções pragmáticas e exequíveis. Procurou-se, igualmente, aferir se o SGD contribui para a construção de uma Inteligência Coletiva, sendo esta definida, na aceção de Lévy (2003) citado por Bembem e Costa (2013, p. 141), como “[...] uma inteligência distribuída por toda parte, incessantemente valorizada, coordenada em tempo real, que resulta em uma mobilização efetiva das competências”. Como complemento, procurou-se refletir sobre a estabilidade do organismo, se indiferente ou não aos ciclos políticos, comparando, na mesma vertente, o período anterior à implementação do mencionado sistema.

METODOLOGIA

Os métodos usados para desenvolver o estudo de caso foram a observação, a análise documental e a análise bibliográfica (figura 2).

A observação realizou-se no SGD/SGI da DGEstE-DSRC. A escolha prendeu-se com o facto de uma das autoras do presente estudo exercer funções no referido organismo, detendo, por isso, condições privilegiadas para a sua realização (Stake, 2007; Yin, 2006). Durante a observação foram recolhidos dados vertidos em grelhas simples, elaboradas e estruturadas em duas vertentes: i) verificação das componentes do SGD e ii) verificação das componentes do SGI (figura 2). Para o efeito, considerou-se o período crítico de 2007-2008, que marcou a implementação do Sistema, até à presente data. Os

resultados dessa observação foram descritos e confrontados com a literatura científica recolhida sobre a matéria, gerando as principais discussões e conclusões do estudo.

A análise documental incidu em textos organizacionais (normativos e técnicos) e em documentação jurídica considerada relevante. Para tal, recorreu-se, como termo de contextualização, ao estudo da história dos 136 anos de existência do organismo, seccionando a linha do tempo no momento de implementação do atual SGD/SGI, obtendo-se, assim, um antes e um depois dessa decisão de implementação, essenciais como termo de comparação. Essa análise foi complementada por uma outra, de caráter bibliográfico, por meio da qual se obteve um conjunto consistente de textos, o que permitiu a elaboração de uma revisão do tema, focalizada na identificação de estudos prévios realizados, especialmente, mas não só, no contexto português (figura 2).



Figura 2. Matriz metodológica indutiva, concebida para o estudo de caso.

Fonte: elaboração nossa.

Ainda, a realçar que com base nessa forma indutiva, gradual e flexível de construir o estudo, resumida na figura 2, os dados obtidos puderam ser devidamente estruturados, para serem, então, analisados e contrastados, de forma a emitirem-se os principais resultados e as conclusões.

CONTEXTUALIZAÇÃO DO ESTUDO

Para se perceber esta problemática no contexto da DGEstE-DSRC, é necessário recuar até ao século XIX, ao âmago cronológico do seu surgimento, que coincidiu com uma série de reformas do ensino aí retomadas ou implementadas.

As Leis de 2/05/1878 e de 11/06/1880 foram disciplinadas pelo Decreto Regulamentar de 28/07/1881, com efeitos a 01/07/1881. Por estes diplomas, ficou o território nacional dividido, para fins de inspeção do ensino primário, em 12 circunscrições e, cada uma destas, subdividas em círculos escolares a abranger os concelhos aí incluídos.

Mas, para ser possível perceber da razão do surgimento dos serviços descentralizados da Educação, é necessário recuarmos ainda um pouco mais. De acordo com Justino e Almeida (2016), a partir de 1835, sucessivas reformas ao nível das instituições e das regulamentações foram sendo realizadas, culminando com a reforma do ensino de 1870, que criou o Ministério da Instrução Pública e, com ele, cerca de 16 circunscrições escolares, repartidas, por sua vez, em 50 circunscrições de concelhos.

Fica assim patente que os organismos desconcentrados e regionalizados da educação surgem pela primeira vez, em Portugal, no ano de 1881. Essa data é consistente com a data dos documentos mais antigos que existem no arquivo histórico da DSRC. Sucessivamente, os arquivos institucionais das instituições extintas foram sendo incorporados às instituições atuais. Fica também patente, desde 1878,

a necessidade de criação de um organismo desconcentrado, que servisse como “ponte” entre as escolas e a administração central, modelo esse que prevaleceu até a atualidade, tendo sido sucessivamente “testado” pelos diversos tumultos político-sociais que ocorreram ao longo destes mais de 100 anos, tais como, a implementação da República (1910), da Ditadura (1926) e da Revolução (1974), revelando-se funcionalmente o mais estável, dentro da instabilidade. Das sucessivas reformas estruturais, e de nome, que o organismo sofreu ao longo do tempo, saliente-se a criação das Direções dos Distritos Escolares e das Direções Regionais de Educação.

A organização da rede escolar sofreu uma mudança profunda entre 1933 e 1936, como consequência da ditadura Salazarista (Leite, 2015), impondo-se a reestruturação dos organismos desconcentrados da Educação. Em 1933, por Decreto n.º 22.369, de 30 de março, foram regulamentados o funcionamento das Direções dos Distritos escolares e também das respetivas Delegações de Zonas Escolares.

Adiante, em 1987, o Decreto-Lei n.º 3/87, de 03 de janeiro, vem criar as Direções Regionais de Educação (DRE). Estas são assumidas como órgãos desconcentrados de coordenação e de apoio aos estabelecimentos de ensino não superior e de gestão dos respetivos recursos humanos, financeiros e materiais, com exclusão das funções de controlo, a cargo da Inspeção-Geral de Ensino. Tratam-se de superestruturas regionais que tinham a cargo a gestão de diversos programas e medidas de apoio às escolas de ensino não superior, numa ótica de proximidade geográfica. Resumidamente, qualquer problema que surgisse numa escola (pública ou mesmo privada) era à DRE da sua área que esta se devia dirigir à procura de uma solução.

Finalmente, pelo Despacho n.º 1.3313/2003, datado de 08 de julho, são extintas as Delegações Escolares e os respetivos arquivos são integrados nas DRE. A partir dessa data, estas instituições passam a deter um expressivo conjunto documental que reflete mais de 130 anos de história da Educação em Portugal.

ESTUDO DE CASO: COMPONENTES DOS SGD/SGI NA DGESTE-DSRC

No que respeita às tecnologias e aos processos, os resultados revelam que, no âmbito da DGEstE-DSRC, é lícito afirmar que existe e encontra-se em funcionamento um SGD, mormente, porque a organização dispõe e usa as ferramentas necessárias à sua prossecução, nomeadamente, uma Portaria de Gestão Documental (PGD), aprovada e publicada em Diário da República (Portaria n.º 536/2009 de 18/05) e um Sistema Eletrónico de Gestão Documental (SEGD), que inclui o plano de classificação preconizado na referida Portaria. Ficou também claro que, de facto, existe um SGI, consubstanciado pelo planeamento, organização e controlo da informação que circula por canais pré-definidos, estando esta acessível em diferentes níveis de usos e de privilégios, atendendo ao grau de responsabilidade dos colaboradores. Saliente-se, ainda, a existência de um *workflow* bem definido e bem documentado (figura 3).

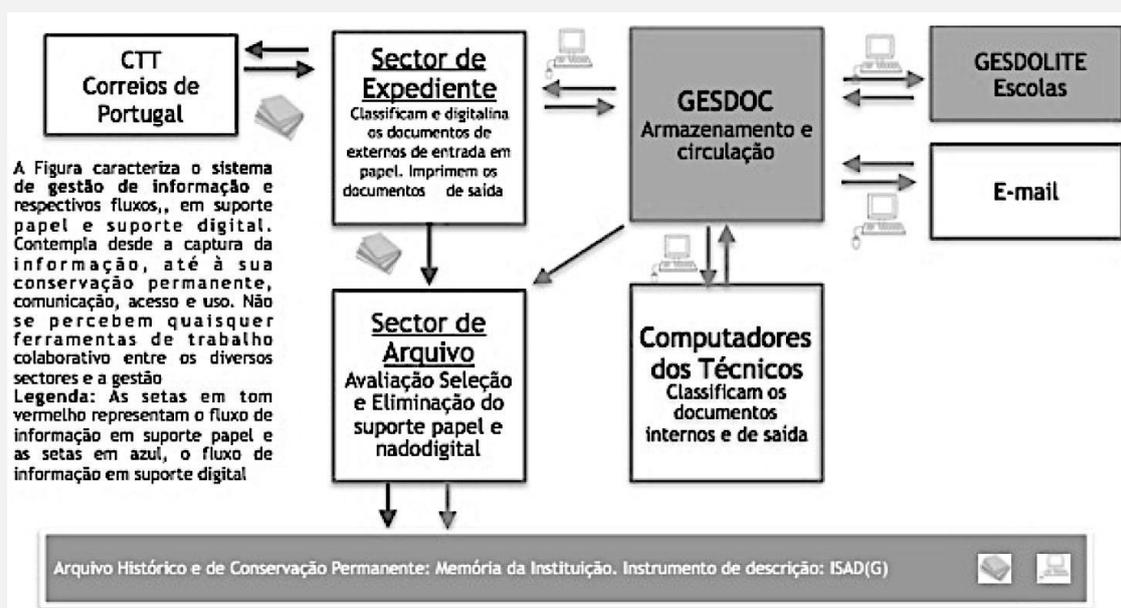


Figura 3. Representação do SGD/SGI da DGEstE-DSRC.

Fonte: Elaboração nossa.

Refira-se, também, o facto de a DGEstE-DSRC já ter realizado duas versões de um outro instrumento imprescindível à GD/GI do seu sistema, o designado Relatório de Avaliação de Documentação Acumulada (RADA), que submeteu, em conformidade com a legislação vigente, à apreciação da autoridade arquivística nacional, a Direção Geral do Livro dos Arquivos e das Bibliotecas (DGLAB), com o intuito de eliminar a documentação acumulada ao longo dos anos. Os RADA enviados à DGLAB foram aprovados e esse processo segue em moto contínuo. Realce-se, contudo, que ainda não foi avaliada nenhuma documentação em suporte eletrónico/digital.

O Arquivo Histórico institucional, por sua vez, existe, e a respetiva informação encontra-se descrita em conformidade com a normativa internacional *General International Standard for Archival Description* (ISAD G), nacionalmente harmonizada nas Orientações para a Descrição Arquivística (ODA) (cf. figura 3), estando esta acessível ao público e comunicável, obedecendo à legislação em vigor. Este é o arquivo com mais procura externa, também porque se encontra publicitado no sítio Web da DGEstE-DSRC¹ e por tal visível ao público.

Quanto ao perfil genérico dos utilizadores, registre-se que o público é frequente e usufrui dessa informação, quer para uso pessoal (história familiar), quer para a elaboração de monografias sobre história local. Entre os utilizadores destacam-se os estudantes de doutoramento e de mestrado, que procuram e aí encontram, fontes para o seu trabalho de investigação.

DISCUSSÃO E CONCLUSÃO

Os resultados coligidos e sumariados corroboram que a instituição pode ser considerada um caso exemplar, no âmbito da APP, na matéria de GD/GI, tendo envidado, nesse plano, e com sucesso,

¹ Disponível em: <http://www.dgeste.mec.pt/index.php/arquivo-historico-da-dgeste-ds-regiao-centro/>.

esforços consideráveis nos últimos 10 anos. No entanto, verifica-se que, apesar de o organismo ter concebido e implementado um SGD/SGI, consubstanciado pelo uso tático, operacional e controlado da sua informação, não detém um SGC, na aceção estratégica da palavra. Perante a descrição dos fluxos de informação no âmbito do seu SGI, infere-se que não existem, em uso, ferramentas destinadas ao incremento de trabalho colaborativo, entre os 100 trabalhadores e chefias que diariamente se encontram em funções na organização. Ou seja, não existe um sistema formal e perfeitamente internalizado de GC. Esse deveria ser o culminar de todo este processo, em que a informação flui, é processada pelos colaboradores, de acordo com as suas construções mentais (Angeloni, 2003), e amplamente partilhada e reutilizada, gerando múltiplas conversões (Nonaka & Takeuchi, 1995; Nonaka & Konno, 1998), obtendo em resposta, vantagens competitivas ou sociais, estas últimas mais de acordo com os objetivos da APP. Não obstante, observa-se que há uma partilha espontânea de informação, ocorrendo principalmente no seio dos gabinetes, entre colegas, o que é benéfico para a criação horizontal do conhecimento. No entanto, verifica-se que essa informação, de cariz eminentemente setorial, não influi sistematicamente no processo de tomada de decisão da gestão de topo ou não é usada de forma regular e consistente em processos formais de SGC, igualmente não detetados por esse estudo.

Por hipótese, este será um problema transversal à APP, em que a comunicação tradicionalmente tende a fazer-se univocamente do topo para a base. Esta situação é geradora de constrangimentos, que se podem refletir em reestruturações sucessivas de serviços, dependentes do ciclo político, que o tempo demonstra, depois, serem desfasadas das reais necessidades dos seus públicos.

Refira-se, igualmente, que o organismo em causa se debate com questões políticas que “emperram” o seu funcionamento, continuando a sua senda de instabilidade, potenciada, pelo facto de o seu âmbito funcional ser muito diferente daquilo que a última Lei Orgânica de 2012 (Decreto-Lei n.º 266-F de 31 de Dezembro de 2012) lhe ditou, ou seja, daquilo que o anterior ciclo político previu para seu destino.

Assim sendo, o estudo revela que na organização em questão, deter um SGD/SGI pode ser condição necessária, mas não suficiente, para alcançar um SGC, pois mesmo depois de ter sido instituído o atual sistema, em pleno funcionamento desde 2008, não nos parece possível, com base nos dados obtidos, considerá-lo como substrato para a tomada de decisão em várias vias (*botton-up* e *top-down*), condição *sine qua non* dos SGC. Tal evidência corrobora a literatura científica consultada. A criação de um SGC é algo complexo e multifatorial (Alavi & Leidner, 1999), não bastando investimentos em TIC, sem um efetivo envolvimento dos *stakeholders* e dos colaboradores, que configuram um dos pilares da tríplice aliança identificada por Cong e Pandya (2003), representado pelas pessoas, sendo de capital importância, numa articulação bem-sucedida com os outros dois: tecnologias e processos. Pelo que se percebe, a instituição ainda não conseguiu articular, e a contento, esses três pilares que fundamentam o exercício da GC. Também, segundo Angeloni (2003), o trabalho colaborativo, a partilha de informação e a tecnologia são aspetos cruciais e que devem ser usados em conjunto.

Atendendo a isso, e de forma a mudar a sua condição atual, recomenda-se à instituição a adoção de um modelo comunicativo, sem falhas na “engrenagem”, pronto a ser partilhado, por meio de ferramentas de trabalho colaborativo, que permitam a troca de informação entre trabalhadores e chefias, como um incremento da sua inteligência competitiva, nos moldes preconizados por Lévy e citados por Bembem e Costa (2013) e das suas competências informacionais individuais.

De modo a atingir esse objetivo, sugere-se, como medida a aplicar, ao nível interno, o incremento de uma política ativa e contínua de promoção de formação profissional dos trabalhadores, no local de trabalho, que funcionará como uma plataforma de partilha de informação geradora de mais valias e de

conhecimento, também pela maior flexibilidade das vias de comunicação. Outra ação impulsionadora, com impactos ao nível externo, seria deter uma página na Internet que incluísse, para além das informações habituais, ferramentas do tipo Web 2.0, de diálogo colaborativo e em tempo real, entre o organismo e os cidadãos, de que poderiam ser exemplo o desenvolvimento de páginas no Facebook® e no Twitter®, onde o público pudesse exprimir-se acerca dos serviços e interagir com o organismo, o que se por um lado acrescentaria responsabilidades pela visibilidade e exposição institucional, assumindo-se os riscos daí advindos, por outro, beneficiaria a interação pela abertura de um canal de comunicação direta entre o organismo e o seu público.

Pressupõe-se, com o culminar desse estudo, que se medidas sistemáticas não são adotadas, o conhecimento não surge espontaneamente a partir da implantação de um SGD/SGI, por mais afinado que ele seja, sendo necessárias intervenções concretas para a sua promoção. A própria rigidez dos fluxos de comunicação, no âmbito da Administração Pública, pode ser apontada como um dos fatores responsáveis por essa realidade. Tal já havia sido referido por Vasques, Borges e Freitas (2013), vindo o presente trabalho corroborar as suas conclusões. Também essas autoras referiram que os projetos de GC podem criar as condições ideais para a obtenção de eficiência na AP, como um equivalente à vantagem competitiva para as empresas privadas, sendo fundamentais para que essas organizações apresentem bons resultados e deem uma resposta positiva às necessidades dos cidadãos, com um mínimo de recursos possível. Esse não é, como diz Alavi e Leidner (1999), um caminho fácil, requerendo por vezes que ajustes e entendimentos, mas sendo necessário.

Refira-se que a história recente da DGEstE-DSRC é um bom exemplo, pois mesmo sofrendo diversas modificações, ao longo dos últimos 136 anos, consubstanciadas em reestruturações e em correções não isentas de custos e que são fruto, muitas vezes, não de necessidades reais, mas de estratégias políticas, o organismo tem sobrevivido, graças ao conteúdo funcional e à missão social que desempenha, sendo esses os valores que se sobrepõem aos aspetos pouco positivos da mudança.

Em jeito de conclusão, espera-se que sejam desenvolvidos mais estudos neste âmbito, que permitam recolher evidências e apresentar soluções que promovam o reposicionamento da APP, do paradigma do acesso ao documento/informação, para o paradigma da partilha colaborativa do conhecimento.

REFERÊNCIAS BIBLIOGRÁFICAS

Alavi, M., & Leidner, D. E. (1999). *Knowledge management and Knowledge Management Systems: conceptual foundations and research issues*. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.581.6535&rep=rep1&type=pdf>.

Angeloni, M. T. (2003). Elementos intervenientes na tomada de decisão. *Ciência da Informação*, 32 (1), 17-22. Disponível em: <http://www.scielo.br/pdf/ci/v32n1/15969.pdf>.

Bembem, A. H. C., & Costa, P. L. V. A. da (2013). Inteligência coletiva: um olhar sobre a produção de Pierre Lévy. *Perspectivas em Ciência da Informação*, 18(4), 139-151.

Bustelo Ruesta, C., & Amarilla Iglesias, R. (2001). Gestión del conocimiento y gestión de la información. *Boletín del Instituto Andaluz de Patrimonio Histórico*, 34, 226-230.

Carta de Lei de 01 de Junho de 1880 da Direção-Geral da Instrução Pública. Diário do Reino. Disponível em: <http://193.137.22.223/pt/patrimonio-educativo/museu-virtual/resenha-de-legislacao/>.

Carta de Lei de 02 de Maio de 1878 da Direção-Geral da Instrução Pública. Diário do Reino. Disponível em: <http://193.137.22.223/pt/patrimonio-educativo/museu-virtual/resenha-de-legislacao/>.

Cong, X. & Pandya K. (2003). Issues of Knowledge Management in the Public Sector. *Electronic Journal of Knowledge Management*, 1(2), 25-33.

Davenport, T. & Prusak, L. (1998). *Working knowledge: how organizations manage what they know*. Boston, MA: Harvard Business School Press.

Decreto n.º 22369, de 30 de Março de 1933. Ministério da Instrução Pública. Disponível em [https://dre.tretas.org/dre/223688/Despacho 13313/2003 de 08 de Julho de 2003](https://dre.tretas.org/dre/223688/Despacho%2013313/2003%20de%2008%20de%20Julho%20de%202003). Diário da República I série. Disponível em: <https://dre.tretas.org/dre/164662/>.

Decreto-Lei n.º 3/87, de 03 de Janeiro. Diário da República I série. Lisboa: Ministério da Educação. Disponível em: <https://dre.tretas.org/dre/36580/decreto-lei-3-87-de-3-de-janeiro#text>

Decreto-Lei n.º 266-F/2012 de 31 de Dezembro. Diário da República I Série n.º 252. Lisboa: Ministério da Educação e Ciência. Disponível em <https://dre.pt/application/dir/pdf1sdip/2012/12/25203/0029200295.pdf>.

Despacho n.º 13313/2003 de 08 de Julho. Diário da República n.º 155/2003-II Série. Lisboa: Ministério da Educação.

Justino, D., & Almeida, S. (2016). Inspeccionar e avaliar: origens da regulação coerciva da educação em Portugal. *Educação, Sociedades e Culturas*, 47, 31-51. Disponível em: http://www.fpce.up.pt/ciie/sites/default/files/ESC47David_0.pdf.

Leite, I. F. (2015). *O 25 de Abril e os elementos performativos*. Disponível em: <http://capazes.pt/cronicas/o-25-de-abril-o-estatuto-da-mulher/2/>.

Nonaka, I., & Takeuchi, H. (1995). *The Knowledge-Creating Company: how Japanese Companies create the dynamics of innovation*. New York: Oxford University Press.

Portaria n.º 536/2009 de 18 de Maio de 2009. Diário da República I Série. Disponível em: <https://dre.tretas.org/dre/252307/>.

Regulamento de 28 de Julho de 1881 da Direção-Geral da Instrução Pública. Diário do Governo. Lisboa: Ministério da Instrução Pública.

Stake, R. E. (2007). *Investigación con estudio de casos*. (4ª ed.). Madrid: Ediciones Morata.

Vasques, S., Borges, M. M., & Freitas, M. C. V. de. (2012). Informação, conhecimento e serviço público: um estudo de caso exploratório com contributos para a gestão da informação e do conhecimento na Câmara Municipal de Viseu. In Francisco, A. S. de Almeida, A. rmando M. alheiro da Silva, M.ário J. B. Franco, & C. arla C. de Freitas (Orgs.), *Educação, gestão da informação e sustentabilidade*. (cap. 10, p. 213-246). Porto: Faculdade de Letras da Universidade do Porto. ISBN 978-972-8932-99-2.

Yin, R. K. (2006). *Estudo de caso: planeamento e métodos*. (3ª ed.). Porto Alegre: Bookman.

**PERSPETIVAS DE INVESTIGAÇÃO
EM REPRESENTAÇÃO E ORGANIZAÇÃO DO
CONHECIMENTO: ATUALIDADE E TENDÊNCIAS**



UMA ABORDAGEM FENOMENOLÓGICA ÀS ORGANIZAÇÕES INTELIGENTES: A PERSPETIVA DOS ESTUDANTES DE PÓS-GRADUAÇÃO

Luís Corujo¹, Jorge Revez²

¹Universidade de Lisboa, Faculdade de Letras, 0000-0003-4411-2453, luiscorujo@campus.ul.pt,

²Universidade de Lisboa, Faculdade de Letras, 0000-0002-3058-943X, jrevez@campus.ul.pt

RESUMO A avaliação dos sistemas de informação tem frequentemente relegado para segundo plano o elemento humano. Partindo da análise de Choo relativa à Organização e Gestão do Conhecimento no seio da Organização Inteligente, e da proposta metodológica da Fenomenologia, são investigadas as experiências de estudantes de cursos pós-graduados (2.º e 3.º Ciclos do Ensino Superior) lecionados em Portugal, relativamente aos procedimentos administrativos nas Instituições de Ensino Superior, perspetivados no âmbito dos sistemas de informação das organizações. Para o efeito, efetuam-se entrevistas eletrónicas estruturadas que permitem colher as perceções dos estudantes face aos sistemas de informação, particularmente no que se refere aos procedimentos administrativos, e que evidenciem fenómenos de insuficiência, falhas de conceção, funcionamento e desalinhamento, captados pelas vivências dos sujeitos, e que tenham resultado, por exemplo, em frustrações, sugestões de melhoria, etc. Os resultados informam sobre fenómenos reais, porque percecionados por várias pessoas, mas que não costumam ser documentados e, logo não estudados. De uma forma geral, as Instituições de Ensino Superior são percecionadas de forma positiva, apresentando ainda algumas falhas no funcionamento dos sistemas de informação.

PALAVRAS-CHAVE *Gestão do Conhecimento, Sistemas de Informação, Procedimentos Administrativos, Ensino Superior*

ABSTRACT The information systems assessment has often relegated the human element to the background. Based on Choo's analysis of the Organization and Knowledge Management within the Intelligent Organization, and Phenomenology's methodological proposal, we investigate the experiences of students of Portuguese postgraduate courses (2nd and 3rd Cycles of Higher Education) regarding the administrative procedures in Higher Education Institutions, within the scope of its information systems. For this purpose, we carried out structured electronic interviews to gather students' perceptions of information systems, particularly in regard to administrative procedures, and specifically highlighting phenomena of insufficiency, misconfiguration, misalignment and lack of knowledge captured in subjects' experiences, and resulting, for instance, in frustration, suggestions for improvement, etc. The results report real phenomena, because they are perceived by several people, but usually are not documented and thus not studied. In general, the Higher Education Institutions are perceived under a positive light, but still display some failures in the functioning of their information systems.

KEYWORDS *Knowledge Management, Information Systems, Administrative Procedures, Higher Education*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

O contacto dos estudantes de pós-graduação com as instituições do ensino superior, no que concerne a procedimentos administrativos e tarefas burocráticas, representa, do ponto de vista organizacional, colocar frente a frente clientes muito exigentes e supostas organizações de vanguarda.

Choo (2002, p. xiv) considera que, no âmbito das organizações, a criação e a utilização de Conhecimento são experiências sociais, nas quais múltiplos atores participam e exercitam os seus diferentes raciocínios e interesses, no sentido de modificar o seu comportamento como reflexo de novo Conhecimento (p. 1-22). Choo sugere que o Conhecimento Organizacional pode ser dividido em Conhecimento Tácito, Conhecimento Explícito e Conhecimento Cultural (p. 263-273). O primeiro refere-se ao Conhecimento implícito utilizado pelos membros da organização para desempenhar o seu trabalho com competência, sendo difícil de verbalizar porque se expressa por intermédio de competências baseadas na ação e não pode ser reduzido a regras ou fórmulas. O segundo remete para o Conhecimento que foi codificado formalmente ou tornado tangível na figura de um artefacto físico, podendo ser facilmente comunicado em objetos ou em regras. O último diz respeito às suposições e crenças partilhadas acerca dos objetivos, capacidades, clientes e competidores da organização, sendo que tais crenças são utilizadas para atribuir valor e significado a nova Informação e Conhecimento.

Uma vez que o Conhecimento é construído com base na experiência acumulada (p. 257), este autor considera que a Organização Inteligente se caracteriza por ser capaz de triar os seus recursos e competências informacionais, para transformar Informação em Conhecimento, e utilizar esse Conhecimento para sustentar e melhorar o seu desempenho num ambiente em mudança, identificando a finalidade da gestão da informação com a transformação da Informação em Conhecimento (p. xiv). Este autor entende a gestão da informação como a gestão de uma rede de processos que adquirem, criam, organizam, distribuem e usam informação (p. xiv).

Vários autores consideram que os sistemas de informação são elementos com potencial facilitador para lidar com o Conhecimento nas Organizações, nomeadamente na criação de novos produtos e serviços, no desenvolvimento de novas ideias e saber-fazer, e na partilha de visões e boas práticas (Davenport & Prusak, 1998; Detlor, 1999). Assim, é possível afirmar a necessidade de desenvolver sistemas de informação utilizados para tramitar informação produzida e recebida no âmbito dos procedimentos administrativos referentes às atividades da organização, e que esse desenvolvimento deve ser feito no seio de uma gestão da informação planificada, que possibilite constituir a informação como um capital organizacional que permite criar e utilizar Conhecimento (Tácito, Explícito e Cultural), e que resulte na identificação da instituição como organização inteligente.

Na perspetiva de Taylor (1982, p. 341), um sistema de informação inclui um conjunto de atividades que se caracterizam por serem processos, definidos formalmente, que selecionam, adquirem, organizam, armazenam, recuperam, apresentam, analisam e/ou interpretam mensagens, aqui entendidas como informação. Esses processos acrescentam valor às mensagens de várias formas, sendo que estas mensagens são consideradas como comunicações formais que as pessoas desenvolvem e emitem de alguma forma, como uma publicação, registo ou apresentação. Estas mensagens são também geradas por intermédio de vários sistemas e processos computacionais, e podem ser tratadas, formatadas ou organizadas através de instruções programadas, podendo ser solicitadas ou encaminhadas automaticamente a outras pessoas que poderão, ou não, fazer uso delas de diversas formas. Este autor

identifica como sistemas de informação, sistemas burocráticos, biblioteconómicos, de processamento de dados, de gestão de informação, de apoio à tomada de decisão, e de recuperação de informação.

Apesar desta perspetiva formal, considera-se que devem ser tomados em consideração outros aspetos, como aqueles que estão ligados à comunicação de informação e/ou interação com o sistema de informação, que estão para lá do formalizado em manuais de procedimentos, ligados ao Comportamento Informacional, Tradição Organizacional e capacidade de resolução de problemas não previstos, e que se manifestam muitas vezes pela existência de graus de informalidade decorrentes da comunicação verbal (por exemplo sistemas com informação resultante de dados fornecidos verbalmente e não por registo escrito), e vícios burocráticos, muitas vezes consuetudinários, em que a razão de ser do procedimento administrativo é ultrapassado pelo rito criado hábito.

No entanto, existem situações em que estes procedimentos administrativos e sistemas de informação relacionados, aparentemente "perfeitos" do ponto de vista das abordagens e modelos de planificação e que servem de base aos projetos que os formalizaram e lhes deram origem, mormente em sede de alterações apelidadas de *Modernização Administrativa*, acabam por transparecer disfunções, desequilíbrios, falhas, problemas de funcionamento, e que geram entropias por não ter sido considerada a perspetiva dos utilizadores externos, sejam eles cidadãos, contribuintes, beneficiários, clientes, alunos, podendo inclusivamente prejudicá-los e atentar contra os seus direitos.

Com efeito, Choo (2002, pp. 238–240) alerta para o facto de os utilizadores da informação, apesar de serem a “razão de ser” da atividade informacional (e organizacional), são apenas episodicamente ou periféricamente envolvidos no desenvolvimento dos sistemas de informação, originando um fosso entre as necessidades reais de informação e a informação captada e fornecida através dos sistemas e serviços de informação, que se torna ainda mais explícito pela proeminência dada às Tecnologias da Informação e os seus especialistas. O autor fornece uma alternativa que se afasta da fragmentação funcional, apelando a uma visão holística e transversal que una os recursos e capacidade da organização para criar, organizar e utilizar o Conhecimento e desenvolver as infraestruturas que efetivem a Gestão do Conhecimento. Tal passa pelo trabalho conjunto de três grupos de especialistas: os especialistas do domínio, empenhados em criar e utilizar o Conhecimento (operadores, profissionais, tecnólogos, gestores); os especialistas da Informação, com competências, formação e experiência na Organização do Conhecimento em sistemas e estruturas que facilitam a utilização produtiva dos recursos de Informação e Conhecimento (bibliotecários, arquivistas, gestores documentais, etc.); e os especialistas nas Tecnologias da Informação, peritos na modelação da infraestrutura informacional da organização (analistas de sistemas, *designers* de sistemas, engenheiros de *software*, programadores, administradores de dados, gestores de redes informáticas, e outros especialistas que desenvolvem sistemas e redes de informação computadorizados).

Neste sentido, considera-se importante estudar a perspetiva dos utilizadores dos sistemas de informação no âmbito dos procedimentos administrativos, encarando-os como clientes externos à organização. Para tal, pretende-se recorrer a uma abordagem fenomenológica que permita captar as experiências e perceções dos sujeitos, e cuja análise possibilite explorar um novo trilha de estudos empíricos sobre questões que não costumam ser documentadas e, logo não estudadas.

Lyotard considera a Fenomenologia como o estudo dos fenómenos, do que aparece à consciência, daquilo que é dado, que deve ser explorado tal como «coisa que se percebe, em que se pensa, de que se fala, evitando forjar hipóteses, tanto sobre o laço que une o fenómeno, com o ser de que é fenómeno, como sobre o laço que o une como o Eu para quem é fenómeno» (2008, p. 9). Com base neste autor

verifica-se que a Fenomenologia tem origem nas ideias de Husserl (1859-1938) e nos desenvolvimentos derivados de Heidegger e Merleau-Ponty, sendo utilizada atualmente como metodologia pelas ciências humanas e sociais (Andrade & Neves, 2015; Coutinho, 2013), e pela Ciência da Informação em particular (Budd, 2005; Figueiredo, 2012). Para Coutinho (2013, p. 349) a investigação fenomenológica caracteriza-se por colocar a tónica no «individual» e na «experiência subjetiva». Assim, o seu propósito é descrever o fenómeno na perspectiva dos participantes, estudando as experiências vivenciadas pelos indivíduos, tal como são descritos (interpretados) pelos sujeitos. A descrição feita pelo investigador aponta para a essência das experiências de vários indivíduos que experienciaram o mesmo fenómeno (Creswell, 2014, p. 14). Destes estudos empíricos poderão surgir estudos, eventualmente, de carácter comparativo, e que permitirão, por um lado, teorizar sobre os problemas identificados e, por outro, procurar corrigir as discrepâncias e desajustes dos sistemas de informação já existentes e desenvolver novos sistemas com maior sensibilidade e que melhor respondam aos anseios e necessidades dos utilizadores externos.

O objeto de estudo deste trabalho inclui os sistemas de informação relativos a procedimentos administrativos organizacionais, sejam eles digitais, analógicos, formalizados em manuais de procedimentos, ou informais, verbais ou registados, no caso particular das Instituições de Ensino Superior. Este objeto é percecionado pelo aluno de pós-graduação (2º e 3º ciclos), aqui considerado sujeito consciencializado como utilizador externo que recorre ao sistema de informação, enquadrado nos procedimentos administrativos como forma normalizada de comunicação com a entidade. Mas este sistema de informação também é percecionado como mais uma camada (burocrática) colocada entre a organização e o sujeito, cuja relação (troca de informação, prestação de bem e/ou serviço) se processa como uma osmose por intermédio dessa camada. O método de recolha de dados escolhido permite captar as perceções dos sujeitos resultantes da interação com as Instituições de Ensino Superior e os seus procedimentos administrativos, por intermédio dos sistemas de informação, adequando-se desta forma à abordagem fenomenológica deste estudo.

Como hipótese de trabalho, a investigação propõe analisar se os resultados mostram ou não que as Instituições de Ensino Superior apresentam falhas graves no funcionamento dos sistemas de informação e que os utilizadores se consideram atores irrelevantes no processo de avaliação desses sistemas e reclamam, em abstrato, vontade de participar no desenvolvimento dos sistemas de informação.

METODOLOGIA

Esta investigação parte da questão: com base nas experiências dos estudantes de cursos pós-graduados, qual é a perceção que estes têm relativamente aos procedimentos administrativos nas Instituições de Ensino Superior, perspetivada no âmbito dos sistemas de informação das organizações? Tal remete para um estudo de carácter qualitativo baseado numa abordagem fenomenológica.

Para o efeito, optou-se, como método de recolha de dados, por efetuar entrevistas eletrónicas estruturadas que permitisse colher as perceções dos estudantes face aos sistemas de informação, e que evidenciasse fenómenos de insuficiência, falhas de conceção, funcionamento e desalinhamento, captados pelas vivências dos sujeitos, e que tenham resultado, por exemplo, em frustrações, sugestões de melhoria, etc.

Coutinho (2013, p. 351) refere a entrevista não estruturada como método de recolha de dados indicado para os estudos fenomenológicos, utilizando uma amostra intencional dos indivíduos e centrando-se nas experiências passadas, sendo a análise de dados orientada pelo significado, procurando temas e padrões através dos participantes, caracterizando-se por serem abertos, relacionados com a experiência e intuição, resultando em descrições temáticas.

Para tal, foi desenhado um guião das entrevistas com dez questões, destinando-se três perguntas para a caracterização dos sujeitos e sete perguntas de resposta aberta que visavam a recolha das perceções dos estudantes em diferentes momentos da sua experiência de contacto com os sistemas de informação, explícitos nos procedimentos administrativos, das Instituições de Ensino Superior. Os dados foram recolhidos entre 17 e 30 de maio de 2017, junto de um conjunto de estudantes do ensino superior português na área da Ciência da Informação, inscritos entre 2012 e 2016. Recorreu-se à plataforma *Google Forms* e os resultados foram posteriormente organizados e preparados para análise numa folha de cálculo de *Excel*. Foi realizada então, com o suporte do software *ATLAS.ti*, a leitura e a interpretação dos dados, a sua codificação por temas e descrições, de forma a tornar possível interpretar os significados desses temas/descrições. Apesar das críticas que os processos de codificação através de *software* têm suscitado na literatura (Goble, Austin, Larsen, Kreitzer, & Brintnell, 2012), procurou realizar-se aqui uma abordagem com um grau de complexidade baixo, que permitisse abranger a totalidade dos dados mas sem recorrer a métodos demasiado exaustivos, que este estudo não exigia.

Realizou-se ainda um pré-teste com dois estudantes de pós-graduação, escolhidos por conveniência e facilidade de contacto. Este pré-teste revelou a necessidade de melhorar algumas formas verbais das perguntas formuladas, que ficaram registadas da seguinte forma:

- 1) Qual o último Ciclo de Estudos em que se inscreveu?
- 2) Qual o ano de início do último Ciclo de Estudos em que se inscreveu?
- 3) Qual a Instituição de Ensino Superior onde se inscreveu?
- 4) Qual a perceção genérica que tem sobre o funcionamento dos serviços administrativos da Instituição do Ensino Superior que frequenta/frequentou?
- 5) No momento da Candidatura ao Ciclo de Estudos, como analisa os procedimentos administrativos em que esteve envolvido (facilidades, constrangimentos, insuficiências, ...)?
- 6) No momento da Inscrição no Ciclo de Estudos, como aprecia os procedimentos administrativos em que esteve envolvido (facilidades, constrangimentos, insuficiências, ...)?
- 7) Qual a sua perceção sobre o acesso à informação das Unidades Curriculares (pautas, sumários, folhas de presenças, ...)
- 8) Como avalia os procedimentos administrativos relativos à informação sobre propinas (prazos, pagamentos, atrasos, multas, ...)
- 9) Como analisa os procedimentos administrativos relativamente a Requerimentos e pedidos de Certidões/Declarações?
- 10) Comente por favor os mecanismos disponíveis para Reclamações, Elogios, Sugestões e Pedidos de ajuda.

A amostra foi escolhida intencionalmente para incluir estudantes dos 2.º e 3.º Ciclos da área da Ciência da Informação, por serem aqueles que poderão construir uma visão mais estruturada dos problemas em análise. Por outro lado, estes estudantes apresentam à partida uma maior disponibilidade para refletir e

escrever sobre este tipo de problemáticas. Para o lançamento da entrevista, foi solicitado o apoio dos responsáveis pela área da Ciência da Informação na Universidade de Lisboa e na Universidade de Coimbra, aos quais os autores agradecem.

RESULTADOS

Os resultados informam sobre fenómenos reais, porque percecionados por várias pessoas, mas que não costumam ser documentados e logo não estudados. Como em qualquer estudo qualitativo, a generalização dos resultados só ocorre quando são estudados casos adicionais e se generalizam os resultados aos novos casos, tal como ocorre com a replicação usada na investigação experimental. No entanto, a repetição dos resultados de um estudo de caso num novo cenário implicam procedimentos qualitativos bem documentados, nomeadamente a existência de um protocolo para documentar detalhadamente o problema e o desenvolvimento de uma meticulosa base de dados de estudos de caso (Creswell, 2014).

Neste estudo empírico foram recolhidas 30 respostas completas. Este número não pretende representar qualquer universo definido nem dele se pode inferir dimensões quantitativas de análise. O que importa são as experiências mais ou menos relatadas nas respostas dos participantes.

As três primeiras questões relativas à caracterização dos respondentes (Figura) mostram uma maior participação de estudantes do 2.º Ciclo, inscritos sobretudo nos dois últimos anos e vinculados às duas universidades às quais foi solicitado apoio na divulgação da entrevista: a Universidade de Lisboa e a Universidade de Coimbra. Neste campo, o segmento da identificação dos participantes não contribuiu com qualquer resultado inesperado.

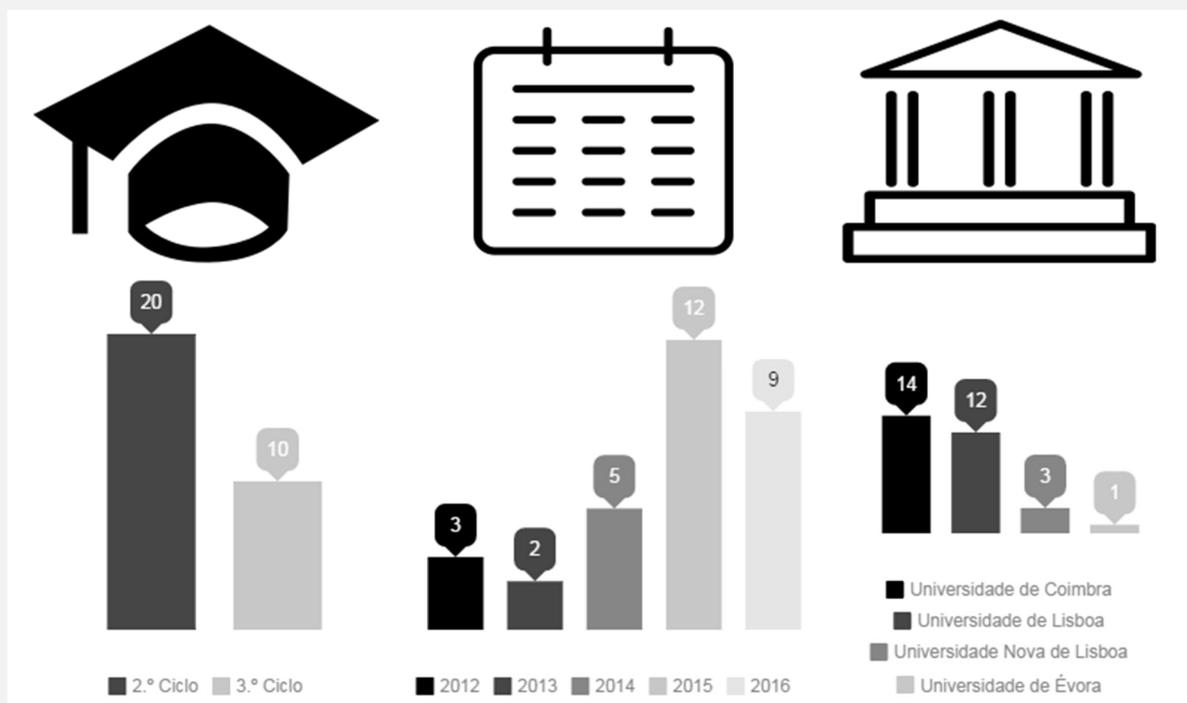


Figura 1. Participantes (N=30) por ciclo de estudos, ano de inscrição e universidade

Entre as várias análises possíveis, mediante as potencialidades qualitativas do software *ATLAS.ti*, optou-se, em primeiro lugar, pela codificação das percepções, tendo por base o conteúdo das respostas, percepções essas entendidas na polarização «Positivo» e «Negativo». Como se pode observar (Figura), a contagem das citações codificadas no texto das respostas demonstra uma prevalência de percepções e opiniões positivas, o que contraria a hipótese inicial de trabalho. Refira-se que, do ponto de vista da densidade do conteúdo das respostas, as opiniões positivas estão maioritariamente associadas a respostas curtas, ao contrário das opiniões negativas que em geral correspondem a respostas mais longas e detalhadas. Isto pode explicar-se pela procura de, no caso negativo, tentar justificar a sua percepção e/ou dar voz às frustrações causadas pelos problemas e entropias associados aos procedimentos administrativos.

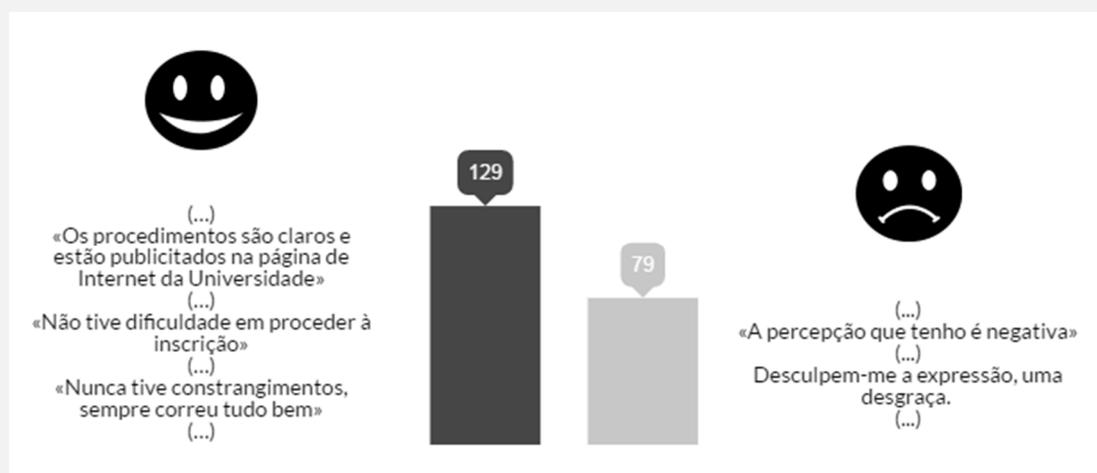


Figura 2. Contagem do tipo de citações após análise das respostas (dados codificados com recurso ao software *ATLAS.ti*), com a inclusão de alguns exemplos

Quando se analisa (Figura) a correlação entre as perguntas colocadas (que correspondem a diferentes momentos de percepção dos fenómenos) e a lista de códigos relativos às opiniões e percepções manifestadas pelos entrevistados, conclui-se que as três categorias mais referidas pelos participantes são o apoio dos serviços, a acessibilidade e a clareza da informação, o que aponta para uma percepção francamente positiva e uma manifestação do que parece ser relevante explicitar durante a análise daqueles fenómenos. Estes elementos devem ser tidos em conta e problematizados no momento em que as instituições pretendam melhorar os procedimentos administrativos e os sistemas de informação, com o fim de se tornarem organizações inteligentes. O papel dos utilizadores torna-se assim, na prática, determinante para essa transformação.

Na quarta e quinta posição da lista ordenada (Figura) surgem a ineficiência (alocação desajustada de recursos) e a insatisfação, que remetem para o pólo negativo. A soma destas cinco primeiras categorias suplanta quase em dobro o total de ocorrências de todas as outras categorias, o que é demonstrativo da sua importância para os entrevistados. Os dados permitiam ainda correlacionar os códigos atribuídos com as Universidades mencionadas, mas esse escopo afigurou-se pouco relevante para este estudo.

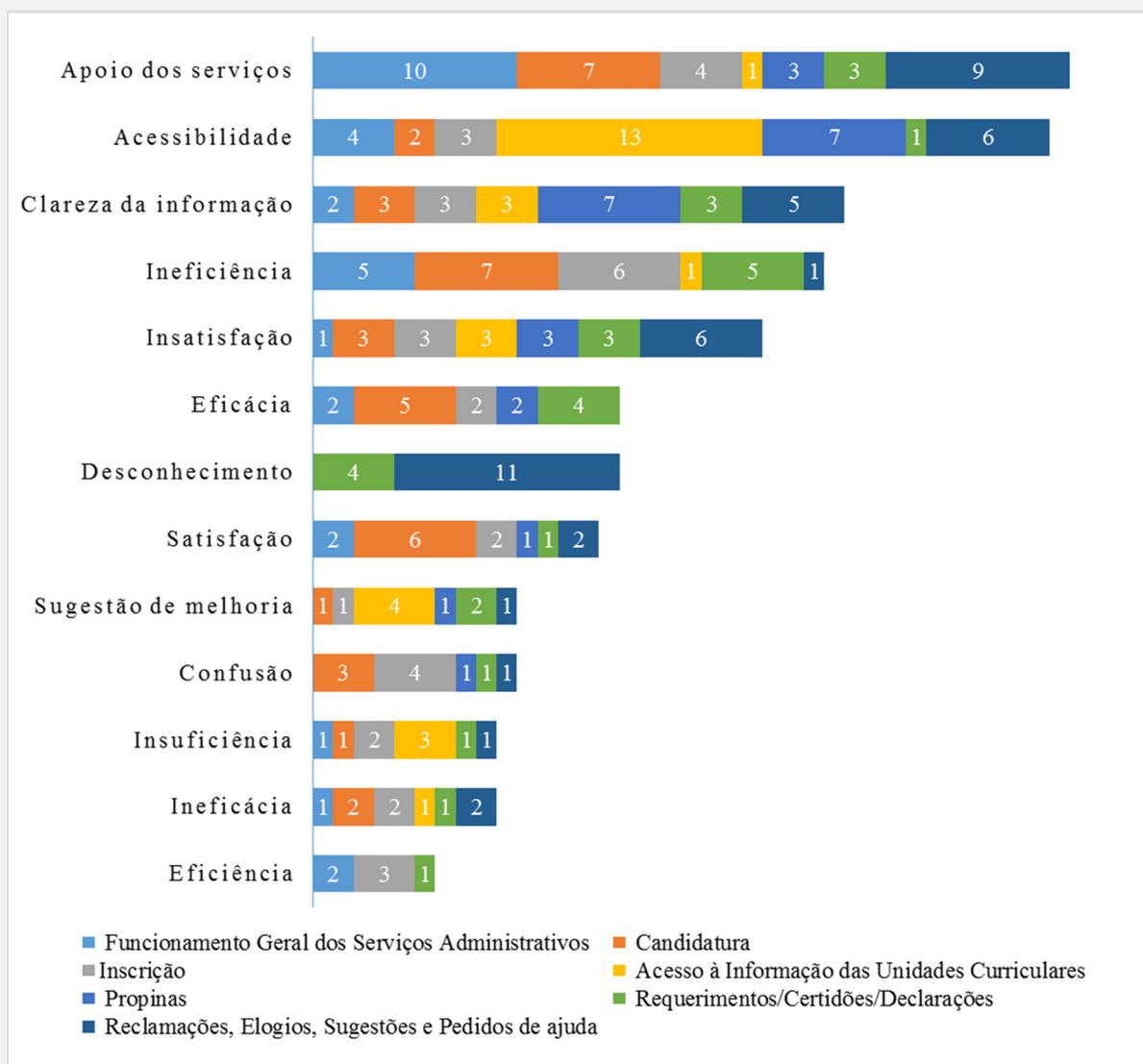


Figura 3. N.º de Ocorrências de cada Opinião/Percepção por Pergunta

Finalmente, foi confrontada a perceção positiva e negativa face a cada um dos momentos da relação entre os estudantes e as Instituições do Ensino Superior (Figura). Verifica-se que o funcionamento dos procedimentos ligados às «Propinas» sobressai em termos de perceção positiva dos alunos, sendo que o único aspeto percecionado negativamente que ultrapassa numericamente as perceções positivas refere-se aos procedimentos ligados às «Reclamações, Elogios, Sugestões e Pedidos de Ajuda». Pode ainda referir-se que o item «Requerimentos/Certidões/Reclamações» apresenta o valor mais equilibrado entre a dimensão positiva e negativa, na qual se pode antever um conjunto de perceções ambivalentes.

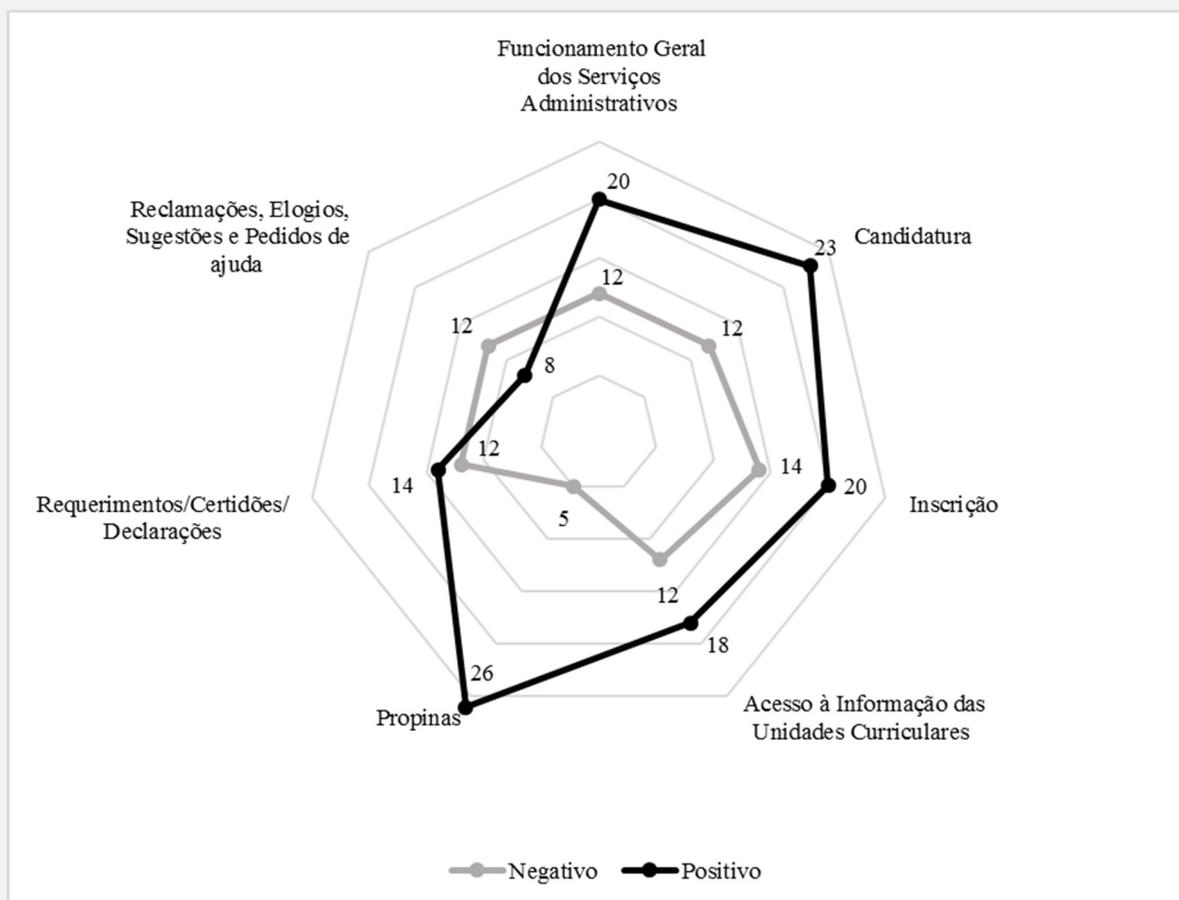


Figura 4. N.º de Ocorrências da Percepção Positiva e Negativa por Pergunta

Na observação da Figura é ainda claro que não existe uma correspondência exata entre o número de respostas (N=30) e o número de ocorrência ou citações codificadas em «Positivo» ou «Negativo», o que demonstra que em cada pergunta respondida surgiam amiúde opiniões diversas, como é exemplo o seguinte excerto de uma resposta ao item «Candidatura»: «A possibilidade de tratar da candidatura através da internet (online) facilita um pouco o processo, no entanto, por causa de falta de informação, informação incompleta, confusa e por vezes, duplicada, há sempre imensas dúvidas que entopem todos os anos os serviços administrativos» (Entrevista 23).

Por fim, anota-se ainda que o momento «Reclamações, Elogios, Sugestões e Pedidos de ajuda» foi aquele que obteve um número de ocorrências (apenas 20) claramente inferior ao número de respostas obtidas, na medida em que foi o elemento mais sujeito a citações do código «Desconhecimento». Este código revela ainda a percepção do afastamento dos entrevistados relativamente aos processos de melhoria e transformação dos sistemas.

CONCLUSÕES

Apesar de não se poder efetuar uma generalização dos resultados obtidos nos estudos de carácter fenomenológico, este estudo permite uma aproximação às disfunções dos sistemas de informação organizacionais. Um dos problemas deste tipo de trabalhos prende-se com a subjetividade inerente às descrições dos sujeitos entrevistados e à análise consequente dos dados. Tal implica a utilização de

estratégias de validação como a triangulação, a verificação pelos membros da equipa de investigação e por outros pares, a descrição detalhada e profunda dos procedimentos e ainda esclarecer o viés que os investigadores possam transportar para a investigação e apresentar informação que contraria os temas apresentados (Creswell, 2014).

Em suma, as respostas dos entrevistados permitem plasmar um maior número de perceções positivas em relação às negativas. É interessante que a identificação dos procedimentos a que se referem essas perceções positivas – nomeadamente o elemento «Propinas» - permita uma interpretação que indicie uma valoração e um maior investimento nos mecanismos de captação de receitas por parte das Instituições de Ensino Superior. Aparentemente, é dada mais importância ao bom funcionamento de procedimentos como as «Propinas», do que aqueles ligados às funções e fins do ensino superior, incluindo os procedimentos que permitiriam melhorar o funcionamento, como as reclamações e sugestões, tal como se verifica na seguinte resposta: «Curiosamente a informação relativa a taxas e emolumentos é sempre muito clara e direta, o mesmo nem sempre se pode dizer acerca de outros assuntos» (Entrevista 23).

No entanto, com base em Choo (2002), consideramos que a Informação resultante deste trabalho tem potencial para ser usada para criar Conhecimento organizacional passível de ser utilizado pelas instituições, conduzindo o seu foco para uma ação sobre os seus sistemas de informação, no sentido de os melhorar e, assim, modificar o seu comportamento, com vista a atingir um melhor desempenho e resposta ao ambiente externo. Para tal propõe-se o desenvolvimento de questionários para estudantes, funcionários, docentes e investigadores que digam respeito à perceção dos fenómenos ligados aos procedimentos identificados por este trabalho.

Respondendo à hipótese de trabalho, que em geral não se confirma, os resultados mostram que as Instituições de Ensino Superior são percecionadas de forma positiva, apresentando ainda algumas falhas no funcionamento dos sistemas de informação. Por outro lado, verifica-se que os utilizadores se consideram atores com uma reduzida influência no processo de avaliação desses sistemas e reclamam, em número reduzido, a vontade de participar no desenvolvimento dos sistemas de informação.

REFERÊNCIAS BIBLIOGRÁFICAS

- Andrade, W. O. de, & Neves, D. A. de B. N. (2015). O uso da fenomenologia e do protocolo verbal na ciência da informação: estudos de usuários. Em *Desafios y oportunidades de las Ciencias de la Información y la Documentación en la era digital*. Madrid: Universidad Complutense de Madrid.
- Budd, J. M. (2005). Phenomenology and information studies. *Journal of Documentation*, 61(1), 44–59.
- Choo, C. W. (2002). *Information Management for the Intelligent Organization: the Art of scanning the Environment* (3ª). Medford, NJ: ASIS&T.
- Coutinho, C. P. (2013). *Metodologia de Investigação em Ciências Sociais e Humanas: Teoria e Prática* (2ª). Coimbra: Almedina.
- Creswell, J. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (3ª). Thousand Oaks, CA: Sage Publications.

Davenport, T. H., & Prusak, L. (1998). *Working knowledge: How organizations manage what they know*. Boston, Massachusetts: Harvard Business School Press.

Detlor, B. (1999). Utilizing Web Information Systems for Organizational Knowledge Work: An Investigation of the Information Ecology and Information Behaviors of Users in a Telecommunications Company. Em *62nd Annual Meeting of the American Society for Information Science* (pp. 1–11). Washington DC: ANSI.

Figueiredo, M. F. de. (2012). Pós-fenomenologia e Ciência da Informação: aportes epistêmicos para acesso ao conhecimento. *InCID: R. Ci. Inf. e Doc.*, 3(1), 21–35.

Goble, E., Austin, W., Larsen, D., Kreitzer, L., & Brintnell, S. (2012). Habits of Mind and the Split-Mind Effect: When Computer-Assisted Qualitative Data Analysis Software is Used in Phenomenological Research. *Forum: Qualitative Social Research*, 13(2), art. 2. Obtido de <http://nbn-resolving.de/urn:nbn:de:0114-fqs120227>

Liotard, J.-F. (2008). *A Fenomenologia*. Lisboa: Edições 70.

Taylor, R. S. (1982). Value-Added Processes in the Information Life Cycle. *Journal of the American Society for the Information Science*, 33(5), 341–346.



O PAPEL DAS DEFINIÇÕES NA FORMAÇÃO DE CORPUS TERMINOLÓGICO PARA ELABORAÇÃO DE ONTOLOGIAS DE DOMÍNIOS: UMA EXPERIÊNCIA A PARTIR DE EXTRAÇÃO AUTOMÁTICA E MANUAL

Maria Luiza de Almeida Campos¹, Renato Rocha Souza²

¹Universidade Federal Fluminense, Departamento de Ciência da Informação, Programa de Pós-Graduação em Ciência da Informação, 0000-0002-9253-3706, maria.almeida@pq.cnpq.br

²Fundação Getúlio Vargas, Escola de Matemática Aplicada & Universidade Federal de Minas Gerais, Escola de Ciência da Informação, 0000-0002-1895-39052, renato.souza@fgv.br

RESUMO Compara-se o procedimento automático e o manual para a formação de corpus terminológico em Ontologia no domínio do Folclore e Cultura Popular Brasileira. O material utilizado para a aplicação dos procedimentos são as definições existentes em fontes do domínio de aplicação. Utilizou-se o método comparativo, o que possibilitou apresentar uma série de aspectos para uma tomada de decisão quanto ao melhor procedimento a ser adotado em domínios de aplicação.

PALAVRAS-CHAVE *Ontologia de Domínio, Formação de Corpus Terminológico, Extração Automática, Extração Manual, Definição.*

ABSTRACT Compares the automatic and manual procedure for the creation of a terminological corpus for a domain Ontology in the field of Folklore and Brazilian Popular Culture. The material used for the application of the procedures are the existing definitions in sources of the application domain. The comparative method was used, which allowed to present a series of aspects for a decision on the best procedure to be adopted in application domains.

KEYWORDS *Domain Ontology, Terminological Corpus, Automatic Text Mining, Manual Extraction, Term Definitions.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A Ontologia de Domínio pode ser definida como um modelo conceitual que representa o conhecimento de uma área de aplicação (área de assunto, de missão ou de problema) e que deve servir como uma representação de consenso (ou modelo de referência) de uma conceituação compartilhada por uma determinada comunidade. Um dos primeiros procedimentos para a elaboração de Ontologias de Domínio é o levantamento de *corpus* terminológico consistente que descreva este domínio. De forma geral, as propostas metodológicas (Fernandez, Gomez-Perz, & Juristo, 1997; Ding & Foo, 2002a e b; Fernandez-López, 2001; Gangemi, Pisanelli, & Steve, 1999; Gruninger & Fox, 1995; Guarino & Welty, 2002; Metaxiotis, Psarras, & Askounis, 2001; Mizoguchi, 2014; Paslaru-Bontas, 2007; Sure, Staab, &

Studer, 2002; Uschold & King, 1995) pressupõem a existência ou mesmo a captura de uma estrutura de conceitos já existentes, sugerindo, a partir dessa estrutura, critérios para verificar a explicitação de propriedades e relações entre os conceitos. Entretanto, temos observado a necessidade de estudos que venham a discutir procedimentos metodológicos para a formação de *corpus* terminológico em Ontologia de Domínio.

Em Tesouros, na maioria das vezes, a obtenção do *corpus* se dá através de procedimento de indexação de documentos, pois ele funciona como um instrumento que auxilia na padronização dos assuntos dos documentos. Ontologias, por outro lado, têm por função permitir a descrição do conhecimento existente em um domínio; sendo o insumo, o conhecimento de especialistas de um domínio, em vez do conteúdo de documentos que, nesse caso, possuem função acessória. Assim, faz-se necessária a investigação de procedimentos que possam permitir a formação de um *corpus* terminológico que representa um dado domínio de especialidade. Esse *corpus*, consistentemente constituído, agregando as relações e axiomas, permite que se criem condições para a implementação de raciocínio automático.

Em domínios de especialidade, as definições são elementos que explicitam o conteúdo semântico de uma entidade. Tendo em vista que as entidades são descritas e definidas umas em relação às outras através de enunciados definitórios (Dahlberg, 1981), consideramos a definição como elemento fundamental para a descoberta deste *corpus* terminológico em Ontologias e consideramos ainda a possibilidade de descoberta deste *corpus* de forma automática ou semiautomática, na perspectiva de otimizar esta tarefa e possibilitar uma maior abrangência, rapidez e efetividade na etapa de identificação do *corpus*.

A determinação de definições como um ponto de partida para o levantamento do *corpus* terminológico em Ontologia de Domínios se sustenta através de uma abordagem onomasiológica. Essa abordagem se caracteriza pela construção do conceito a partir do entendimento de uma concepção consensual da realidade sobre um objeto social, em que se identificam processos de seleção, realce e iluminação de alguns aspectos do ser e, por conseguinte, atenuação, ou mesmo ocultamento, de outros aspectos. Esta “versão do referente” está consolidada nas definições (Dahlberg, 1978a; Riggs, 1979, 1989a, b, 1996).

Dessa forma, consideramos que as definições podem ser o elemento básico para o entendimento do objeto em um determinado domínio: elas evidenciam as propriedades de um dado objeto através de relações que podem ser identificadas nos elementos de seu enunciado definitório. O enunciado definitório é composto em sua sintaxe linguística de verbos que indicam conceitos antecedentes e sucessores relacionados. Esses elementos antecedentes e sucessores formam o *corpus* terminológico do domínio à medida que deverão ser identificados. Por outro lado, selecionados os verbos, estes podem ser estudados sob o ponto de vista da identificação da tipologia de relações existentes entre os elementos antecedentes e sucessores na definição, que podem ser determinados e analisados por humanos. Entretanto, pretendemos investigar em que medida é possível a formação de *corpus* terminológico por mecanismos automáticos que possam ocorrer de forma eficiente, possibilitando uma maior otimização da tarefa em menor tempo possível. A partir deste levantamento, pretendemos futuramente também identificar e classificar relações semânticas que ocorrem em determinado domínio e, além disso, padrões de enunciados definitórios.

A formação dos *corpora* dar-se-á através de captura automática de textos – em um processo conhecido como *crawling* (que consiste na captura programática de informações textuais de sítios da Internet, usualmente através de programas específicos ou genéricos, desenvolvidos em alguma linguagem de programação) – seguida de seleção e curadoria, segundo a relevância das fontes em que as definições

são identificadas. A partir dessa etapa, os textos são processados através de um marcador morfossintático, necessário à descoberta das instâncias sintáticas de interesse (verbos). Esta etapa é desenvolvida com o uso do Parser Palavras (Bick, 2000). Uma vez que os verbos que denotam definições são identificados, sua vizinhança no enunciado é inspecionada para a identificação das entidades relacionadas e que explicitam informações adicionais sobre o conceito em análise.

O campo empírico para a identificação do *corpus* terminológico foi o domínio de Folclore e Cultura Popular Brasileira. Utilizamos o método comparativo para analisar o procedimento automático e manual, em que identificamos semelhanças e diferenças na formação do *corpus*. A abordagem manual foi realizada a partir da análise das definições, seguindo princípios estabelecidos por Dahlberg em diversos estudos relacionadas às definições conceituais (Dahlberg, 1978a, b, 1981, 1983).

Esta pesquisa é parte de dois estudos desenvolvidos com apoio do CNPq relacionados aos interesses de investigação de seus autores; um trata de aspectos teóricos e metodológicos para a elaboração de Modelos de Domínios visando à representação do conhecimento em Ontologias de Domínios, e o outro pesquisa a possibilidade de incorporação de técnicas de processamento de linguagem natural para as tarefas de organização de conhecimento e indexação automática. É uma pesquisa interinstitucional que envolve duas Instituições de Ensino e Pesquisa Brasileira.

METODOLOGIA

Esta investigação se apoia em metodologia qualitativa utilizando o método comparativo. A “abordagem qualitativa apresenta certos elementos – como a recursividade, por exemplo – que implicam o fato de a análise poder se iniciar até mesmo ao longo da fase de coleta de dados” (Appolinário, 2006, p. 159). Algumas características podem ser identificadas em uma pesquisa que utiliza uma perspectiva qualitativa: a análise é sistemática e compreensiva, ou seja, exige rigor, mas não rigidez, mas encontrando sua compreensão a partir de padrões que possam ser significantes para o estudo; o processo de análise mais utilizado é a comparação, podendo construir e/ou aprimorar categorias, revelando padrões e, por fim, pode-se chegar ao exame de categorias e padrões revelados com base em teorias e estudos anteriores. O método comparativo, de forma direta, “permite analisar o dado concreto, deduzindo do mesmo os elementos constantes, abstratos e gerais” (Lakatos & Marconi, 2007, p. 107), ou, ainda, “procede pela investigação de indivíduos, classes, fenômenos ou fatos, com vistas a ressaltar as diferenças e similaridades entre eles” (Gil, 1995, p. 35). Segundo Colino (2002, p. 2), os objetivos da análise comparativa podem ser listados, como se seguem: (1) investigação de semelhanças e diferenças (variação) entre casos, que dá lugar à (2) observação de regularidades, que devem ser explicadas mediante (3) o descobrimento da covariação ou da interpretação da diversidade, que habitualmente faz-se através da (4) comprovação de hipóteses explicativas, que levam à (5) explicação da complexidade de relações causais e ao estabelecimento de generalizações ou pautas particulares que permitam, finalmente, (6) a generalização de teorias ou sua refutação.

Nesse sentido, utilizamos o método comparativo para identificar em que medida o levantamento automático é eficaz e eficiente para a formação de *corpus* terminológico em Ontologias de Domínios. A partir do resultado encontrado, após a utilização do recurso automático, comparamos com o resultado encontrado em levantamento de *corpus* de forma manual através das definições do domínio do Folclore e Cultura Popular Brasileira. Para tanto, esta pesquisa utilizou os seguintes procedimentos metodológicos, que são classificados em três categorias de operação, a saber: Operação Automática

(AO), Operação Manual (OM) e Comparação entre AO e OM. Na AO, foram os seguintes os passos desenvolvidos: 1. Estabelecimento de um problema/questão sobre o domínio do Folclore e Cultura Popular Brasileira; 2. Identificação das entidades que compõem este problema no domínio especificado; 3. Levantamento em fontes terminológicas selecionadas sobre a definição das entidades identificadas; 4. Análises automáticas das definições levantadas das entidades; 5. Identificação automática de elementos antecedentes e sucessores dessas definições; 6. Identificação de novo *corpus* terminológico a partir dos elementos antecessores e sucessores. Na OM, os passos foram os que se seguem: 1. Estabelecimento de um problema/questão sobre o domínio do Folclore e Cultura Popular Brasileira; 2. Identificação das entidades que compõem este problema no domínio especificado; 3. Levantamento em fontes terminológicas selecionadas sobre a definição das entidades identificadas; 4. Análise manual das definições levantadas das entidades, a partir dos pressupostos apresentados por Dahlberg no que concerne às características dos conceitos; 5. Identificação de outras entidades a partir desses pressupostos; 6. Identificação de novo *corpus* terminológico. Na terceira categoria de operação, foi realizada a comparação entre os fatores considerados semelhantes e diferentes, tendo em vista a otimização do processo de levantamento de *corpus* terminológico, a partir de critérios estabelecidos.

A ABORDAGEM ONOMASIOLÓGICA E A PERSPECTIVA DA CONCEPTOLOGIA

A abordagem onomasiológica se caracteriza pelo uso de critérios que permitem uma concepção consensual da realidade sobre um objeto social, refletindo a forma como o referente é conceitualmente construído. Em outras palavras, ao iluminar uma faceta do ser, a abordagem onomasiológica produz uma “versão” do referente que constitui o modo como determinado grupo de falantes interage socialmente com ele. Esta forma especial de apreender esta dimensão conceitual considera que a linguagem é um instrumento de construção de uma dada realidade e que é na língua em uso e referendada por um grupo de especialidade que se instauram, sustentam e alteram os processos sociais.

Esta concepção baseia-se no postulado da funcionalidade construtivista da linguagem e encontra alguns de seus elementos na “Lexicologia Construcionista”. De acordo com esse postulado, a linguagem desempenha, ao lado das outras funções (comunicativa, interacional, etc.), a tarefa fundamental de “construir” o referente do discurso (Lima, 2007).

Segundo Riggs (1979, 1989 a e b), dentro do campo da Linguística, existe uma variedade de disciplinas conexas, como Semântica, Sociolinguística, Psicolinguística e Planejamento da Linguagem, que podem reivindicar, todas elas, como uma de suas associadas, o campo do conhecimento chamado “Lexicologia”. De outro lado, relacionados com a Filosofia, encontramos o campo da Lógica, da Filosofia da Ciência, e a pesquisa em Classificação. Nesse contexto, emergiu um subcampo que focaliza a análise de conceitos, incluindo suas relações com a ciência como unidades do conhecimento teoricamente significativas, e suas relações com observações empíricas julgadas por critérios “operacionais” ou “denotativos”. Este subcampo foi denominado por Riggs conceptologia e está aberto a diferentes abordagens filosóficas e não precisa se unir a nenhuma delas.

Riggs considera a terminologia e a lexicografia áreas complementares e acredita que cada campo pode beneficiar-se do outro. Reconhece, no entanto, que os terminólogos têm familiaridade e usam a lexicografia, enquanto o inverso não é verdadeiro. O ruído existente para o entendimento das atividades nas duas áreas parece provir do fato que tanto uma área como outra usam por vezes o mesmo termo

para designar o produto de suas atividades: terminologia, glossário e até mesmo dicionário (Gomes, Campos, & Guimarães, 2010).

A IMPORTÂNCIA DAS DEFINIÇÕES EM ONTOLOGIAS DE DOMÍNIOS

Os estudos têm mostrado discussões sobre um vasto número de modos de definição e estes modos estão relacionados com o compromisso ontológico (acordo) assumido em cada domínio. Consideramos que as ontologias, entre outros aspectos, necessitam de um padrão para seus enunciados definitórios que possam apoiar o entendimento pelos agentes inteligentes (*software*). Consideramos, também, que partimos do conteúdo conceitual de uma unidade de representação, seja ela uma entidade ou uma instância, para que seja possível inferir sobre estas unidades um dado entendimento que possa ser “compreendido” por máquinas.

Os problemas de definição têm sido objeto de estudos diversos na língua geral. Na língua especializada, nos últimos 50 anos, a definição passou a ser objeto de estudo sistemático. Até a década de 1960, as definições tinham um caráter mais conceitual e filosófico. Entretanto, para atender às necessidades da ciência, foi necessário um esforço teórico visando à elaboração de definições operacionais – aquelas cujos elementos da definição não identificam somente o gênero próximo e a diferença específica – mas, também, elementos que procuravam definir certas operações nas quais o conceito seria aplicado (Campos, 2010).

No caso das ontologias, as definições propiciam a possibilidade de entendimento do objeto e de compatibilização semântica, pois descrevem o conteúdo significativo de um termo. Esta descrição, quando formalizada, possibilita que agentes inteligentes possam entender o significado de um termo e estabelecer inferências sobre esses significados, pois a definição é composta de características do objeto que está sendo definido. Entretanto, as ontologias existentes ressentem-se hoje de um padrão de seu enunciado definitório para a sua elaboração (Michael, Mejino, & Rosse, 2001; Smith & Kumar, 2004; Smith et al., 2007). Entendemos aqui um enunciado definitório como o conjunto de características que descrevem um objeto em um dado contexto e que têm por função, além de permitir o posicionamento desse objeto em um dado domínio, possibilitar o entendimento de toda sua complexidade, ou seja, que elementos ônticos (relações funcionais e partitivas) possam ser agregados aos elementos lógicos (relações de gênero-espécie) (Dahlberg, 1983).

Outra questão que se coloca é que a falta de padrão para as definições em ontologias é bastante problemática quando se coloca a questão da compatibilização de linguagens que operam em bases cooperativas em sistemas automatizados. Bouaud, Bachimont, Charlet, e Zweigenbaum (1994) discutem princípios para a aquisição e a estruturação de ontologias que, em certa medida, abordam a importância de definições em Ontologias de Domínio a partir de questões que problematizam como é possível representar conhecimento para que um mecanismo automatizado possa “compreender” significados. Seus argumentos remontam à própria Inteligência Artificial.

De modo geral, a Inteligência Artificial procura soluções usando o conhecimento que, em geral, é expresso em linguagem natural. Como não existe qualquer expressão operacional desse conhecimento, isto é, o domínio não é formal, temos que lidar com a riqueza semântica da língua e suas múltiplas interpretações potenciais. O problema básico é que o conhecimento humano para o qual queremos prover um *software* não é formal. O conhecimento que as pessoas têm sobre os objetos em um domínio

é principalmente *descritivo*. As pessoas raciocinam com este conhecimento descritivo empregando as leis semânticas do sentido, como analogia, metáfora, etc. Em domínio não formal, não temos definições completas para as categorias de objetos considerados no conhecimento descritivo. Para que um computador lide com esta espécie de conhecimento, precisamos padronizá-lo: precisamos estar de acordo com as noções que ele mobiliza e com seu uso. O conhecimento precisa ser considerado como uma noção objetiva, isto é, compartilhada. Basicamente, temos que fixar o significado dos termos para que todos tenham a mesma compreensão quando usarem um termo. O conhecimento será descrito adequadamente se seu objetivo e seus aspectos forem explicitados. O uso de termos, conseqüentemente, não se apoiará em todo o potencial da interpretação semântica, mas na manipulação de suas descrições explícitas. No entanto, as restrições postas ao conhecimento pelo processo de padronização não devem modificar o significado das noções representadas (Bouaud et al., 1994).

Outros trabalhos que ressaltam a importância de estabelecimento de definições coerentes para a elaboração de ontologias vêm sendo apresentados na literatura (Michael et al., 2001; Kohler, Munn, Ruegg, Skusa, & Smith, 2006; Hovy, 2002; Smith & Rosse, 2004). Nesses trabalhos, questões relacionadas à importância da descrição dos aspectos dos objetos, ou seja, suas características e propriedades, vêm sendo discutidas visando permitir o raciocínio automático, fundamental para que as ontologias possam possibilitar inferência e descoberta de conhecimento. Os estudos têm mostrado que as definições em ontologias devem funcionar como um sistema integrado em que cada elemento da definição deve estar relacionado como em um sistema.

Michael et al. (2001) registram a diferença entre a produção de definições para atender às necessidades de dicionários e léxicos em contraposição àquelas que pretendem atender às ontologias. Argumentam que nos dicionários a unidade de informação é uma palavra e que, nesse sentido, as definições especificam os significados das palavras. As palavras são apresentadas de forma alfabética e as definições não possuem nenhuma relação com as palavras vizinhas. Essa forma de definir atende às necessidades dos dicionários, em que a definição é realizada de forma isolada, sem qualquer relação com as outras palavras. Ao contrário dos dicionários, uma definição no âmbito de uma ontologia pode ser considerada incompleta se todos os conceitos pais (conceitos superordenados) não tiverem sido também definidos. Este aspecto é um dos fatores que permite a consistência da estrutura lógica, muito importante para possibilitar o raciocínio automático.

Em ontologias, diferentemente de outras linguagens que também têm por função permitir a representação de informação em bases compartilhadas, como Tesouros, por exemplo, é necessário que se possa produzir o raciocínio inteligente através de mecanismos de inferência que dão suporte à manipulação do conhecimento a ser explicitado. Dessa forma, considera-se que a definição em ontologias necessita ser formalizada para permitir a interpretação por agentes inteligentes e não só por humanos. Ela necessita de um padrão para seus enunciados definitórios e de uma linguagem lógica que possa apoiar o entendimento dos conteúdos conceituais dos termos por esses agentes inteligentes.

CONCLUSÕES

A partir do estabelecimento de critérios que visaram identificar a eficácia e a eficiência dos procedimentos automáticos frente aos procedimentos manuais, foi possível chegar a um resultado que ressaltou pontos negativos e positivos nos dois procedimentos. Essa forma de análise permitiu levantar questões que possibilitem uma tomada de decisões quanto ao procedimento de levantamento de *corpus*

para elaboração de Ontologias de Domínios em contextos específicos. Nesse sentido, por exemplo, se por um lado o levantamento automático possibilita uma agilidade na identificação das entidades que estão em relação, ou seja, os elementos antecessores e sucessores, por outro, a análise manual consegue identificar com mais precisão os sentidos que, por vezes, não estão evidenciados através de uma análise sintática, dependem de uma análise semântica, que possibilita identificar os diversos significados das expressões verbais.

A partir das abordagens apresentadas, consideramos que, para a elaboração de Ontologias de Domínios, é fundamental analisar todos os fatores intervenientes na formação do *corpus* terminológico. A formação deste *corpus* é que permitirá consistência de expressividade no momento de estabelecer enunciados de conhecimento sobre o domínio. As estratégias metodológicas utilizadas para se chegar a estes enunciados dependem das bases teóricas e metodológicas que serão adotadas. Pode-se beneficiar-se dos recursos computacionais e de processamento de linguagem natural para automatizar tarefas e ampliar a cobertura de insumos a serem analisados, mas é fundamental a curadoria e a análise hermenêutica para que os resultados sejam satisfatórios. Esta investigação possibilitou também que trabalhos futuros venham a ser realizados visando à identificação de padrões definitórios em domínios específicos.

REFERÊNCIAS BIBLIOGRÁFICAS

- Appolinário, F. (2006). *Metodologia da Ciência, Filosofia e Prática da Pesquisa*. São Paulo, Brasil: Pioneira Thomson Learning.
- Bick, E. (2000). *The Parsing System Palavras: Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework* (Doctoral thesis, University Of Aarhus). Retrieved from <http://visl.sdu.dk/~eckhard/pdf/PLP20-amilo.ps.pdf>
- Bouaud, J., Bachimont, B., Charlet, J., & Zweigenbaum, P. (1994). Acquisition and Structuring of an Ontology within Conceptual Graphs. In *Proceedings of ICCS Workshop on Knowledge Acquisition using Conceptual Graph Theory*, College Park: MD, 1994, (pp. 1-25). Retrieved from <https://pdfs.semanticscholar.org/ce80/a41aa2e3d8139b5cbcd3de722d42a6c13f81.pdf>
- Campos, M. L. A. (2010). O papel das definições na pesquisa em ontologia. *Perspectivas em Ciência da Informação*, 15 (1), 220-238.
- Colino, C. (2002). El metodo comparativo. In R. Reyes (Dir.), *Diccionario Crítico de Ciencias Sociales*. Madrid, España: Universidad Complutense.
- Dahlberg, I. (1981). Conceptual definitions for interconcept. *International Classification*, 8 (1), 16-22.
- Dahlberg, I. (1978a). *Ontical structures and universal classification*. Bangalore: Sarada Ranganathan Endowment for Library Science.
- Dahlberg, I. (1978b). A Referent-oriented analytical concept theory of interconcept. *International Classification*, 5 (3), 142-150.
- Dahlberg, I. (1983). Terminological definitions: characteristics and demands. In *Problèmes de la définition et de la synonymie en terminologie* (pp. 13-51). Québec, Canada: GIRSTERM.

- Ding, Y., & Foo, S. (2002). Ontology research and development Part 2: a review of ontology generation. *Journal of Information Science*, 28 (5), 375-388.
- Fernandez, M., Gomez-Perz, A., & Juristo, N. (1997). Methontology: From ontological art towards ontological engineering. In *Proceedings of the AAAI Spring Symposium on Ontological Engineering* (pp. 33-40). Menlo Park, CA: The AAAI Press.
- Fernandez-López, M. (2001). Overview of methodologies for building ontologies. In *IJCAI Workshop on Ontologies and Problem-Solving Methods: Lessons Learned and Future Trends, 1999* (pp. 26-34).
- Gangemi, A., Pisanelli, D. M., & Steve, G. (1999). An Overview of the ONIONS Project: Applying Ontologies to the Integration of Medical Terminologies. *Data Knowledge Engineering*, 31 (2), 183-220.
- Gil, A. C. (1995). *Métodos e técnicas de pesquisa social*. (4^a ed.). São Paulo: Atlas.
- Gomes, H. E., Campos, M. L. A., & Guimarães, L. S. (2010). Organização da informação e terminologia: abordagem onomasiológica. *DataGramZero Revista Ciência da Informação*, 11 (5).
- Gruninger, M., & Fox, M. S. (1995). Methodology for the Design and Evaluation of Ontologies. In *Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing at the 14th International Joint Conference on Artificial Intelligence*, Montreal, Canada, 1995.
- Guarino, N., & Welty, C. (2002). Evaluating Ontological Decisions with OntoClean. *Communications of the ACM*, 45 (2), 61-65.
- Hovy, E. (2002). Comparing Sets of Semantic Relations in Ontologies. In R. Green, C. A. Bean, & S. H. Myaeng (Eds.), *Semantics of Relationships: An Interdisciplinary Perspective* (pp. 91-110). [S.l.]: Kluwer.
- Kohler, J., Munn, K., Ruegg, A., Skusa, A., & Smith, B. (2006). Quality control for terms and definitions in ontologies and taxonomies. *BMC Bioinformatics*, 7 (212).
- Lakatos, E. M., & Marconi, M. A. (2007). *Fundamentos de metodologia científica*. São Paulo, Brasil: Atlas.
- Lima, W. F. (2007). A Lexicologia Construcionista: uma proposta alternativa de estudo do léxico na linguagem em uso. In A. M. Isquierdo & I. M. ALVES (Orgs.). *As Ciências do léxico: lexicologia, lexicografia, terminologia* (pp. 125-136). São Paulo, Brasil: Humanites.
- Metaxiotis, K., Psarras, J. E., & Askounis, D. (2001). Building ontologies for production scheduling systems: towards a unified methodology. *Information management & computer security*, 9 (1), 44-50.
- Michael, J., Mejino, J. L. V., & Rosse, C. (2001). The role of definitions in biomedical concept representation. *JAMIA Symposium Supplement*, 1, 463-467.
- Mizoguchi, R. (2014). What Ontological Engineering Can Do for Solving Real-World Problems. *Knowledge and Systems Engineering Advances in Intelligent Systems and Computing*, 244.
- Paslaru-Bontas, E. (2007). *A Contextual Approach to Ontology Reuse: Methodology, Methods and Tools for the Semantic Web* (Master's Dissertation, Fachbereich Mathematik und Informatik, Freien Universität at Berlin, Berlin, Germany).

Riggs, F. W. (1989a). Information science and social science: the need for onomantics. *International Forum for Information and Documentation*, 14 (1), 12-21.

Riggs, F. W. (1989b). Terminology and lexicography: their complementarity. *Terminology and lexicography: their complementarity*, 2 (2), 90-110.

Riggs, F. W. (1979). Terminology for the Social Sciences. In *Theoretical and methodological problems or terminology, Proceedings of an international symposium, Moscow, Russia*, (pp. 591-606).

Riggs, F. W. (1996). Onomantics and terminology. *Knowledge Organization*, 23, 25-3.

Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., ... & Mungall, C. J. (2007). The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nature Biotechnology*, 25 (11), 1251-1255.

Smith, B., & Kumar, A. (2004). On Controlled Vocabularies in Bioinformatics: A Case Study in the Gene Ontology. *BIOSILICO: Drug Discovery Today*, 2, 246-252.

Smith, B., & Rosse, C. (2004). The Role of Foundational Relations in the Alignment of Biomedical Ontologies. In M. Fieschi et al. (Eds.), *Medinfo* (pp. 444-448). Amsterdam, Netherlands: IOS Press.

Sure, Y., Staab, A., & Studer, R. (2002). Methodology for development and employment of ontology based knowledge management applications. *SIGMOD Record*, 31 (4), 18-23.

Uschould, M., & King, M. (1995). Towards a Methodology for Building Ontologies. In *Proceedings of the International Joint Conference on Artificial Intelligence, Workshop on Basic Ontological Issues in Knowledge*.



REQUISITOS PARA FERRAMENTAS DE VISUALIZAÇÃO DE ONTOLOGIAS

Cristiane Mendes Netto¹, Gercina Angela Borém de Oliveira Lima²

¹Universidade Vale do Rio Doce, 0000-0002-5725-8323, cris.netto@gmail.com

²Universidade Federal de Minas Gerais, 0000-0003-0735-3856, limagercina@gmail.com

RESUMO Com o objetivo de auxiliar o uso de ontologias, no que se refere a ampliar a compreensão humana do domínio representado por elas, a visualização de ontologias tem sido um tema presente na literatura científica. Com o objetivo de analisar as propostas de ferramentas de visualização de ontologias, apresenta-se, neste artigo, o resultado de um estudo de avaliação de características de quatorze ferramentas relatadas na literatura, bem como testes de uso com duas dessas ferramentas. A metodologia para seleção das referências foi desenvolvida com aplicação de etapas do processo Proknow-C (KnowledgeDevelopmentProcess – Construtivist), que possibilitou a organização de um conjunto de vinte referências científicas, publicadas de 2007 a 2016. A contribuição do trabalho consiste na sistematização de requisitos funcionais e de usabilidade para ferramentas de visualização de ontologias. Nos requisitos funcionais básicos, foi recomendada a visualização de classes, instâncias, relações e propriedades, bem como as funções de ampliar e reduzir visões, pesquisar, filtrar, exportar imagem. Como requisitos inovadores, sugeriu-se: anexar anotações, compartilhar, enviar notificações e salvar preferências dos usuários. Nos requisitos de usabilidade, propõe-se que se tenha: consistência, atalhos, retroalimentação, diálogo, prevenção de erros e personalização. Observou-se no estudo realizado que a visualização da informação aplicada às ontologias tem potencial para oferecer aos usuários um suporte para auxiliar a recuperação de informação. Como trabalhos futuros a serem desenvolvidos, sugerem-se estudos com usuários para identificação de novos requisitos, a partir dos interesses e necessidades desses, bem como uma análise de tarefas, em um processo centrado no usuário. Espera-se, com este trabalho, colaborar para o desenvolvimento de ferramentas de visualização de ontologia de boa usabilidade que proporcionem aos usuários alcançarem seus objetivos.

PALAVRAS-CHAVE *Ontologia, Organização do Conhecimento, Visualização de Ontologias.*

ABSTRACT The ontology visualization has been a present theme in the scientific literature, in order to increase the human comprehension of the domain represented by the ontologies. This work presents the results of an evaluation study of the characteristics of fourteen tools reported in the scientific literature, as well as the tests of two of these tools. The objective of this study is to analyze the proposed ontology visualization tools and the presented requirements for the ontology visualization tools to attend the user needs. The methodology used to select the references was developed using the steps of the Proknow-C process (Knowledge Development Process - Constructivist) that allowed the organization of a set of twenty scientific references, published from 2007 to 2016. The contribution of the work consists in the systematization of functional and usability requirements for the ontology visualization tools. For the basic functional requirements there were recommended the classes visualization, instances, relations and properties, such as the functions of zooming and reducing views, researching, filtering, and exporting image. As innovative requirements it was suggested: to attach annotations, share, send notifications and save the user preferences. In the requirements of usability, it is proposed to have: consistency, shortcuts, feedback, dialogue, error prevention and personalization. It was observed in the study that the information visualization applied to the ontologies has the potential to offer to the users a support to help with the information

retrieval. As future work it is suggested to be developed studies with the users to identify new requirements based on their interests and needs, as well as a task analysis in a user-centered process. It is hoped that this work will collaborate to the development of the ontology visualization tools of good usability and that it will allow the users to reach their objectives.

KEYWORDS *Ontology, Knowledge Organization, Ontology Visualization*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

No campo da organização da informação e do conhecimento, os estudos de ontologias têm tratado principalmente das suas aplicações como Sistemas de Organização do Conhecimento para criação de bases de conhecimento. Pode-se considerar que o interesse comum desses estudos reside no potencial que as ontologias possuem para fornecerem uma conceituação compartilhada de um domínio para ser comunicada a pessoas e sistemas computacionais.

No âmbito computacional, a criação de ontologias tem sido beneficiada pelo desenvolvimento de linguagens formais e ferramentas, que oferecem os recursos formais para profissionais especialistas em ontologias. No entanto, dada a complexidade dos domínios a serem representados, as ontologias comumente podem conter centenas e até milhares de classes e instâncias. Essa estrutura pode se tornar de difícil visualização e extremamente complexa para a compreensão humana, principalmente por usuários não especialistas em ontologias.

Com vistas a investigar essa questão, este trabalho buscou analisar as aplicações da área de Visualização de Informação (VI) para ontologias. A VI é definida por Card, Mackinlay, e Shneiderman (1999) como o uso de representações de dados abstratos suportadas por computador e interativas para ampliar a cognição. Para Freitas, Chubachi, Luzzardi, e Cava (2001), a VI é uma área de aplicação de técnicas da computação gráfica, geralmente interativas, visando auxiliar o processo de análise e compreensão humana de um conjunto de dados, por manipulações das representações gráficas.

Observou-se que estudos dessa área têm sido aplicados em ferramenta de visualização de ontologias com o propósito de auxiliar o uso de ontologias e a compreensão humana do domínio representado por elas. Visando analisar as características dessas ferramentas, realizou-se uma revisão sistematizada da literatura científica, em trabalhos publicados no período de 2007 a 2016. A metodologia utilizada para essa revisão foi baseada em etapas do processo ProcKnow-C (Knowledge Development Process – Construtivist), conforme proposto por Tasca, L. Ensslin, Ensslin, e Alves (2010).

Como resultado, este trabalho descreve um recorte advindo dessa revisão de literatura, destacando a análise de ferramentas para visualização de ontologias, identificadas em referências da literatura científica. Por meio de um estudo comparativo das características de interface e recursos dessas ferramentas, apresentam-se doze requisitos funcionais e seis requisitos de usabilidade considerados essenciais para ferramentas de visualização de ontologia. A contribuição deste trabalho se alicerça no delineamento deixado às futuras pesquisas e no desenvolvimento de ferramentas de visualização de ontologias.

A organização do artigo apresenta-se da seguinte forma: a Seção 2 descreve fundamentos sobre a visualização de ontologias. A seção 3 a metodologia da pesquisa. Em seguida, são abordados, nas respectivas seções, os resultados e as conclusões.

A VISUALIZAÇÃO DE ONTOLOGIAS

A visualização de informação aplicada à ontologia consiste em representar graficamente os elementos que constituem as ontologias: classes, relações, axiomas e instâncias, com possibilidades de realização de tarefas de interação sobre esses elementos. Observa-se que nos estudos de visualização de ontologias, encontrou-se que os termos modelos e técnicas estão sendo usados para designar o as características do tipo de visualização aplicada a ontologia. Com base nas características predominantes, oferecidas pelas ferramentas de visualização de ontologias, Katifori e Halatsis (2007) agruparam os métodos de visualização de ontologias em seis categorias, conforme apresentado a seguir.

- a) *Lista indentada*: uma lista indentada apresenta a taxonomia da ontologia em uma estrutura semelhante à organização de pastas do Windows Explorer.
- b) *Árvore de links e nós*: na árvore de links e nós a ontologia é representada como um conjunto de nós interconectados (grafos) que geralmente podem ser expandidos e retraídos pelo usuário.
- c) *Visualização Zoomable*: uma visualização baseada em aproximação (*zoomable*) apresenta os nós filhos aninhados dentro de seu nó pai e permite ao usuário fazer ampliação para e sair dos nós.
- d) *Preenchimento de espaço*: a vista de preenchimento de espaço apresenta os nós subdividindo o espaço da tela. O tamanho de cada subdivisão depende dos atributos do nó, como o número de nós secundários.
- e) *Foco e Contexto*: essa visão apresenta contexto e foco ao mesmo tempo exibindo o nó em foco ampliado e o resto dos nós colocados em torno dele.
- f) *Paisagem 3D*: a visão da paisagem da informação apresenta objetos 3D codificados por cor e tamanho em um plano usando a metáfora da paisagem.

Existem vários exemplos de ferramentas descritas na literatura que permitem a visualização com características dessas categorias de classificação. Cabe destacar que uma ferramenta indicada como exemplo de um método pode também conter características de métodos distintos. Mediante a diversidade de métodos e ferramentas que se apresentam para visualização de ontologias, tem-se que a escolha de uma solução é um desafio para os usuários, assim como é complexa a definição de requisitos para o desenvolvimento de novas soluções. Logo, avaliar as principais características do domínio e dos métodos é necessário para se propor uma visualização de informação mais adequada. Além disso, tem-se que alguns elementos que compõem as ontologias geralmente são priorizados em detrimento de outros, conforme a especificidade de ferramenta e método de visualização.

METODOLOGIA

No que se refere ao objetivo deste trabalho, esta pesquisa se caracteriza por ser exploratória e descritiva. Segundo Gil (2002), a primeira característica se apresenta em pesquisas que têm como objetivo proporcionar maior familiaridade com um problema, com vistas a torná-lo mais explícito, envolvendo

o aprimoramento de ideias ou a descoberta de intuições. A pesquisa descritiva visa à descrição das características de determinada população ou fenômenos, ou ao estabelecimento de relações entre variáveis.

Quanto ao delineamento da pesquisa, no que se refere ao planejamento em sua dimensão mais ampla, ela classifica-se como pesquisa bibliográfica, desenvolvida com base em material já elaborado, constituída de trabalhos científicos (Gil, 2002). A abordagem da análise dos dados baseia-se na pesquisa indutiva, em que, partindo-se de dados particulares constatados, infere-se uma verdade geral, não contida nas partes examinadas (Lakatos & Marconi, 2010).

A natureza da pesquisa é aplicada, tipo que se caracteriza, segundo Lakatos e Marconi (2003), por seu interesse prático, isto é, que os resultados sejam aplicados, imediatamente, na solução de problemas que ocorrem na realidade.

Como instrumento para sistematização da revisão de literatura, utilizou-se o processo Procknow-C, proposto por Tasca et al. (2010). A identificação desse processo ocorreu a partir de pesquisas em periódicos nacionais da área de Ciência da Informação, nos quais encontrou-se a aplicação do processo nos trabalhos de L. Ensslin, Ensslin, e Pacheco (2012), L. Ensslin, Waiczuk, Chaves, e Ensslin (2015) e Machado, Ensslin, William, e L. Ensslin (2016).

No contexto do Procknow-C, as referências selecionadas na literatura constituem um conjunto de publicações com reconhecimento científico, identificadas conforme o interesse de pesquisa e delimitações estabelecidas pelo pesquisador. Neste trabalho, o portfólio foi selecionado por pesquisas no mês de janeiro de 2017, nas bases de dados *Information Science & Technology Abstracts* (ISTA); *Library, Information Science & Technology Abstracts* (LISTA); *Library and Information Science Abstracts* (LISA); *Web of Science* e SCOPUS, disponíveis para acesso pelo Portal Capes (<http://www.periodicos.capes.gov.br>).

O recorte temporário para seleção das referências consistiu de publicações entre 2007 a 2017 e, como critério de delimitação, consideraram-se a publicação em periódicos científicos, livros e anais de conferências científicas. A *string* de busca escolhida para condução da recuperação foi (ontolog* visuali*) OR (ontolog* navigat*). Os caracteres curingas (*) foram usados para cobrir variações de plural e dos termos de pesquisa. No caso do termo referente à visualização, em inglês, a grafia pode ser tanto *visualization* quanto *visualisation*.

A partir da análise manual das referências retornadas pela pesquisa, identificou-se um conjunto de vinte artigos que tratam de apresentações de ferramentas de visualização de ontologias. Esses trabalhos foram analisados e, a partir deles, foram selecionadas duas ferramentas para testes de uso com uma ontologia e avaliação da interface e dos recursos. Os resultados dessa avaliação são relatados na seção a seguir.

RESULTADOS

O conjunto de artigos que fundamentou a realização deste estudo consistiu de vinte referências, sendo 25% dos trabalhos advindos de periódicos e outros 75% de conferências científicas. O Quadro 1 apresenta o conjunto das quatorze ferramentas de visualização de ontologias identificadas, junto às suas características e referências. Algumas ferramentas são relatadas em mais de uma referência, como foi

o caso das ferramentas VOWL, Knoocks e OWLGrEd, que apresentam detalhamento distinto em diferentes publicações.

Quadro 1. Ferramentas de visualização de ontologias relatadas na literatura

Ferramenta	Características da visualização	Referências
VOWL	Baseada em grafos. A visualização é feita por uma representação visual projetada para mapear os elementos da OWL. Criada para usuários especialistas e não especialistas em ontologias.	Lohmann, Negru, Haag, e Ertl (2016) Bach, Pietriga, e Ertl (2014) Lohmann, Negru, e Bold (2014) Negru e Lohmann (2013)
Knoocks	A visualização é dedicada em representar as instâncias relacionadas à estrutura nas ontologias.	Kriglstein e Wallner (2010) Kriglstein e Wallner (2011) Kriglstein (2009)
OntoSumViz	Desenvolvida com base no uso de vetores descritivos para gerar uma visualização que represente uma visão geral de ontologia. Proposta para auxiliar a avaliação e a seleção de ontologias para um domínio.	Machova, Vrana, Mach, e Sincak (2016)
MemoGraph	Visualização desenvolvida visando ao uso universal e voltada especificamente para pacientes com Alzheimer. Adota uma representação com imagens que podem ser escolhidas por usuários ou pelo sistema.	Ghorbel et al.(2016)
OWLGrid	Visualização gerada com uma notação baseada em diagrama de classes em UML.	Cerans, Liepins, Sprogis, Ovcinnikova, e Barzdins (2015)
Osmosys	Visualização com diferentes formatos para representação dos conteúdos de ontologias, com acesso via web.	Psyllidis (2015)
UML-like	Visualização para uso no Protégé, baseado na representação da linguagem UML.	Cai, Shi, e Yang (2015)
OLSViz	Desenvolvida especificamente para visualização das ontologias do repositório <i>Ontology Lookup Service (OLS)</i> .	Vercruysse, Venkatesan, e Kuiper (2012)
KC-Viz	A visualização é gerada a partir de um método para extração de conceitos-chaves que são usados para representar as ontologias de forma concisa.	Motta et al. (2011)
OntoTrix	Ferramenta projetada para permitir aos usuários visualizar e navegar em conjuntos de instâncias e suas relações em ontologias.	Bach, Pietriga, Liccardi, e Legostaev (2011).
SSN-XG	Visualização baseada em representação de diagramas conceituais.	Howse, Stapleton, Taylor, e Chapman (2011)
OWLGrEd	Ferramenta para edição e visualização de ontologias, com uma representação gráfica baseada na linguagem UML.	Bārzdiņš, Bārzdiņš, Čerāns, Liepiņš, e Sproģis (2010a) Bārzdiņš, Bārzdiņš, Čerāns, Liepiņš, e Sproģis (2010b).
OWLEasyViz	Editor de ontologia que combina uma representação textual e gráfica de ontologias com uma proposta de interface para atender a usuários especialistas e não especialistas em ontologias.	Catenazzi, Sommaruga, e Mazza (2019)
GrOWL	Visualização baseada em grafos para uso no Protégé.	Krivov, Williams, e Villa (2007)

Referência: autoras (2017).

Percebeu-se que grande parte das ferramentas relatadas na literatura não estão disponíveis para uso e que podem ter sido descontinuadas no seu desenvolvimento, como é o caso do Knoocks e OntoTrix. Além disso, observa-se que as propostas relatadas na literatura ainda carecem de melhores avaliações através de testes com usuários e conhecimento mais aprofundado sobre as suas necessidades.

Considera-se que, para cada grupo de usuários e domínio, pode haver uma necessidade distinta, tornando mais complexa, portanto, uma ferramenta que seja ideal para visualização de ontologias em todos os contextos.

Com o objetivo de conhecer e avaliar as funções das ferramentas de visualização de ontologias, escolheu-se testar as ferramentas VOWL e Knoocks, por terem maior número de publicações científicas. Para uso da ferramenta Knoocks, foi realizado um contato com os autores da ferramenta, para obter o programa para instalação, tendo em vista que o projeto de seu desenvolvimento foi descontinuado. Já a ferramenta VOWL foi utilizada com a versão disponível na internet. Para os testes, utilizou-se a ontologia *Travel*, encontrada disponível publicamente por meio do *Protégé Ontology Library* (https://protegewiki.stanford.edu/wiki/Protege_Ontology_Library). Esta ontologia foi escolhida por representar um domínio de fácil conhecimento e por ter uma quantidade de classes suportada pelas duas ferramentas.

A ferramenta Knoock foi projetada para visualizar as classes, instâncias e propriedades. A representação das classes é feita com retângulos, chamados de blocos. Cada bloco representa uma classe com suas subclasses e os retângulos à direita são subclasses dos retângulos à esquerda. As instâncias de cada classe são listadas dentro dos blocos de suas respectivas classes. As propriedades dos objetos são representadas pelas linhas e as propriedades de tipos são identificadas por cores. Ao clicar sobre o nome de cada instância, suas propriedades são apresentadas.

A Figura 1 apresenta a interface principal da ferramenta Knoocks. A organização desta ferramenta é feita em quatro áreas principais. Na área superior esquerda (A), encontram-se, organizados em abas, os recursos de filtros que podem ser aplicados para mostrar ou esconder blocos, busca por palavra-chave, histórico e marcações que podem ser aplicadas para identificar itens das visualizações. A área inferior esquerda (B) é utilizada para mostrar uma visão geral da ontologia. Os itens apresentados nessa área podem ser vistos de forma detalhada na terceira e maior parte da interface (C). Na lateral direita (D), são apresentadas as opções de mostrar todas as conexões da ontologia, organizar a visualização de forma circular, esconder detalhes e exportar a visualização central para um arquivo de imagem PNG e ainda esconder ou mostrar detalhes de domínio.

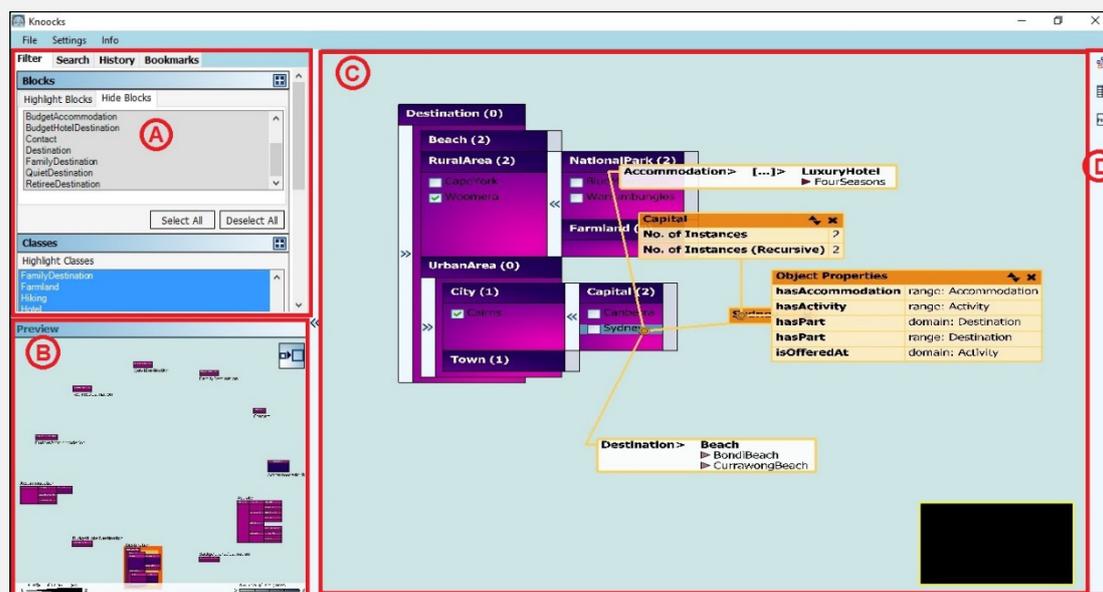


Figura 1. Organização da interface da ferramenta Knoocks

Na análise da ferramenta Knoocks, percebeu-se que a organização da interface e a representação por blocos favorecem o uso por não especialistas em ontologias. Considerou-se positiva a estratégia utilizada na tela inicial do programa para orientar o usuário quanto à representação utilizada. Considerou-se também uma potencialidade os recursos de filtros para o usuário controlar o volume de informação a ser visualizado e a interface, que apresenta ao usuário a definição de preferências no momento de abrir a ontologia. Percebeu-se, pelo tempo já decorrido do desenvolvimento da ferramenta, que sua interface pode ser considerada ultrapassada, vistas as tendências atuais de formatos e cores. No entanto, sua organização e nível de controle oferecido ao usuário podem servir de inspiração para novas ferramentas a serem desenvolvidas.

A ferramenta VOWL foi testada na versão Web, disponível em: <<http://visualdataweb.de/webvowl/>>. A interface desta ferramenta está organizada em três partes, conforme mostra a Figura 2. O modelo de visualização consiste em um grafo que pode ser manipulado com movimentos de clicar e arrastar, conforme interesse de visualização. A visualização do grafo é apresentada na parte central da interface (A). Na parte inferior da interface (B), estão contidas as principais funções da ferramenta. Dentre essas funções, tem-se: a) pesquisar por palavra-chave, com recurso de sugestão de termos pela ferramenta; b) abrir um arquivo OWL de uma ontologia; c) exportar imagem da visualização para os formatos JSON e SVF; d) ajustar a proximidade dos itens de visualização; e) aplicar filtros sobre os conteúdos que são mostrados na visualização; e) personalizar alguns itens da visualização e f) voltar à visualização inicial. O painel lateral direito (c) apresenta a descrição dos itens selecionados da ontologia.

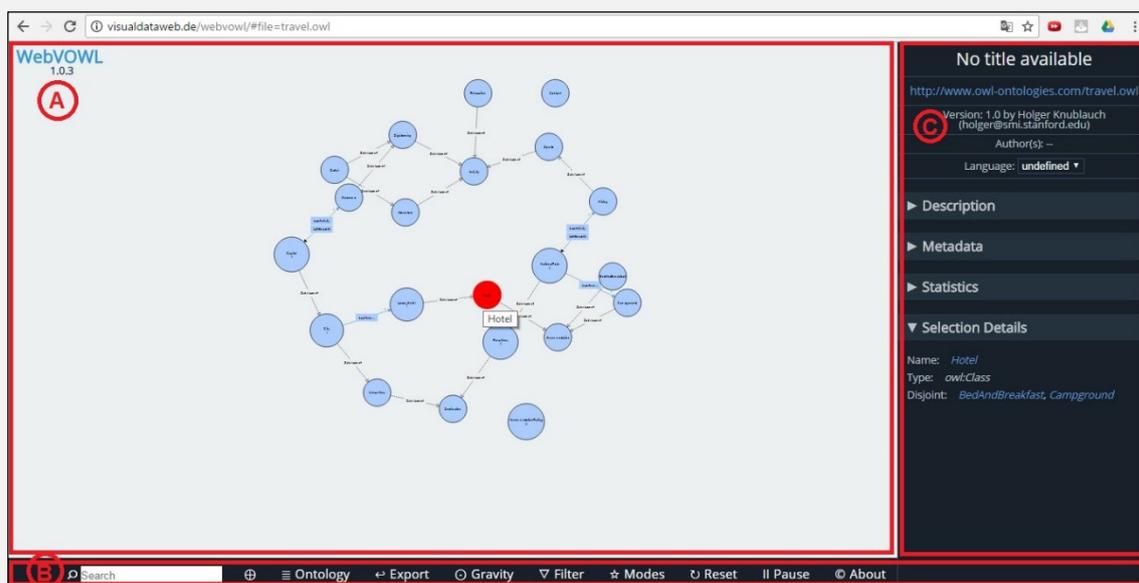


Figura 2. Organização da interface da ferramenta VOWL

Entre as potencialidades consideradas nesta ferramenta, tem-se a sua interface via web, com interação rápida sobre os grafos que representam a ontologia e opções de ajustes para visão geral e detalhamento. A descrição na lateral direita referente ao item selecionado na ontologia favorece o usuário no acesso à informação. O recurso de sugestão de termos para fazer a busca por palavras-chaves permite que o usuário direcione a sua busca com facilidade. Entre as limitações da ferramenta, tem-se a dificuldade de visualizar ontologias maiores, pois a representação fica confusa para o usuário, visto o grande número de nodos no grafo.

A partir dos estudos realizados, apresenta-se a seguir a indicação de requisitos que podem ser considerados essenciais no contexto atual para visualização de ontologias. Quanto aos requisitos funcionais, conforme apresenta Paula Filho (2009), eles descrevem as funções que um produto deve realizar em benefício dos seus usuários. Nesse contexto, a partir de uma análise das ferramentas apresentadas na literatura, elaborou-se o detalhamento de dez requisitos funcionais sugeridos como básicos e inovadores para ferramentas de visualização de ontologias. Os requisitos considerados básicos referem-se aos já encontrados, em geral, nas ferramentas de visualização. Os requisitos inovadores foram sugeridos com base em experiências das autoras com outras ferramentas de gestão do conhecimento.

Requisitos básicos:

- 1) *Visualizar classes*: mostrar uma visão clara e de destaque das classes de uma ontologia é importante para que os usuários identifiquem o conteúdo do domínio representado.
- 2) *Visualizar instâncias*: mostrar as instâncias em evidência pode ser do interesse do usuário, visto que estas são os dados concretos de uma representação ontológica.
- 3) *Visualizar relações*: evidenciar as relações e os seus tipos representados em uma ontologia pode favorecer a compreensão do domínio.
- 4) *Visualizar propriedades*: apresentar o detalhamento das propriedades dos tipos de dados para conhecimento do usuário.
- 5) *Ampliar e reduzir visões*: mostrar uma visão geral e detalhada da ontologia, conforme escolha do usuário, possibilitando inferências da organização semântica da ontologia.
- 6) *Pesquisar*: incluir uma opção de consulta à ontologia por palavra-chave e com recurso de indicação dos termos da ontologia.
- 7) *Filtrar*: gerar visualizações baseadas em filtros, conforme escolha dos usuários, permitindo evidenciar múltiplos aspectos de uma ontologia.
- 8) *Exportar imagens*: Gerar arquivos de imagem da representação gráfica da ontologia.

Requisitos inovadores:

- 1) *Anexar anotações*: dar ao usuário a opção de incluir anotações e marcações que o permita compreender melhor o domínio dentro da sua experiência.
- 2) *Compartilhar*: fornecer opção de trabalho colaborativo para compartilhar o estudo de uma ontologia com outros usuários.
- 3) *Enviar notificações*: permitir que o usuário defina interesses de assunto e fornecer notificações de atualização.
- 4) *Salvar preferências*: permitir que o usuário escolha suas preferências de uso na ferramenta para personalizar o seu uso.

Quanto aos requisitos de usabilidade, que visam aumentar a satisfação dos usuários de uma ferramenta, apresenta-se a lista a seguir, baseada em Agner (2009):

- 1) *Consistência*: Repetir os padrões para facilitar o uso da interface, mantendo o *layout* das cores, tipologia e menus. Fazer com que ações de operações semelhantes sejam repetidas para facilitar o aprendizado dos usuários.
- 2) *Atalhos*: Incluir teclas de atalhos para acesso às funções do sistema, visando aumentar a velocidade de uso para os usuários que se tornarem mais experientes.

- 3) *Retroalimentação*: Comunicar ao usuário as ações executadas, podendo usar metáforas de interface no caso de operações que demandam respostas mais demoradas.
- 4) *Diálogo*: Sinalizar para o usuário, por meio de mensagens, a sequência de tarefas realizadas.
- 5) *Prevenção de erros*: Incluir opção de backup e desfazer ações para permitir a recuperação de um estado, no caso de erros cometidos pelo usuário.
- 6) *Personalização*: Permitir que os usuários ajustem diferentes modos de exibir os recursos da ferramenta, evitando sobrecarga de informações, conforme uso e interesse.

Além dos requisitos, considera-se que, para favorecer o uso multiplataforma da ferramenta, torna-se essencial que a ferramenta seja desenvolvida para uso via Web. Também é importante que a notação a ser utilizada pela ferramenta favoreça a compreensão do usuário. As propostas de notação já sugeridas na literatura, como as propostas por Cai et al. (2015) e também Negru e Lohmann (2013), podem ser um ponto de partida para as implementações.

CONCLUSÕES

Este artigo apresentou uma indicação de requisitos funcionais e de usabilidade para ferramentas de visualização de ontologias. O resultado foi obtido a partir da análise de referências da literatura científica, publicadas entre 2007 e 2016 e testes realizados com duas ferramentas. Nos requisitos funcionais básicos, foi recomendada a visualização de classes, instâncias, relações e propriedades, bem como as funções de ampliar e reduzir visões, pesquisar, filtrar, exportar imagem. Como requisitos inovadores, sugeriu-se: anexar anotações, compartilhar, enviar notificações e salvar preferências dos usuários. Nos requisitos de usabilidade, foi proposto que se tenha: consistência, atalhos, retroalimentação, diálogo, prevenção de erros e personalização.

Observou-se no estudo realizado que a visualização de informação aplicada às ontologias tem potencial para oferecer aos usuários um suporte para auxiliar a recuperação de informação. Além disso, os recursos visuais podem contribuir para melhorar a realização de tarefas, ampliando a compreensão humana do conhecimento no domínio representado pelas ontologias, bem como a recuperação da informação. Acredita-se que a representação visual favorece a capacidade humana para processamento da informação e reduz o tempo de busca, visto que possibilita o uso da inferência perceptiva e a manipulação dos conteúdos com visualização de detalhes e organizações sob demanda.

Como trabalhos futuros, sugerem-se estudos com usuários para conhecimento aprofundado de suas necessidades e especificação de novos requisitos. Faz-se necessária, também, uma análise de contexto de uso, em um projeto centrado no usuário, com vistas à usabilidade, possibilitando melhor qualidade e atendimento dos objetivos. A partir disso, acredita-se que o desenvolvimento de uma ferramenta para visualização de ontologia, com avaliações de usabilidade sendo parte do processo, poderá proporcionar uma solução bem-sucedida para auxiliar os usuários no uso de ontologias.

REFERÊNCIAS BIBLIOGRÁFICAS

- Agner, L. (2009). *Ergodesign e arquitetura de informação: trabalhando com o usuário*. (2. ed.). Rio de Janeiro, RJ, Brasil: Quartet.
- Bach, B., Pietriga, E., Liccardi, I., & Legostaev, G. (2011). OntoTrix: a hybrid visualization for populated ontologies. *Proceedings of the international conference companion on World wide web 20*, 177-180. doi: 10.1145/1963192.1963283
- Bārzdīņš, J., Bārzdīņš, G., Čerāns, K., Liepiņš, R., & Sproģis, A. (2010a). OWLGrEd: a UML Style Graphical Notation and Editor for OWL 2. In Clark, K., Sirin, E. (Eds.), *Proceedings of the 7th International Workshop OWL: Experience and Directions*. Retrieved from http://www.webont.org/owled/2010/papers/owled2010_submission_5.pdf
- Bārzdīņš, J., Bārzdīņš, G., Čerāns, K., Liepiņš, R., & Sproģis, A. (2010b). UML style graphical notation and editor for OWL 2. In P. Forbrig, & H. Günther (Eds.), *Perspectives in Business Informatics Research, 9th International Conference on Business Informatics Research* (pp. 102-114). doi: 10.1007/978-3-642-16101-8
- Cai, Z., Shi, K., & Yang, H. (2015). A Novel Visualization for Ontologies of Semantic Web Representation. In *International Conference on Computational Intelligence and Communication Networks* (pp. 1371-1374). doi: 10.1109/CICN.2015.266
- Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in information visualization: using vision to think*. San Francisco, CA: Morgan Kaufmann.
- Catenazzi, N., Sommaruga, L., & Mazza, R. (2009). User-friendly ontology editing and visualization tools: the OWLeasyViz approach. *Proceedings of the International Conference on Information Visualisation 13*, 283-288. doi: 10.1109/IV.2009.34
- Cerans, K., Liepins, R., Sproģis, A., Ovcinnikova, J., & Barzdins, G. (2015). Domain-Specific OWL Ontology Visualization with OWLGrEd. In Simperl E. et al. (Eds.), *Lecture Notes in Computer Science: Vol. 7540. Extended Semantic Web Conference* (pp. 419-424). doi: 10.1007/978-3-662-46641-4_38
- Ensslin, L., Ensslin, S. R., & Pacheco, G. C. (2012). Um estudo sobre segurança em estádios de futebol baseado na análise da literatura internacional. *Perspectivas em Ciências da Informação*, 17(2), 71-91.
- Ensslin, L., Waiczysk, C., Chaves, L. C., & Ensslin, E. R. (2015). Processo para evidencição do estado da arte aplicado ao tema gestão de produção científica. *Transinformação*, 27(3), 219-228.
- Freitas, C. S., Chubachi, O. M., Luzzardi, P. G., & Cava, R. A. (2001). Introdução à Visualização de Informações. *Revista de Informática Teórica e Aplicada*, 8(2), 143-158.
- Ghorbel, F., Nebrasse, E., Elisabeth, M., Fayçal, H., Faiez, G., & Noura, H. (2016). MEMO GRAPH: An Ontology Visualization Tool for Everyone. *Procedia Computer Science*, 96, 265-274.
- Gil, A. C. (2002). *Como elaborar projetos de pesquisa*. (4. ed.) São Paulo, SP, Brasil: Atlas.

- Howse J., Stapleton, G., Taylor, K., & Chapman, P. (2011). Visualizing Ontologies: A Case Study. In Aroyo, L. et al. (Eds.), *Lecture Notes in Computer Science: Vol 7031. The Semantic Web: International Semantic Web Conference* (pp. 257-272). doi: 10.1007/978-3-642-25073-6_17
- Katifori, A., & Halatsis, C. (2007). Ontology visualization methods: a survey. In *ACM Computing Surveys*, 39(4), 10.
- Kriglstein, S. (2009). User requirements analysis on ontology visualization. *Proceedings of the International Conference on Complex, Intelligent and Software Intensive Systems*, 694-699. doi: 10.1109/CISIS.2009.37
- Kriglstein, S., & Wallner, G. (2011). Development process and evaluation of the ontology visualization tool knoocks-a case study. *Proceedings of the International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications 1*, 187-197. doi: 10.5220/0003324101870197
- Kriglstein, S., & Wallner, G. (2010). Knoocks: A visualization approach for owl lite ontologies. In *International Conference on Complex, Intelligent and Software Intensive Systems* (pp. 950-955). doi: 10.1109/CISIS.2010.55
- Krivov, S., Williams, R., & Villa, F. (2007). GrOWL: A tool for visualization and editing of OWL ontologies. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(2), 54-57.
- Lakatos, E. M., & Marconi, M. A. (2003). *Fundamentos de metodologia científica*. (5. ed.). São Paulo, SP, Brasil: Atlas.
- Lakatos, E. M., & Marconi, M. A. (2010). *Fundamentos da metodologia científica*. São Paulo, SP, Brasil: Atlas.
- Lohmann, S., Negru, S., Haag, F., & Ertl, T. (2016). Visualizing ontologies with VOWL. *Semantic Web*, 7(4), 399-419. Retrieved from <http://www.semantic-web-journal.net/content/visualizing-ontologies-vowl>
- Lohmann, S., Negru, S., Haag, F., & Ertl, T. (2014a). VOWL 2: User-Oriented Visualization of Ontologies. In Janowicz K. et al (Eds.), *Lecture Notes in Computer Science: Vol. 8876. Knowledge Engineering and Knowledge Management* (pp. 266-281). doi: 10.1007/978-3-319-13704-9_21
- Lohmann, S., Negru, S., & Bold, D. (2014b). The ProtégéVOWL plugin: ontology visualization for everyone. In Presutti V. et al (Eds.), *Lecture Notes in Computer Science, Vol. 8798. Proceedings of the European Semantic Web Conference* (pp. 395-400). doi: 10.1007/978-3-319-11955-7_55
- Machado, K. C., Ensslin, S. R., William, V., & Ensslin, L. (2016). Avaliação de Desempenho na Gestão da Informação Digital: contribuições da literatura científica. *Informação & Sociedade*, 26(2), 155-172.
- Machova, K., Vrana, J., Mach, M., & Sincak, P. (2016). Ontology Evaluation Based on the Visualization Methods, Context and Summaries. *Acta Polytechnica Hungarica*, 13(4), 53-76.
- Marconi, M. A., & Lakatos, E. M. (2003). *Fundamentos de metodologia científica*. (5. ed.) São Paulo, SP, Brasil: Atlas.

Motta, E., Mulholland, P., Peroni, S., d'Aquin, M., Gomez-Perez, J. M., Mendez, V., & Zablith, F. (2011). A novel approach to visualizing and navigating ontologies. In Aroyo L. et al. (Eds.), *Lecture Notes in Computer Science: Vol 7031. Proceedings of the International Semantic Web Conference* (pp. 470-486). doi: 10.1007/978-3-642-25073-6_30

Negru, S., & Lohmann, S. (2013). A Visual Notation for the Integrated Representation of OWL Ontologies. *Proceedings of the International Conference on Web Information Systems and Technologies 9*, 308-315. Retrieved from <http://vowl.visualdataweb.org/v1/VOWL2013.pdf>

Paula Filho, W. P. (2009). *Engenharia de Software: Fundamentos, Métodos e Padrões*. (3. ed.). Rio de Janeiro, RJ, Brasil: LTC.

Psyllidis, A. (2015). OSMoSys: a web interface for graph-based rdf data visualization and ontology browsing. *Proceedings of the International Conference on Web Engineering 15*, 679-682. doi: 10.1007/978-3-319-19890-3_56

Tasca, J. E., Ensslin, L., Ensslin, S. R., & Alves, M. B. (2010). An approach for selecting a theoretical framework for the evaluation of training programs. *Journal of European Industrial Training*, 34(7), 631-655.

Vercruyssen, S., Venkatesan, A., & Kuiper, M. (2012). OLSVis: an animated, interactive visual browser for bio-ontologies. *BMC bioinformatics*, 13(1), 116.



COERÊNCIA NA REPRESENTAÇÃO TEMÁTICA DE ARTIGOS CIENTÍFICOS NO REPOSITÓRIO DE SAÚDE PÚBLICA DA FUNDAÇÃO OSWALDO CRUZ

Fatima Cristina Lopes dos Santos¹, Maria Cecília Magalhães Mollica²,
Vânia Lisboa da Silveira Guedes³

¹Fundação Oswaldo Cruz, Escola Nacional de Saúde Pública Sérgio Arouca, flopes@ensp.fiocruz.br

²Universidade Federal do Rio de Janeiro, Instituto Brasileiro de Informação em Ciência e Tecnologia, Programa de Pós-Graduação em Ciência da Informação, ceciliamollica@terra.com.br

³Universidade Federal do Rio de Janeiro, Programa de Mestrado Profissional em Letras, vanialisboa@facc.ufrj.br

RESUMO Trabalho originado da dissertação de mestrado defendida no Instituto Brasileiro de Informação em Ciência e Tecnologia/Universidade Federal do Rio de Janeiro, em fevereiro de 2017. A pesquisa investiga a coerência semântica na indexação de artigos de periódicos na área de saúde pública, em língua portuguesa, indexados no Repositório de Produção Científica da Escola Nacional de Saúde Pública Sérgio Arouca (ENSP) da Fundação Oswaldo Cruz (Fiocruz), visando a contribuir com a interoperabilidade semântica entre os repositórios dessa Fundação. Para tal, foram analisados 10 artigos de periódicos científicos brasileiros, publicados em acesso aberto entre 2012-2014 e cadastrados no Repositório de Produção Científica da ENSP/Fiocruz. Os artigos foram processados nos softwares RankWord e WordSmith 6.0, aplicando-se depois as leis de Zipf e Ponto de Transição de Goffman para identificação dos termos de alta carga semântica dos textos. Os termos selecionados em cada artigo foram comparados nos 5 (cinco) indicadores utilizados na pesquisa: Região T de Goffman; Expressões WordSmith 6.0, Palavras-chave; DeCS (Descritores em Ciências da Saúde) e Título. Os resultados foram compilados utilizando-se o plano discursivo das orações, com ênfase em termos usados como figura (foregrounding) e em termos usados como fundo (backgrounding), auxiliando na verificação do grau de coerência entre os artigos. Os dados coletados revelam baixo grau de coerência semântica na maioria dos artigos estudados. Sugere-se uma investigação com amostra ampliada para confirmação dos dados obtidos, principalmente se tratando de artigos de periódicos publicados em português na área de saúde e sua importância na comunicação científica nacional.

PALAVRAS-CHAVE Coerência, Indexação Temática, Área de Saúde, Artigos Científicos.

ABSTRACT The study investigates the semantic coherence in the indexing of articles of journals in the area of public health, in the area of public health, in the area of public health. Portuguese language, indexed in the Scientific Production Repository of the National School of Public Health Sérgio Arouca (ENSP) of the Oswaldo Cruz Foundation (Fiocruz), aiming to contribute with the semantic interoperability between the repositories of this Foundation. For that, 10 papers from Brazilian scientific journals, published in open access between 2012-2014 and registered in the Scientific Production Repository of ENSP/Fiocruz, were analyzed. The articles were processed in the software RankWord and WordSmith 6.0, applying later the laws of Zipf and Goffman Transition Point to identify the terms of high semantic load of the texts. The terms selected in each article were compared in the 5 (five) indicators used in the research: Goffman's T region; Expressions WordSmith 6.0, Keywords; DeCS (Descriptors in Health Sciences) and Title. The results were compiled using the discourse plan of the sentences, with emphasis on terms used as foregrounding and in terms used as background, helping to verify the degree of coherence between the articles. The data collected

reveal a low degree of semantic coherence in most articles studied. It is suggested an investigation with an enlarged sample to confirm the data obtained, especially when dealing with articles published in Portuguese in the health area and its importance in national scientific communication.

KEY WORDS *Coherence, Thematic Indexing, Health Area, Scientific Articles.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

1 INTRODUÇÃO

Para um alto grau de precisão na recuperação da informação na web - além da especificidade da linguagem de indexação - faz-se necessária a coerência semântica na representação da informação, visando a busca orientada em ambientes virtuais organizados para tal. Quando esses ambientes virtuais tratam de informação em saúde e do processo de comunicação científica, a recuperação da informação envolve o complexo modelo de negócio de periódicos científicos - principal canal de comunicação e divulgação entre pares.

Assim, este estudo analisa a gradação - redução ou aumento gradual e contínuo - da coerência semântica na indexação de artigos de periódicos na área de saúde pública, sistematizados no Repositório de Produção Científica da Escola Nacional de Saúde Pública Sergio Arouca (ENSP) da Fundação Oswaldo Cruz (Fiocruz). Do ponto de vista da função do periódico de canal formal de comunicação científica, a verificação do grau de coerência semântica na representatividade dos artigos de periódicos científicos poderá contribuir para a organização da informação, unificando padrões na representação temática e possibilitando a interoperabilidade semântica entre bases eletrônicas de dados.

A pesquisa se fundamenta em pressupostos teóricos e metodológicos das áreas de representação temática e da linguística documentária, no âmbito da Organização do Conhecimento na CI, bem como na Bibliometria, e tem como metas: (a) verificar comparativamente a coerência semântica na indexação de artigos científicos na área de saúde pública; (b) investigar a ocorrência de termos de indexação à luz de modelos quantitativos bibliométricos utilizados na indexação na Ciência da Informação; (c) fornecer evidências da importância do gênero discursivo artigo de periódico científico na área de saúde pública; (d) identificar os termos-figura e os termos-fundo dos artigos analisados para maior entendimento da informatividade de textos científicos; e) contribuir para o avanço do conhecimento acerca de fenômenos que norteiam a indexação da informação científica, particularmente sobre a coerência semântica na indexação da informação na área de saúde. As áreas de Ciência da Informação e Ciências da Linguagem oferecem sustentação para o desenvolvimento dessa pesquisa, na medida em que o desafio que se impõe é o de evidenciar a importância da coerência semântica na indexação de informações para a interoperabilidade de dados científicos.

O conceito de interoperabilidade não é recente na biblioteconomia. O compartilhamento de conteúdo entre bibliotecas é feito desde meados do século XX (Sayão & Marcondes, 2008). Porém, com a consolidação da internet, como via de acesso à informação e à educação, esse conceito tornou-se fundamental para a otimização da interação entre os repositórios. Andrade e Lara (2016) definem a interoperabilidade como a capacidade de dois ou mais sistemas trocar e usar informações sem um esforço especial de qualquer dos sistemas. Sendo assim, garantir a interoperabilidade torna-se o desafio de projetar serviços coerentes para uma diversidade de usuários, a partir de componentes que são

tecnicamente diferentes e gerenciados por diferentes organizações, exigindo um sofisticado grau de cooperação (Sayão & Marcondes, 2008).

A produção, a representação e a organização do conhecimento científico, especialmente publicado em artigos de periódico, gênero discursivo reconhecido como canal formal de comunicação científica, deve ser priorizada com o propósito de evitar buscas desordenadas e recuperação de informações irrelevantes. No que se refere à produção e publicação da informação científica, Asnake diz que:

Quando os resultados de um estudo de pesquisa não são publicados, outros pesquisadores não poderão apreciar o valor das evidências geradas, nem poderão construir sobre elas, e o conhecimento científico não poderá, de modo geral, crescer nem se desenvolver (Asnake, 2015, p.1972).

Para Marcondes e Campos (2008), existem dois obstáculos para o acesso e a utilização desse conhecimento, em larga escala: (i) grande número de publicações, impulsionado pela explosão informacional após a WWW; (ii) falta de estruturação desse conhecimento, tornando-o legível somente por pessoas. Berners-Lee, Hendler e Lassila (2001) mencionam que a maior parte do conteúdo disponibilizado na Web é projetada para a leitura por pessoas; eles ainda acrescentam que esse formato não estruturado impede que o conteúdo de um artigo científico seja processado por programas. Nesse sentido, os autores propõem a Web Semântica, como uma web mais estruturada, contendo informações com significados bem definidos.

No artigo *The Semantic Web*, Berners-Lee *et al.* (2001) esclarecem que os computadores deverão ter acesso a coleções estruturadas de informações e conjuntos de regras e inferências para conduzir o raciocínio automatizado, especificamente na representação e na recuperação da informação, onde o conteúdo de um texto será identificado e compreendido, semanticamente, por um programa eletrônico. Lancaster, ao reportar-se à representação da informação, menciona que:

As representações de textos consistirão no próprio texto completo, partes dele ou outra forma de representação construída por meios humanos ou automáticos. As representações de pedidos serão termos, termos apresentados em relações lógicas, enunciados textuais ou itens (Lancaster, 2004, p.284).

Na Web Semântica, um dos principais sistemas de organização do conhecimento - embasado na padronização terminológica - é a Ontologia, que possui funcionalidades que permitem que a máquina processe o raciocínio, automatizado, através de regras e inferências (Sales, Campos & Gomes, 2008). Segundo Guedes (2015), na estruturação de uma ontologia, deve-se definir um vocabulário comum, a partir do entendimento compartilhado formalmente pelos membros de uma comunidade de discurso científico. A autora esclarece que “as ontologias têm foco no estabelecimento de relações. Para tanto, são estruturadas linguagens formais que possibilitem a compreensão e o processamento, por parte das máquinas, de tais relações” (Guedes, 2015).

Como um dos pilares da Web Semântica, as ontologias são pensadas para viabilizar a interoperabilidade entre sistemas heterogêneos (conceitos) na Web semanticamente mais rica (Pickler, 2007). Marcondes e Campos (2008, p.110) afirmam que “quando o conhecimento é representado em uma linguagem declarativa, o conjunto de objetos que podem ser representados é chamado de universo do discurso”.

No que tange à indexação temática, como um dos procedimentos de representação temática, Lancaster (2004, p.10) afirma que “quanto mais especializada a clientela de um centro de informação, maior a

probabilidade de que a indexação possa e deva ser feita sob medida, ajustando-se com precisão aos interesses do grupo”. Quando uma ciência elabora seus conceitos e escolhe os termos que irão representá-los, ela constrói a terminologia da comunidade de discurso. Nesse sentido, ela se apresenta à comunidade científica indicando como gostaria de ser reconhecida e legitimada, especialmente por meio de um discurso específico, composto por termos e conceitos criados para denominar seus objetos concretos e abstratos. Benveniste diz:

Uma ciência só consegue se impor na medida em que define seus conceitos, por meio de denominações. Não existe outro modo de estabelecer sua legitimidade sem especificar, denominar seu objeto (Benveniste, 2006, p.249).

Questões relacionadas ao campo da indexação temática e recuperação da informação são objetos de estudo da Biblioteconomia e da Ciência da Informação. No capítulo sobre tradução técnico-científica, redação científica e gestão de informação, escrito por Krieger e Finatto, elas mencionam que:

Documentalistas, bibliotecários, enfim, profissionais que gestionam informação técnica e científica e a colocam à disposição de usuários, sabem o quanto é importante a familiarização com determinados assuntos e terminologias. As terminologias, para esses profissionais, valem, como elementos de indexação, como expressões de busca ou como palavras-chave de linguagem documentária. Na sua apreensão e relação com o texto especializado, diferente do que se passa com o tradutor, acentuam-se aspectos conceituais e temáticos (Krieger & Finatto, 2004, p.177)

Portanto, aspectos ligados ao mapeamento de relações lógica e ontológica, conceituais e temáticas devem ser considerados nas atividades de análise documentária e de indexação, onde se busca a adequação do termo utilizado no texto à linguagem de indexação estabelecida pelo sistema. Essas atividades estão relacionadas com a linguagem natural e a linguagem documentária, onde, a partir da análise do texto em linguagem natural, os conceitos mais relevantes são traduzidos para uma linguagem de especialidade própria da área em que será representado (Sales, 2007). Andrade, ao discorrer sobre a linguagem de especialidade, diz que ela é

Caracterizada pelo seu uso por profissionais nas diversas áreas de pesquisa e atuação, representando as áreas do conhecimento, as atividades, os processos, produtos e serviços de uma determinada área, servindo de instrumento de comunicação entre os sujeitos envolvidos nesses contextos (Andrade, 2010, p.31).

A indexação temática busca identificar os assuntos contidos no texto de um documento, traduzindo-os para uma linguagem de indexação natural - palavras que ocorrem no texto - ou linguagem de indexação controlada - termos/descriptores autorizados em vocabulário controlado (Lancaster, 2004). Quanto mais específica a representação das informações identificadas no texto em análise, maior probabilidade de recuperação de informações relevantes pelo sistema.

Na indexação, termos e descritores podem ser vistos como elementos de representação e disseminação do conhecimento, mas distinguem-se pela sua natureza e características básicas. O termo é uma unidade léxica que assume valor semântico próprio numa área de conhecimento, pela razão de integrar uma comunicação especializada. O descritor é um componente de uma linguagem constituída por gestores de informação, sendo assim, artificial (Krieger & Finatto, 2004). Por isso, a etapa de tradução de um documento, que segundo Lancaster (2004) envolve a conversão da análise conceitual num conjunto de termos de indexação, é de suma importância para a representação e está relacionada com a escolha de palavras-chave e de termos autorizados em um vocabulário controlado.

O vocabulário controlado utilizado pelo Repositório de Produção Científica da ENSP é denominado Descritores em Ciências da Saúde (DeCS), produzido pelo Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde (Bireme). O DeCS foi desenvolvido a partir do Medical Subject Headings (MeSH) da United States National Library of Medicine (NLM) e é um vocabulário trilingue que serve como linguagem única de indexação de artigos de revistas científicas, livros, anais de congressos, relatórios técnicos e outros tipos de materiais (DECS, 2014). O uso desse vocabulário no Repositório da ENSP objetiva auxiliar o processo de indexação de textos, indexados na base de dados do repositório, contribuindo assim para a interoperabilidade entre repositórios da área de saúde.

Sales (2007) afirma que as linguagens artificiais auxiliam as atividades de representação de conteúdos informacionais, viabilizando um armazenamento de informações que poderão ser facilmente recuperáveis. Bocatto e Torquetti (2012, p. 79) defendem que “o uso de uma linguagem de indexação controlada permite a representação na indexação e a pesquisa por assunto com maior exatidão e equidade com as necessidades informacionais dos usuários”.

Desta forma, são investigados o grau de coerência semântica entre (i) o conteúdo dos artigos selecionados, no repositório da ENSP; (ii) as palavras-chave atribuídas pelos autores dos artigos em análise; (iii) os descritores escolhidos no DeCS para a indexação desse conteúdo. Lancaster (2004, p.68) acrescenta que a coerência na indexação refere-se à extensão com que há concordância quanto aos termos a serem usados para indexar o documento. O autor afirma que quanto mais termos retirados do próprio conteúdo do texto, para indexação do documento, maior será a probabilidade de uma representação de qualidade.

1.1 CONTRIBUIÇÃO DA LINGÜÍSTICA DOCUMENTÁRIA

A Linguística Documentária surgiu na Espanha na década de 1990, e objetiva dar subsídios aos processos de representação para organização e recuperação de informações (Andrade, 2010). O termo Linguística Documentária foi utilizado por García Gutierrez em 1990, sugerindo uma disciplina que parte do princípio de que os problemas relacionados à informação são problemas de linguagem e, sendo assim, busca seus fundamentos nas ciências da linguagem (Lara & Tálamo, 2017).

Na Linguística Documentária, a partir da análise do mecanismo de funcionamento da língua e da Linguagem Natural, são estabelecidos os elementos para construção e uso de linguagens que darão conta de organizar, representar e comunicar informações em ambientes institucionais (Andrade, 2010). Para Tálamo e Lara (2006), a Linguística Documentária impõe-se como campo de estudos dos meios de representação da informação, em cujo vértice encontram-se as linguagens de processamento e produção da informação.

Segundo García Gutiérrez (1998 apud Lara & Tálamo, 2017) a Linguística Documentária recorre à linguística, à semântica, à gramática aplicada, à gestão da informação, como aos campos afins, como a análise do discurso, a análise do conteúdo e, de modo geral, as ciências cognitivas.

Sales (2007) entende a Linguagem Documentária como um sistema de signos que visa à uniformização do uso da linguagem de especialidade, proporcionando a representação padronizada do conteúdo informacional de textos científicos, bem como a recuperação da informação pertinente. Andrade (2010, p.33) cita que as “Linguagens documentárias delimitam os significados dos seus descritores para ofertar sentidos, de acordo com as necessidades de disponibilização e uso das informações”. Para Krieger e

Finatto (2004), o desafio está na apreensão e na divulgação da informação especializada, no sentido de que os conteúdos dos documentos devem ser identificados e adequadamente registrados, por meio do processo de indexação.

Essas Linguagens são constituídas por termos que, segundo Lara (2004), são designações que correspondem a um conceito em uma linguagem de especialidade, o termo é um signo linguístico que difere da palavra, unidade da língua geral, por ser qualificado no interior de um discurso de especialidade.

Nesta pesquisa, essas questões foram analisadas sob a perspectiva da Linguística Documentária em aproximação com perspectivas teóricas e metodológicas da Sociolinguística e do Funcionalismo, na Linguística. O Funcionalismo estuda a língua como um organismo não autônomo, produto e instrumento de comunicação, de persuasão, expressão e simulação, das manifestações humanas (Oliveira & Cezario, 2007). Para Fóris (2013), a função básica da linguagem é codificar, armazenar e transferir o conhecimento acumulado pela sociedade. A mesma autora acrescenta que a estrutura da linguagem e seus complexos processos de uso são abordados de formas diferentes, por vários ramos da Linguística. Sob esse ponto de vista, independentemente do processo de indexação temática realizado em linguagem natural ou controlada, o autor, ao elaborar um texto, seleciona, da terminologia de seu campo de conhecimento, termos/conceitos que ele julga adequados para explicitação e comunicação de suas ideias e reflexões.

Oliveira e Cezario (2007), sob a concepção do Funcionalismo, afirmam que a forma linguística tende a ser motivada pela função. Assim, eles explicam que, na produção de uma situação discursiva, fazemos escolhas lexicais e estruturais de acordo com nossos objetivos para obter sucesso na comunicação e, de alguma forma, atuar sobre os interlocutores. O autor de um texto molda o seu discurso de acordo com seus propósitos comunicativos e, nesse processo, ele elabora seu texto (aqui representado pelos artigos científicos), selecionando palavras e as representando tematicamente por palavras-chave, com o objetivo de chamar a atenção do leitor para os conceitos que ele considera mais importantes no texto.

Segundo Fóris (2013, p.423), o texto, como uma rede semântica construída por conceitos que precisam estabelecer coerência entre eles, não só “codifica a informação, mas garante que ela seja distribuída através do espaço e do tempo”. Lancaster (2004, p.68) esclarece que a coerência refere-se à extensão com que há concordância quanto aos termos a serem usados para representar conceitos, no processo de indexação temática do documento. Para Moreiro Gonzalez (1993), a coerência está relacionada ao significado alcançado através da disponibilidade e da relevância das relações, na estrutura do texto, e “é necessária para a distribuição da informação no texto” (Moreiro Gonzalez, 1993, p.134).

Em suma, no texto, o autor discursa chamando a atenção para partes ou conceitos onde acredita estar o âmago da sua produção. Neste processo, faz-se um paralelo com o estudo de Hopper (1979), citado por Oliveira e Cezario, sobre orações figura e orações-fundo:

[...] ao contar uma história, os usuários constantemente embalam as informações de acordo com suas percepções acerca do fato, das necessidades do ouvinte e dos objetivos comunicativos. Os pontos principais de uma história, suas ações, são vistos como figura (foregrounding) e os comentários, avaliações ou ações secundárias são o fundo (backgrounding) (Oliveira & Cezario, 2007, p.203).

Oliveira e Cezario (2007) explicam que, enquanto as orações-figura (foregrounding) se referem a situações objetivas, dinâmicas, as orações-fundo (backgrounding) muitas vezes se referem a elementos

subjettivos, estáticos. Segundo eles, isso não significa que as orações-figura (foregrounding) sejam mais importantes que as orações-fundo (backgrounding):

Podemos contar uma história para avaliar as atitudes de um participante, ou argumentar a favor de uma tese, e, assim, as orações fundo tornam-se mais importantes para atingir os objetivos comunicativos (Oliveira & Cezario, 2007, p.204).

Neste estudo, com base na perspectiva linguística de orações-figura e orações-fundo, foram analisados termos ou descritores. Acredita-se que a análise de termos-fundo e termos-figura pode ser processada e validada com o uso das Leis de Zipf e do Ponto de Transição de Goffman, associados à utilização do programa WordSmith 6.0, que possibilita a investigação qualitativa dos termos identificados, descrevendo a sua contextualização nos textos que compõem a amostra de pesquisa.

2 METODOLOGIA

Esta pesquisa quanti-qualitativa teve como propósito estimar o grau de coerência semântica de conceitos e termos abordados em 10 artigos científicos indexados no Repositório de Produção Científica da ENSP/Fiocruz, e está estruturada em 7 principais etapas: (i) composição dos corpora; (ii) contagem da frequência de uso das palavras que compõem cada corpus; (iii) cálculo da frequência que corresponde ao Ponto de Transição de Goffman, para cada corpus processado; (iv) delimitação da Região de Transição de Goffman, para cada corpus processado; (v) identificação das palavras com alto conteúdo temático, de cada corpus; (vi) contextualização dessas palavras, utilizando o software WordSmith; (vii) investigação da categorização discursiva dos termos, sob a perspectiva de termos-figura ou termos-fundo. Em suma, a pesquisa baseou-se na aplicação e análise de três principais conjuntos de procedimentos, representados na figura a seguir:

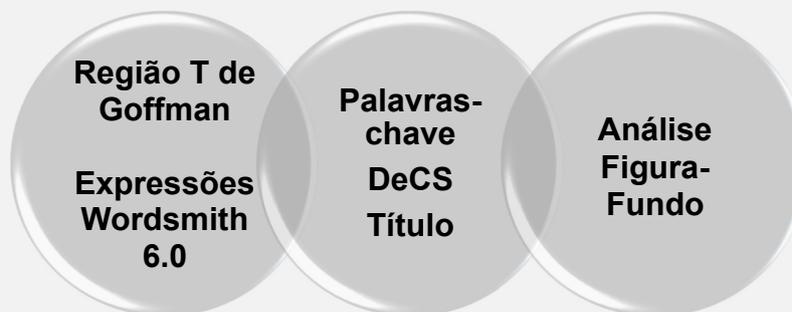


Figura 1 - Conjunto de informações estudadas nos artigos selecionados

Para a composição da amostra, foram selecionados 10 artigos de periódicos científicos sobre saúde pública, corroborando a abordagem defendida em Mollica e Guedes (2015), referente à importância do gênero discursivo artigos de periódicos, especialmente em estudos na área da Ciência da Informação. A seleção dos artigos seguiu os subseqüentes critérios: (i) artigos científicos cadastrados no Repositório de Produção Científica da ENSP; (ii) artigos científicos publicados no período de 2012-2014; (iii) artigos científicos publicados em periódicos científicos nacionais; (iv) artigos científicos publicados em língua portuguesa; (v) artigos científicos publicados em periódicos científicos de acesso aberto; (vi) artigos científicos publicados em periódicos classificados entre os estratos A1 a B3 no Sistema WebQualis da CAPES.

A amostra passou pelo seguinte procedimento: verificação da frequência de uso das palavras que ocorrem nos artigos selecionados com a utilização do programa RankWords; aplicação do procedimento de delimitação da região de transição de Goffman, onde ocorrem os termos com alta carga semântica dos artigos selecionados; contextualização dos termos de alta carga semântica, localizados na Região de Transição de Goffman, em expressões/sentenças dos artigos, localizadas pelo programa WordSmith 6.0; investigação crítica da ocorrência dos termos selecionados, nos indicadores de conteúdo (região T de Goffman e Expressões WordSmith 6.0) e de representatividade (palavras-chave, DeCS e título); categorização dos resultados das atividades citadas acima, utilizando as categorias referentes ao plano discursivo das orações-figura (foregrounding) e das orações-fundo (backgrounding) (OLIVEIRA; CEZARIO, 2007), no caso desta pesquisa, termos-figura e termos-fundo dos artigos analisados; representação figurativa do conteúdo dos artigos analisados, baseando-se nos termos-figura e termos-fundo dos textos, para auxiliar a identificação da coerência semântica desses artigos.

3 RESULTADOS

Os artigos selecionados foram publicados em 5 títulos de periódicos científicos. Nesses periódicos, 5 artigos foram publicados em 2014, 1 artigo publicado em 2013 e 4 artigos publicados em 2012. No que se refere à classificação dos títulos no sistema WebQualis, verificou-se que os estratos situam-se entre A2 e B3. Os artigos analisados apontam o uso de 41 palavras-chave, determinando a média de 4 palavras-chave por artigo. Dessas palavras-chave, 24 foram identificadas como descritores autorizados pelo DeCS, chegando-se à média de 2 (2,4) descritores autorizados por artigo. Assim, no processamento da amostra, a porcentagem de palavras-chave autorizadas pelo DeCS é de 59% (24) e a porcentagem de palavras-chave não autorizadas pelo DeCS é de 41% (17). O aproveitamento dos termos de maior frequência de uso para elaboração de palavras-chave foi nulo em 4 textos, e nos outros 6 artigos o percentual de aproveitamento foi baixo, chegando a 14% de aproveitamento, no máximo.

A relação das palavras-chave com o DeCS foi maior. Dos artigos estudados, 6 chegaram a 100% de aproveitamento, isto é, todas as palavras-chave desses artigos são descritores autorizados por esse vocabulário controlado. Percebeu-se a preocupação com a adequação ao DeCS, sem se dar conta de que isso pode acarretar num distanciamento da representatividade do conteúdo dos textos. Para Lancaster (2004), a etapa de tradução não deve ser influenciada pelas características do vocabulário a ser usado, pois se o mesmo não for representado adequadamente poderá não ser localizado corretamente numa pesquisa. O texto é uma rede de conceitos que devem estabelecer coerência, se essa rede não estiver bem construída por termos representativos desses conceitos, ela poderá se tornar inconsistente (Fóris, 2013). A inconsistência dessa rede poderá prejudicar a elaboração da interoperabilidade semântica entre bases de dados. A unidade terminológica, isto é, o termo, é simultaneamente, tanto elemento constitutivo da produção do saber quanto componente linguístico, cujas propriedades favorecem a univocidade da comunicação especializada (Krieger & Finatto, 2004).

A representação figurativa do conteúdo dos artigos, sob a perspectiva teórica e metodológica funcionalista figura-fundo (Oliveira & Cezario, 2007), ratifica este resultado. Nos artigos analisados, a minoria dos termos extraídos dos textos são termos-figura, isto é, termos mais frequentes e com maior relação com o texto, indicando um baixo grau de concordância entre o conteúdo do texto e sua representatividade.

4 CONCLUSÃO

Percebeu-se a necessidade de um olhar mais aprofundado sobre o grau de coerência semântica na representação de artigos publicados na área de saúde pública, pois a amostra estudada resultou em baixo grau de coerência semântica na maioria dos artigos. Para Tonello, Lunardelli e Almeida Júnior (2012), a informação deve estar organizada e representada o mais fielmente possível, sendo necessária a mediação entre o conteúdo informacional do documento e aquele que dele necessita. Tal processo pode ser concretizado por meio da elaboração de palavras-chave e descritores. Segundo Krieger e Finatto (2004) a existência e a circulação de terminologias em distintos cenários comunicativos são testemunhos de que essas cumprem, prioritariamente, a dupla função de fixar o conhecimento técnico-científico e de promover sua transferência de modo pontual. Na comparação com o vocabulário DeCS, os termos de maior frequência de uso nos textos analisados, isto é, os termos-figura, não foram representados pelo vocabulário, o que leva a supor que alguns desses vocabulários podem ser feitos em bases intuitivas, sem pesquisa. Nesse caso específico, a questão pode estar na tradução/adaptação do mesmo pois, como já foi dito, o DeCS foi desenvolvido a partir do Medical Subject Headings (MeSH).

Constata-se, portanto, a importância da coerência semântica na construção de um texto científico, por conseguinte, sua análise conceitual e tradução. É importante ressaltar que na indexação de um documento não há um conjunto correto de termos, isso depende de que tipo de clientela esse documento irá atender, mas é de suma importância a pesquisa dos itens e expressões utilizadas, tanto para a construção de um vocabulário controlado quanto para escolha dos termos que irão representar um texto científico (Lancaster, 2004). Quando o texto científico - principalmente o artigo de periódico que possui grandes restrições de formato - se constrói de forma coerente, numa construção de termos que representarão devidamente os conceitos inseridos no texto, a escolha dos termos para sua representação será mais eficaz. Guedes (2010) diz que o artigo de periódico permanece no centro de uma teia formada por gêneros textuais de publicações de processo de pesquisa e que as interrelações são bem conhecidas e reconhecidas no ambiente acadêmico. Migués *et al.* (2013) dizem que o uso das palavras-chave potencializa o acesso ao conteúdo dos documentos, para além da informação que é representada pelo título e resumo; traduz o pensamento dos autores, e mantém o contato com a realidade da prática cotidiana, acompanhando a evolução científica e tecnológica, que é refletida pelos documentos.

Sugere-se uma investigação com amostra ampliada para confirmação desses dados, principalmente se tratando de artigos de periódicos publicados em português na área de saúde e sua importância na comunicação científica nacional. Os artigos de periódicos das ciências da saúde predominam na distribuição das citações por tipo de literatura (Packer, 2011), e o Brasil está em 23º lugar no ranking internacional de produção científica de artigos, ocupando a melhor posição na América Latina (Nature, 2015).

REFERÊNCIAS BIBLIOGRÁFICAS

- Andrade, J. (2010). *A linguística documentária e a análise de domínio na organização da informação e do conhecimento*. Dissertação (Mestrado em Ciência da Informação). Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo.
- Andrade, J. & Lara, M. L. G. (2016). Interoperability and mapping between knowledge organization systems: metathesaurus-unified medical language system of the National Library of Medicine. *Knowledge Organization*, 43(2), 107-112.
- Asnake, M. (2015). A importância da publicação científica para o desenvolvimento da saúde pública. *Ciência & Saúde Coletiva*, 20(7), 1972.
- Benveniste, É. (2006). Gênese dos termos científicos. In: _____. *Problemas de linguística geral II*. 2.ed. Campinas: Pontes.
- Berners-Lee, T., Hendler, J. & Lassila, O. (2001). The semantic web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, New York, 5, May 2001. Recuperado em 2 dezembro, 2014, de http://www.ryerson.ca/~dgrimsha/courses/cps720_02/resources/Scientific%20American%20The%20Semantic%20Web.htm
- Bocato, V. R. C. & Torquetti, M. C. (2012). Interoperabilidade entre linguagens de indexação como recurso de modelagem de repertório terminológico de coordenadorias de comunicação social em ambientes universitários: uma proposta metodológica. *Informação & Informação*, 17(3), 76-101. Recuperado em 8 agosto, 2013, de <http://www.uel.br/revistas/uel/index.php/informacao/article/view/10800/pdf>
- Descritores em Ciências da Saúde (DeCS). Recuperado em 20 setembro 2014, de <http://decs.bvs.br/P/decsweb2014.htm>
- Fóris, Á. (2013). Network theory and terminology. *Knowledge Organization*, 40(6), 422-429.
- Guedes, V. L. S. (2010). *Nominalizações deverbais em artigos científicos: uma contribuição para a análise e a indexação temática da informação* (Tese de Doutorado em Linguística). Programa de Pós-Graduação em Linguística, Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- Guedes, V. L. S. (2015). Organização de domínios do conhecimento: taxonomias, mapas conceituais, tesouros, ontologias, folksonomias. In: Mollica, M. C. M., Batista, H. R. & Guimarães, L. S. (Orgs.). *Cybercorpora e inovação com práticas de ensinagem*. Curitiba: CRV.
- Krieger, M. G. & Finatto, M. J. B. (2004). *Introdução à terminologia: teoria e prática*. São Paulo: Contexto.
- Lancaster, F. W. (2004). *Indexação e resumos: teoria e prática* (2a. ed. rev. atual). Brasília: Briquet de Lemos/Livros.
- Lara, M. L. G. (2004). Diferenças conceituais sobre termos e definições e implicações na organização da linguagem documentária. *Ciência da Informação*, 33(2), 91-96.

Lara, M. L. G. & Tálamo, M. F. G. M. (2017). *Linguística Documentária e Terminologia: experiência didática na interface das disciplinas*. Recuperado em 19 janeiro 2017, de <http://enancib.ibict.br/index.php/enancib/vii/enancib/paper/viewFile/2426/1557>

Marcondes, C. H. & Campos, M. L. A. (2008). Ontologia e Web Semântica: o espaço da pesquisa em Ciência da Informação. *Ponto de Acesso*, 2(1), 107-136. Recuperado em junho 2010, de <http://www.portalseer.ufba.br/index.php/revistaici/article/view/2669/1885>

Migués, A., Neves, B., Silva, A. L., Trindade, Á., & Bernardes, J. A. (2013). A importância das palavras-chave dos artigos científicos da área das Ciências Farmacêuticas, depositados no Estudo Geral: estudo comparativo com os termos atribuídos na MEDLINE. *InCID: Revista de Ciência da Informação e Documentação*, 4(2), 112-125.

Mollica, M. C. M. & Guedes, V. L. S. (2015). *Gêneros discursivos e Ciência da Informação*. Recuperado em 13 maio de 2015, de <http://www.ppgci.ufrj.br/teste.html?download=213%Ageneros-discursivos-e-ciencia-da-informacao>

Moreiro González, J. A. (1993). *Aplicación de las ciencias del texto al resumen documental*. Madrid: Universidad Carlos III de Madrid.

Nature Index tables. (2015, 18 de junho). *Nature*. Recuperado de http://www.nature.com/nature/journal/v522/n7556_supp/full/522S34a.html

Oliveira, M. R. & Cezario, M. M. (2007). PCN à luz do funcionalismo linguístico. *Linguagem & Ensino*, 10(1), 87-108.

Packer, A. L. (2011). Os periódicos brasileiros e a comunicação da pesquisa nacional. *Revista USP*, 89, 26-61.

Pickler, M. E. V. (2007). Web Semântica: ontologias como ferramentas de representação do conhecimento. *Perspectivas em Ciência da Informação*, 12(1), 65-83.

Sales, L. F., Campos, M. L. A. & Gomes, H. E. (2008). Ontologias de domínio: um estudo das relações conceituais. *Perspectivas em Ciência da Informação*, 13(2). Recuperado em 1 julho 2010, de <http://www.scielo.br/pdf/pci/v13n2/a06v13n2.pdf>

Sales, R. (2007). Suportes teóricos para pensar linguagens documentárias. *Revista Digital de Biblioteconomia e Ciência da Informação*, 5(1), 96-114.

Sayão, L. F. & Marcondes, C. H. (2008). O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. *TransInformação*, 20(2), 133-148.

Tálamo, M. F. G. M. & Lara, M. L. (2006). O campo da linguística documentária. *TransInformação*, 18(3), 203-211.

Tonello, I. M. S., Lunardelli, R. S. A. & Almeida Junior, O. F. (2012). Palavras-chave: possibilidades de mediação da informação. *Ponto de Acesso*, 6(2), 21-34.



A PROFUNDIDADE TEMÁTICA DA ANÁLISE DE ASSUNTO DE INDEXADORES NO BRASIL E PORTUGAL: ESTUDO DA COGNIÇÃO COM MAPAS CONCEITUAIS

Dulce Amélia de Brito Neves¹, Mariângela Spotti Lopes Fujita²

¹Universidade Federal da Paraíba, João Pessoa – PB – Brasil, damelia1@gmail.com

²Universidade Estadual Paulista Júlio de Mesquita Filho, Campus de Marília, Marília, SP, Brasil, fujita@marilia.unesp.br

RESUMO Especificidade e exaustividade são dois aspectos divergentes, porém complementares, do nível de profundidade temática da análise de assunto durante o processo de indexação. Ao estabelecer a política de indexação os níveis de exaustividade e especificidade são determinados em função da identificação e seleção de termos durante a análise de assunto na indexação. A questão que norteou a pesquisa foi: Qual o nível de profundidade temática que o bibliotecário alcança na análise de assunto? Realizou-se observação da análise de assunto para elaboração de mapas cognitivos com o objetivo de avaliar procedimentos e analisar os níveis de profundidade temática dos termos identificados pelos bibliotecários. Os resultados da pesquisa apontam a importância do controle de vocabulário como ferramenta essencial ao desempenho do indexador na análise de assunto. Conclui-se que a profundidade temática na análise de assunto está relacionada ao nível de precisão das linguagens de indexação.

PALAVRAS-CHAVE *Análise de Assunto. Indexação. Mapa Conceitual. Representação da Informação. Ciência da Informação*

ABSTRACT Specificity and completeness are two different aspects, but complementary, level of thematic depth subject analysis during the indexing process. By establishing the indexing policy the completeness and specificity levels are determined depending on the identification and selection of terms for the subject analysis in index-tion. The question that guided the research was: What level of thematic depth that reaches the librarian in the subject analysis? Held observation of subject analysis for development of cognitive maps in order to evaluate procedures and analyse the thematic depth levels of the terms identified by librarians. The survey results point to the importance of vocabulary control as an essential tool to index performance in the subject analysis. We conclude that the thematic depth in the subject of analysis is related to the level of accuracy of indexing languages.

KEYWORDS *Subject Analysis. Indexing. Conceptual map. Representation of Information. Information Science.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A Ciência da Informação (CI) opera com a linguagem de indexação numa perspectiva da compreensão do fenómeno informacional para organizá-lo e representá-lo estabelecendo convergência com estudos voltados aos processos comunicacionais dos registros do conhecimento e, com isso, estimular investigações científicas consistentes relacionadas ao estudo dos vários fenómenos referentes à noção de informação, sejam eles encontrados nos processos biológicos, psicológicos da dimensão humana, ou nas máquinas.

A CI dialoga e interage com áreas que envolvem a comunicação humana, o conhecimento da informação e seus registros, as necessidades e os usos da informação, em contextos sociais, institucionais e individuais.

Os mapas conceituais pretendem traçar a representação no cérebro abrangendo as representações, modelos mentais ou esquemas que os indivíduos elaboram a partir de suas interações no ambiente social e durante a aprendizagem ou na resolução de problemas.

O uso de mapas conceituais na CI é de grande interesse, tendo em vista que possibilita a construção da representação conceitual dos processos mentais dos indivíduos em suas buscas e/ou em seu trabalho cotidiano em sistemas de informação, visto que o mapeamento de um conhecimento permite conhecer a hierarquia do conceito e sua formação na mente.

Nesse sentido, descortinam-se várias possibilidades voltadas aos profissionais bibliotecários como usuários de sistemas de informação que podem esclarecer seu conhecimento de habilidades utilizadas, tanto no tratamento da informação quanto no seu cotidiano.

No que tange ao tratamento da informação é instigante a compreensão do processo mental da variável leitor sua cognição profissional que norteia a atividade de análise de assunto de conteúdos textuais para realizar a indexação orientada pelo conteúdo conforme nível de profundidade temática determinada por uma política de indexação da biblioteca ou sistema de bibliotecas.

A profundidade da indexação é definida por duas variáveis: a exaustividade e a especificidade. Embora diferentes uma da outra são complementares na medida em que juntas definem o nível de profundidade da indexação e são determinantes para a precisão e revocação na recuperação da informação mesmo sem uso de controle de vocabulário. São pois interdependentes e definidas pela política de indexação em função dos requisitos de um sistema de informação tais como a organização, os usuários e a infraestrutura.

Considerando-se que a identificação de termos pelo indexador é realizada durante a análise de assunto e orientada pelo conteúdo documental sob influência das variáveis exaustividade e especificidade, esta pesquisa teve como questão norteadora: Qual o nível de profundidade temática que o bibliotecário alcança na análise de assunto para a identificação de termos que representarão o conteúdo dos documentos?

Assim, o objetivo é investigar a atividade de análise de assunto durante o processo de indexação realizado por bibliotecários do Brasil e Portugal, visando analisar por meio de mapas conceituais os níveis de profundidade temática dos termos selecionados pelos bibliotecários sem o uso de instrumentos de controle de vocabulário, tais como linguagens de indexação.

ANÁLISE DE ASSUNTO E A PROFUNDIDADE DA INDEXAÇÃO

O tratamento temático da informação implica a realização de diferentes operações de análise de conteúdos documentais para diferentes finalidades de representação documental. Essas operações são conhecidas como classificação, indexação e elaboração de resumos. A indexação é uma operação de tratamento temático cujo processo realiza análise de assunto pela identificação e seleção de termos para a representação do conteúdo documental com a finalidade de recuperação por assuntos mediante índices ou estratégia de busca em sistemas online.

A análise de assunto é a etapa inicial do processo de indexação seguida pela etapa de tradução, assim denominada por utilizar uma linguagem de indexação que "traduz" os termos identificados e os seleciona para a representação do conteúdo documental. A análise de assunto se realiza durante a leitura documental para a identificação e seleção dos termos do conteúdo documental abordados pelo autor. Nesse sentido, o indexador necessita conhecer o conteúdo dos documentos, pois do seu desempenho na identificação de termos depende a precisão da recuperação desse conteúdo.

Para Pinto Molina (1993), o indexador precisa ter habilidades cognitivas que o levem, como profissional especializado, a processar a informação textual na busca pelo conteúdo intrínseco do documento durante a análise de assunto.

A análise de assunto é uma das atividades fundamentais dos profissionais da CI, considerado o momento crucial do processo de indexação. Como é baseada na experiência e no conhecimento adquirido pelo indexador é revestida de inerente subjetividade que pode ser controlada pela definição de uma política de indexação para determinação de elementos e variáveis com vistas a obter precisão na recuperação da informação.

Durante sua formação o bibliotecário-indexador é instruído a ater-se a algumas partes do documento tidas como mais importantes: título e subtítulo; resumo se houver; sumário; introdução; ilustrações, diagramas, tabelas e seus títulos explicativos; palavras em destaque e referências do documento, como estabelece a NBR 12.676, da Associação Brasileira de Normas Técnicas (1992). Para a identificação dos termos a Norma propõe um questionamento direcionado aos conceitos identificadores de termos tais como a ação, objeto da ação, agente que praticou a ação e etc.

A complexidade da análise de assunto está na identificação dos termos por meio da análise conceitual que a norma expõe por meio do questionamento considerando-a uma abordagem sistemática do conteúdo documental. Entretanto, ao identificar os termos, o indexador, ao mesmo tempo, seleciona os termos com base na política de indexação que define, *a priori*, a profundidade da indexação pelas variáveis de exaustividade e especificidade.

Cleveland e Cleveland (2001, p.105) consideram que "exaustividade e especificidade são dois lados de uma moeda denominada profundidade da indexação. Quando um item é indexado exaustivamente e especificamente, então o item foi indexado em profundidade".

As definições de exaustividade na literatura referem-se a dois aspectos: ao aspecto quantitativo de termos atribuídos pelo indexador (ABNT, 1992, Lancaster, 2004); e, ao aspecto de extensão da cobertura temática do conteúdo representado pelos termos atribuídos pelo indexador (Soergel, 1994, Gil Leiva, 2008). A especificidade tem relação direta com a exaustividade, pois a quantidade de termos não é suficiente para se ter exaustividade, é necessário que cada termo represente de modo específico o

conteúdo do documento como na figura abaixo na qual se demonstra que "o indexador poderá fazer a cobertura temática do conteúdo com termos mais específicos e genéricos de modo que na recuperação o usuário poderá ter mais opções de acesso, maior precisão e até mais revocação" (Fujita, 2012, p.28).

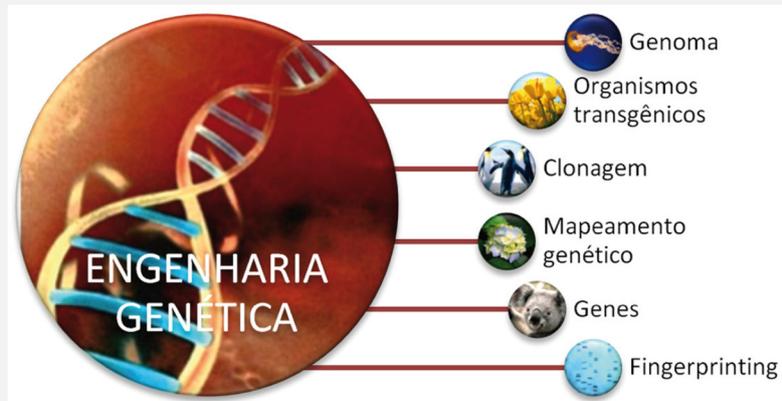


Figura 1: Quantidade de termos específicos e exaustividade

Fonte: Fujita (2012, p.28)

Desse modo, não basta o indexador identificar os termos na análise de assunto é necessário que a seleção dos termos seja orientada pela profundidade da indexação definida pela exaustividade e especificidade. Essa investigação avança para verificar qual o nível de profundidade temática que o bibliotecário alcança na análise de assunto para a identificação de termos que representarão o conteúdo dos documentos, serão analisados os termos identificados e selecionados por bibliotecários de Brasil e Portugal tomando como referência o uso de linguagens de indexação que indiquem os níveis de exaustividade e especificidade de cada termo nas respectivas estruturas lógico-hierárquicas.

PROCEDIMENTOS COM USO DE MAPAS CONCEITUAIS

Não há uma linearidade no pensamento dos seres humano que é formado por uma rede de associação de idéias, semelhante ao hipertexto, cujos links são estabelecidos entre palavras, imagens, gráficos, sons, etc.

Os estímulos diversos causam diferentes reações, assim, ao ouvirmos uma palavra, um som ou uma imagem, fazemos inúmeras conexões a possíveis significados e sua influencia em nossa integridade. O conhecimento é armazenado de duas formas: linguística (semântica), que inclui fala e leitura, e não-linguística, a qual inclui imagens mentais e até mesmo os sentidos físicos como olfato, audição, tato etc. Um novo saber é constituído a partir da conjunção de conhecimentos a partir das informações recebidas e selecionadas tendo em vista o interesse dos individuais.

As técnicas de mapeamento de informação surgiram na década de 60 e , em 1988, Joseph Donald Novak desenvolveu estudos sobre aprendizagem utilizando como ferramenta o mapa conceitual, constituído de um esquema que registra as informações localizando-as em um mapa, mostrando as ligações entre si. Expressando visualmente a representação do conhecimento a partir das conexões entre estes conceitos expressos em mapas.

No entanto, não foi identificado o uso de mapas conceituais em pesquisas na CI, apenas existem estudos teóricos que os apontam como uma técnica adequada em estudos relacionados às necessidades de

usuários da informação (Silva Ferreira; Borges, 2002), em estudos bibliométricos (Alvarenga, 1998), ao processo de indexação (Fujita, (2003), Neves, 2007a), entre outros.

Como diz Moreira Gonzalez (2004) há uma necessidade crescente em adotar métodos que venham a melhor contribuir para preencher as lacunas multidisciplinares e “atender precisamente a un mundo de información cambiante y enrecimiento, se empezó a trabajarenlosconocidos mapas conceptuales de navegación” (p. 01).

Assim, nosso objetivo foi “elaborar mapa cognitivo da atividade de análise de assunto durante o processo de indexação, no Brasil e em Portugal, visando explicitar os procedimentos dos indexadores com vistas a sugerir abordagens mais adequadas traçando um paralelo nos procedimentos” (Neves 2011a, 2011b).

As leituras nos possibilitaram observar similaridades nos procedimentos tendo em vista que na Licenciatura em Ciência da Informação da Universidade do Porto, Portugal, tem apenas uma disciplina voltada à classificação que é ministrada no 2º ano: Metainformação Descritiva. Enquanto na Universidade Federal da Paraíba, Brasil, tem duas disciplinas: Representação Temática da Informação I, no 3º semestre e Representação Temática da Informação II no 4º semestre.

Assim, acreditamos que para uma comparação mais acurada faz-se necessário estudo mais curado do curriculum de ambas as Universidades à luz da legislação vigente em ambos os casos, como também diferenças culturais.

No primeiro momento, fizemos um convite a seis bibliotecários que preenchiam o perfil necessário e explicitamos a motivação da pesquisa que se refere ao tempo de trabalho em indexação, pois segundo pesquisas de Farrow (1995), o uso frequente das linguagens de indexação possibilita a automatização da análise do conteúdo dos documentos.

Dentre os bibliotecários convidados, três brasileiros e o mesmo número de portugueses participaram da pesquisa. Buscando minimizar diferenças entre os sujeitos dos dois grupos, observaram-se semelhanças no perfil dos participantes da pesquisa. Todos tinham mais de dois anos na atividade de indexação, tendo em vista ser este o tempo mínimo para o desenvolvimento cognitivo desta habilidade (Bertrand e Cellier, 1995, Farrow, 1995, Neves, 2004).

Para o desenvolvimento dos mapas cognitivos, foram selecionados três textos, sendo um sobre Astrofísica intitulado “Astrofísica observacional” do Instituto Nacional de Pesquisa Espacial (2003), um de Ciência da Informação de Fernanda Ribeiro (2010) "O Papel mediador da Ciência da Informação na construção da sociedade em rede" e um de análise do discurso de João Bôsko Cabral dos Santos (2008) de "A instancia enunciativa sujeitidual". Os textos, apesar de estarem redigidos em linguagem científica, contemplaram áreas diversas, o que para o indexador não deve ser uma barreira, pois com técnica e experiência eles têm capacidade e condições para procederem à indexação com eficiência e eficácia. Não foi delimitado tempo nem número de descritores.

Após a leitura, os sujeitos da pesquisa elaboraram uma lista sem limitação do número de conceitos que formaram os mapas de cada indexador e dos indexadores, comparando os mapas dos bibliotecários-indexadores da Universidade do Porto/Portugal com os mapas elaborados por bibliotecários-indexadores do Brasil.

ANÁLISE E DISCUSSÃO

Com os mapas foi possível realizar uma avaliação da atividade de análise de assunto dos indexadores para explicitação dos procedimentos utilizados, bem como a análise de profundidade temática a partir dos termos selecionados pelos bibliotecários-indexadores tomando como parâmetro de análise os níveis hierárquicos de três linguagens de indexação utilizadas por bibliotecas: a Terminologia de assuntos da Biblioteca Nacional do Brasil¹, a *Library of Congress Subject Headings*² e o Vocabulário Controlado do SIBI/USP³ para os termos selecionados por bibliotecários-indexadores nos textos 1, 2 e 3 que são, respectivamente, das áreas de Astronomia, Ciência da Informação e de Análise do Discurso.

As três linguagens de indexação foram construídas para uso em diversas modalidades de bibliotecas, sendo que apenas o Vocabulário USP tem a especificidade de haver sido desenvolvido para bibliotecas universitárias. A fonte da Terminologia de Assuntos da Biblioteca Nacional é a LCSH, ou seja, seu vocabulário controlado é estruturado e traduzido conforme a LCSH (Grings, 2015). O Vocabulário Controlado do SIBI/USP constitui-se de uma linguagem interdisciplinar, que abrange as diversas áreas do conhecimento (Universidade de São Paulo, 2015). A escolha dessas três linguagens de indexação teve como critérios serem gerais e conterem todas as áreas de conhecimento especializadas.

Dessa forma, as linguagens de indexação, a Terminologia de assuntos da Biblioteca Nacional do Brasil (TBN), a *Library of Congress Subject Headings* (LCSH) e o Vocabulário Controlado do SIBI/USP (VocaUSP) fornecerão a informação sobre o grau de especificidade e exaustividade dos termos pelo nível hierárquico, como termo genérico ou específico, que ocupam na estrutura lógico hierárquica como por exemplo:

Animais (termo genérico)
...Marsupiais (termo específico)
.....Canguru (termo específico)

AVALIAÇÃO DA ATIVIDADE DE ANÁLISE DE ASSUNTO

Para Tálamo (1987) o conhecimento prévio da estrutura textual é de grande importância. Como todos os bibliotecários tinham experiência de mais de dois anos na atividade de indexar conheciam tinham plenas condições de explorar os textos em sua estrutura de organização.

Estas afirmações podem ser demonstradas no Mapa 01 (Apêndice I) no qual os termos foram mais fáceis de identificar por ser o artigo sobre Astrofísica um texto técnico. Assim, observamos uma grande semelhança entre os conceitos escolhidos pelos indexadores brasileiros e portugueses.

Como já dissemos anteriormente, os indexadores foram instruídos a não usarem instrumentos de controle terminológico e com isto percebemos um “descontrole” na elaboração dos conceitos.

¹ http://acervo.bn.br/sophia_web/index.html

² <http://id.loc.gov/authorities/subjects.html>

³ <http://143.107.154.62/Vocab/Sibix652.dll>

No segundo artigo de autoria da Professora Fernanda Ribeiro intitulado “O Papel mediador da Ciência da Informação na construção da sociedade em rede” (Apêndice II) já se percebe uma divergência não apenas entre os indexadores brasileiros e portugueses, mas entre eles próprios. Mesmo sendo um artigo da área de formação dos indexadores começamos a perceber a identificação de apenas um termo por alguns deles, como por exemplo: “Informação”, “História”, “Profissão”, “Evolução”, dentre outros, cuja representatividade enquanto termo indexador é vazio de significado.

Por outro lado, foram escolhidos como “conceitos” frases inteiras, por exemplo: “Fundamentos teóricos e epistemológicos da Ciência da Informação”, “História da Ciência da Informação”.

O terceiro artigo do Professor João Bôsko Cabral dos Santos “A instancia enunciativasujeitudinal” (Apêndice III) trata de análise do discurso que é um conhecimento carregado de subjetividade e não muito familiar a Biblioteconomia. Alie-se a isto o vocabulário não muito usual usado pelo autor.

Assim, também percebemos uma pulverização dos conceitos selecionados semelhante a tematização, ou seja termos únicos, soltos sem agregar valor ou peso terminológico, como: Significações, Significado, Sentido, Discurso, Linguagem, dentre outros.

ANÁLISE DA PROFUNDIDADE TEMÁTICA

Observamos que, no primeiro texto, houve quantidade quase semelhante de descritores como se observa no Quadro 1. A partir do segundo texto, os sujeitos da pesquisa identificaram grande número de descritores sem distinção de nacionalidade:

Quadro 1: termos identificados na leitura do primeiro artigo da área de Astronomia

Indexadores	Brasil	Portugal
Indexador 1	7	3
Indexador 2	4	5
Indexador 3	6	4

Quadro 2: termos identificados na leitura do segundo artigo da área de Ciência da informação

Indexadores	Brasil	Portugal
Indexador 1	7	5
Indexador 2	5	7
Indexador 3	5	7

Quadro 3: termos identificados na leitura do terceiro artigo da área de Análise do Discurso

Indexadores	Brasil	Portugal
Indexador 1	9	5
Indexador 2	6	8
Indexador 3	4	5

Considerando-se a quantidade de termos identificados pelos indexadores é possível atribuir maior grau de exaustividade ao indexador 1BR que atribuiu 23 termos nos 3 artigos e ao indexador 2PT que atribuiu 20 termos. De maneira geral, todos os indexadores obtiveram um grau de exaustividade com média de

5 termos por indexador no artigo da área de Astronomia, 6 termos no artigo da área de Ciência da Informação e 6 termos no artigo da área da Análise do Discurso.

Além da quantidade, a exaustividade é entendida pela abrangência temática que os termos alcançam na representação do conteúdo textual. Na análise dos termos identificados pelos indexadores apresentados nos mapas 01, 02 e 03 constatam-se termos genéricos e específicos conforme estrutura lógico-hierárquica das linguagens de indexação consultadas.

Na configuração dos termos identificados na leitura do primeiro artigo termo "Astrofísica" foi identificado por todos os indexadores. Tem uma abrangência maior que todos os demais termos e, portanto, representa exaustivamente o conteúdo do artigo. De acordo com as linguagens de indexação Terminologia de Assuntos da Biblioteca Nacional (TBN), Vocabulário Controlado da USP (VocaUSP) e Library Congress Subject Headings (LCSH) observa-se que o termo "Astrofísica" é um descritor específico, de primeiro nível hierárquico dos termos genéricos "Astronomia", "Física" e "Física Cósmica":

Quadro 4: Categorias do termo "Astrofísica" nas linguagens de indexação consultadas

VocaUSP, TBN, LCSH	TBN, LCSH	TBN, LCSH
TG ASTRONOMIA	TG FÍSICA	TG FÍSICA CÓSMICA
TE Astrofísica	TE Astrofísica	TE Astrofísica

Os termos precedidos por TG são termos genéricos denominativos de categorias de assunto, ou seja, são termos superordenados com níveis hierárquicos de termos subordinados que não possuem outro termo genérico acima.

Outro termo identificado por todos os indexadores é "Meio interestelar" que na TBN e LCSH é um termo não preferido substituído pelo termo "Matéria interestelar" que, por sua vez, é termo subordinado de segundo nível hierárquico à "Astrofísica". Outros termos identificados, "Galáxias", "Estrelas", "Via-Láctea", "Planetas" e "Sistema Solar", são pertencentes à mesma hierarquia de "Astronomia". Os níveis hierárquicos podem ser representados, conforme adaptação das linguagens de indexação consultadas, como segue:

Quadro 5: Hierarquias dos termos identificados no artigo 1

TBN	LCSH	VocaUSP
TG ASTRONOMIA	BT ASTRONOMY	BT ASTRONOMIA
. Astrofísica	. Galaxies	. Astrofísica
.. Matéria interestelar .Galáxias	.. Stars	..Astrofísica estelar
.. Estrelas	.. Spiral Galaxies	...Estrelas
... Sistema Solar	... Milky Way	. Meio interestelar
.... PlanetasSolar System	. Mecânica Celeste
.. Via-LácteaPlanets	.. Planetas
	. Astrophysics	. Sistema solar
	..Interstellarmatter	

Os níveis hierárquicos entre os termos identificados pelos indexadores demonstram que existem 4 níveis hierárquicos demonstrando que existe um nível de especificidade com grau de profundidade de 4 níveis entre os termos. O indexador 1BR utilizou maior quantidade de termos com grau de profundidade de 4 níveis e os demais de 3 níveis. O que diferenciou o indexador 1BR dos demais foi a identificação do termo "Planetas".

No **segundo artigo** da área de Ciência da Informação, embora se suponha, em princípio, que o vocabulário seja de domínio dos indexadores, observa-se pelas análises de profundidade temática que o grau de profundidade dos termos identificados em Ciência da Informação foi até o nível 3 conforme consulta às linguagens de indexação:

Quadro 6: Hierarquias dos termos identificados no artigo 2

TBN	LCSH	VocaUSP
TG COMUNICAÇÃO . Ciência da Informação .. Biblioteconomia ... Estudo de Usuário .. Documentação .. Arquivística .. Informação ...Mediação da Informação	TG COMMUNICATION . Information Science RT Library Science ..Documentation ... Archives TG HUMANITIES . Philosophy .. Psychology .. Social Groups Social Psychology Interpersonal relations UF Social Behavior Social Network	TG CIÊNCIA DA INFORMAÇÃO . Biblioteconomia . Arquivística . Documentação . Informação .. Mediação da informação . Usuários da informação
TG FILOSOFIA . Epistemologia		TG FILOSOFIA . Teoria do conhecimento .. Epistemologia
TG EDUCAÇÃO . Ensino .. Interdisciplinaridade		TG EDUCAÇÃO . Ensino .. Interdisciplinaridade
TG PSICOLOGIA . Psicologia aplicada .. Psicologia social ... Processos interpessoais e de grupo Comportamento social Grupos sociais (psicologia) Redes sociais	TG TEACHING RT Education	TG PSICOLOGIA . Psicologia aplicada .. Psicologia social ... Processos interpessoais e de grupo Comportamento social Grupos sociais (psicologia) Redes Sociais

Os demais termos identificados pertencem, segundo linguagens de indexação, à outras categorias hierárquicas. O termo "redes sociais" é o único de nível 6 e pertence à categoria de "Psicologia" o que, de certa forma, comprova os "empréstimos conceituais" de diferente áreas com as quais a Ciência da Informação mantém relações interdisciplinares para os seus diversos aportes teóricos e metodológicos.

Outros termos identificados tais como "bibliotecário" e "profissional da informação", apesar de serem específicos do domínio da Ciência da Informação não estão presentes nas linguagens de indexação. Outros termos como "profissão", "era digital", "sociedade em rede", "história", "evolução" e "transdisciplinaridade", além de não estarem presentes nas linguagens, podem ser considerados menos representativos do conteúdo do artigo e até vazios de significado como é o caso de "evolução".

No **terceiro artigo** da área de análise do discurso a quantidade de termos identificados é grande, porém, somente o termo "análise do discurso" foi selecionado por todos os indexadores e, mesmo assim, com algumas variações. O termo "instância enunciativa sujeitudinal" é um termo importante, sob ponto de vista da representatividade do conteúdo textual e foi selecionado por 5 indexadores, mas não está presente nas linguagens de indexação, por isso tomamos a decisão de incluí-lo como termo subordinado ao termo "enunciação".

Dessa forma e considerando-se a consulta às linguagens de indexação observa-se a seguinte hierarquização entre termos genéricos e mais exaustivos com os termos mais específicos segundo grau de profundidade hierárquica:

Quadro 7: Hierarquias dos termos identificados no artigo 3

TBN	LCSH	VocaUSP
TG LINGUAGEM E LÍNGUAS . Linguística . Oratória TG SEMIÓTICA . Análise do discurso TG SEMÂNTICA .. Análise do discurso	TG LINGUISTIC TG LANGUAGE TG LANGUAGE AND LANGUAGES . Speech . Semantics ..Semiotics ..Discourse analysis TG LITERARY THEORY	TG LINGUÍSTICA . Linguagem .. Enunciação ... Instância enunciativa sujeitucional . Semântica .. Análise do discurso . Semiótica .. Análise do discurso . Signo linguístico .. Significado .. Significante . Gramática .. Oração ... Termos da oração Sujeito TG TEORIA LITERÁRIA . Gênêros literários ..Oratória ... Discurso

O grau de profundidade temática alcançado pelos termos identificados alcançou o nível 4 na hierarquia com o termo "sujeito", selecionado pelos 6 indexadores, considerando-se as variações do termo "sujeito discursivo", "Manifestações - sujeito", "Instâncias enunciativas - sujeito" e "Formas - sujeito". Nota-se o uso do hífen na composição desses termos que, infere-se, seja devido à familiaridade dos indexadores com linguagens de indexação pré-coordenadas que compõem dois ou mais termos ligados por hífen para fazer um cabeçalho de assunto, o que consideramos uma comprovação de que os indexadores têm familiaridade com vocabulário de linguagem de indexação e a utilizam como conhecimento prévio.

Além desses termos compostos com hífen outros termos identificados pelos indexadores não estão representados nas linguagens de indexação tais como: "Interdiscursividade de interpretação", "Memória discursiva", "Processo enunciativo", "Processo interpretativo", "Sentido", "Significações" e "Sujeitucional". De certa forma, são termos ligados aos significados dos termos da hierarquia, como por exemplo, "Processo enunciativo" que se infere seja ligado à "Enunciação", assim como "Sentido" e "Significações" pressupõe-se que tenham ligação com "Significado". Portanto, pode-se considerar que houve um esforço dos indexadores em garantir uma profundidade temática com a identificação de vários termos interligados aos termos mais representativos com a intenção de oferecer mais opções de acesso na recuperação, embora esses termos não existam nas linguagens de indexação. Isso denota que durante a análise de assunto o indexador orienta-se, também, pelo conteúdo do texto para a identificação de termos.

A análise de profundidade temática dos três artigos apresentou diferenças nos graus hierárquicos alcançados pelos termos possivelmente relacionados às características das áreas de assunto e suas estruturas textuais, porque se observou que no artigo da área de Astronomia obteve-se quantidade

combinada à identificação com maior grau de profundidade pela especificidade dos termos. Trata-se de um texto com vocabulário mais técnico e preciso sem variações sinonímicas cujos termos identificados foram quase todos encontrados nas linguagens de indexação o que demonstra alta compatibilidade e precisão de vocabulário.

Os dois outros artigos possuem estrutura textual diversa da do primeiro artigo e vocabulário com variação sinonímica que dificultam o trabalho do indexador na identificação de termos precisos e compatíveis com as linguagens de indexação.

Com a análise da profundidade temática constata-se que é possível realizar uma avaliação do processo de identificação dos termos durante a análise de assunto do conteúdo documental e demonstrar aos indexadores vários aspectos importantes:

- a familiaridade com o vocabulário da área de assunto é fator influente na compreensão do texto para fins de representação temática da informação;
- auxilia o indexador na identificação dos termos mais representativos de cada conteúdo o que confirma, por um lado, a necessidade de experiência de indexadores na análise de assunto de vários documentos de diferentes áreas e, por outro lado, o conhecimento em profundidade das linguagens de indexação; e,
- propicia, além do domínio de vocabulário mais preciso e representativo, também a compreensão da macroestrutura lógico-hierárquica das áreas de assunto de tal modo que o indexador saberá ao identificar o termo, se é um termo genérico e, portanto, exaustivo, ou, se trata de termo mais específico que poderá aumentar o nível de profundidade temática de representação e de sua consequente recuperação do conteúdo textual.

CONSIDERAÇÕES FINAIS

A identificação de conceitos desenvolvida na representação de uma área do conhecimento necessita do domínio de vocabulários especializados no conhecimento prévio do indexador a fim de garantir profundidade temática na análise de assunto e precisão na compatibilização da linguagem de indexação.

Na linguagem, mesmo a científica, existe uma variedade de termos técnicos e científicos e com a produção, cada vez maior de pesquisas e estudos que necessitam ser disseminados, o mais rápido possível, por isso é de grande importância o controle de vocabulário visando a organização e a recuperação da informação.

Portanto, o indexador, mesmo com larga experiência e habilidade no ato de indexar necessita do conhecimento prévio de linguagens de indexação que os auxiliem no momento da identificação do termo durante a análise de assunto para garantir a representação do conteúdo textual na recuperação da informação.

Acreditamos que necessitamos de mais estudos que apontem soluções importantes para a leitura durante o processo de indexação. Para tanto, necessitamos realizar pesquisas que possibilitem a compreensão dos procedimentos de análise de assunto durante a leitura, ou seja, a identificação e a seleção de conceitos de uma unidade informacional. Pretendemos que a nossa investigação seja uma aliada no

apoio e melhor compreensão ao processo de análise de assunto, viabilizando novos caminhos para sua efetivação.

REFERÊNCIAS BIBLIOGRÁFICAS

Associação Brasileira de Normas Técnicas (1992). *NBR 12676: Métodos para análise de documentos - determinação de seus assuntos e seleção de termos de indexação*. Rio de Janeiro. 4 p.

Alvarenga, Lídia (1998). Bibliometria e arqueologia do saber de Michel Foucault – traços de identidade teórico-metodológica. *Ciência da Informação*, Brasília, v. 27, n.3, set./dez.

Ausubel, D. P (1982). *A aprendizagem significativa: a teoria de David Ausubel*. São Paulo: Moraes.

_____; Novak, J.D.; Hanesian, H. (1980). *Psicologia educacional*. Rio de Janeiro: Interamericana.

Bertrand, Annick; Cellier, Jean Marie (1995). Psychological approach to indexing: effects of the operator's expertise upon indexing behaviour. *Journal of Information Science*, v.21, n.6, p. 459-472.

Brown, S. M. (1992). Cognitive mapping and repertory grids for qualitative survey research: some comparative observations. *Journal of Management Studies*, v. 29, n. 3, p. 287-308.

Chu, Clara M., O'brien, Ana (1993). Subject analysis: the crucial first stage in indexing. *Journal of Information Science*, n. 6, v. 19, p. 439-454.

Cleveland, D. B., Cleveland, A. D. (2001). *Introduction to indexing and abstracting*. 3.ed. Greenwood Village, Co.: Libraries Unlimited. 283p.

Cossette, P.; Audet, M. (1992). Mapping of an idiosyncratic schema. *Journal of Management Studies*, v. 29, n. 3, p. 325-348.

Eden, C. (1988). Cognitive mapping. *European Journal of Operational Research*, n. 36, p. 1-13.

Farrow. John F. (1995). All in the mind: concept analysis in indexing. *The Indexer*, v.19, n. 4, p. 243-247, Oct.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. *American Psychologist*, v. 34, 906-911.

Fujita, Mariângela S. (2004). A Leitura Documentária na perspectiva de suas variáveis: leitor-texto-contexto. *DataGramZero, Revista de Ciência da Informação*, v.5 n.4 ago.

_____. (2003). *Revista Digital de Biblioteconomia e Ciência da Informação*, Campinas, v .1, n. 1, p. 60-90, jul./dez.

Fujita, M. S. L. (2012). A política de indexação para representação e recuperação da informação. In: Gil leiva, I., Fujita, M.S.L. *Política de indexação*. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica.

Gil Leiva, I. (2008). *Manual de indización: teoría y práctica*. Gijón: Trea.

Grings, L. (2016). Controle de autoridades na Biblioteca Nacional do Brasil: breve histórico e práticas atuais. // Revista Brasileira de Biblioteconomia e Documentação, v.11,n.15, p.139-154, 2015.

Hjørland, B.; Albrechtsen, H. (1995). Toward a new horizon in information science: domain-analysis. Journal of the American Society for Information Science, v. 46, n. 6.

Lancaster, F. W. (2004). Indexação e resumos: teoria e prática. Tradução de Antonio Agenor Briquet de Lemos. 2. ed. Brasília: Briquet de Lemos. 452 p.

Moreiro Gonzalez, J. et al. (2004). De los tesauros a los topicmaps: nuevo estandar para la representación y la organización de la información. Encuentros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, n. 18, p. 1-19, 2º sem.

Neves, Dulce Amélia B. (2012). Representação temática da informação e mapas cognitivos: interações possíveis. Informação & Sociedade, v. 22, p. 39 - 47.

_____. Martins, Fernanda (2011b). Mapas conceituais em pesquisa sobre representação da informação. Páginas a&b: Arquivos e Bibliotecas, Porto/Portugal. n. 2-7, p. 73 – 89.

_____. (2011). Mapeamento cognitivo de la identificación y selección de términos durante el análisis de asunto del proceso de indexación. In: V Encuentro Ibérico EDICIC, 2011, Badajoz. Actas V Encuentro Ibérico EDICIC. Badajoz: Abecedario. p. 499-457.

Neves, Dulce Amélia B. (2004). Aspectos metacognitivos na leitura do indexador. Belo Horizonte: UFMG.. Tese de Doutorado Escola de Ciência da Informação, UFMG.

Pinto Molina, Maria (1994). Interdisciplinary approaches to the concept and practice of written text documentary content text. Journal of Documentation, v. 50, n. 2, p. 111-133, Jun.

_____. (1993). Analisis documental: fundamentos y procedimientos. 2 ed. rev.yaum.Madrid: EUEDEMA. cap. 7, 9, 10, 11.

Ribeiro, Fernanda (2010). O Papel mediador da Ciência da Informação na construção da sociedade em rede. Informação & Sociedade: Estudos, João Pessoa, v.20, n.1, p. 63-70, jan./abr.

Santos, João Bosco C. (2008) A instância enunciativa sujeitidual. Uberlândia: LEP/GPAD/ILEEL/UFU.

Soergel, D. (1994). Indexing and retrieval performance: the logical evidence. Journal of the American Society for Information Science, v. 45, n. 8, p. 589-599.

Sternberg, Robert J. (2008). Psicologia cognitiva. Porto Alegre: Artmed.

Vigotski, L S. (1998). A formação social da mente. São Paulo: Martins Fontes.

_____. (2001). A construção do pensamento e da linguagem. São Paulo: Martins Fontes.

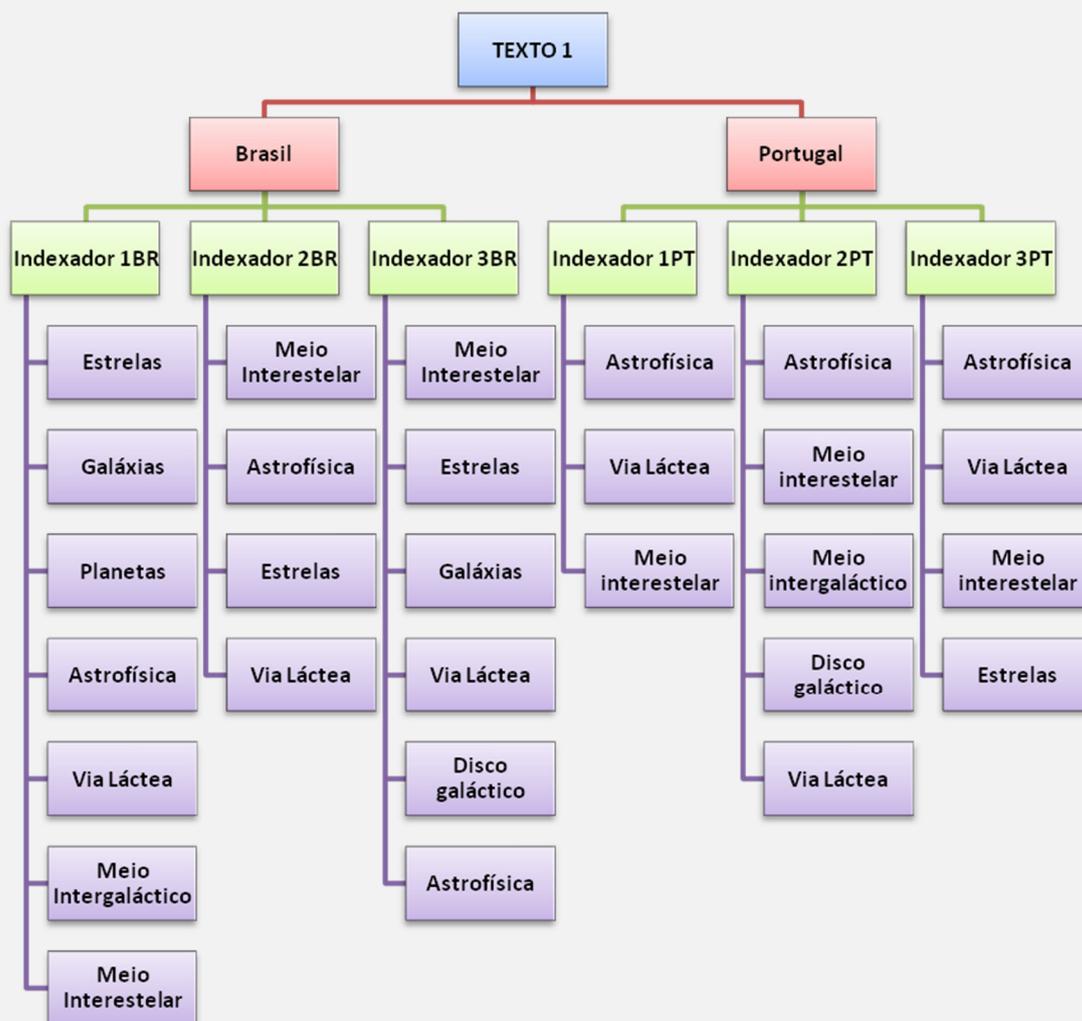
Tálamo, M. F. G. M. (1987). Elaboração de resumos. Escola de Comunicação e Artes. 14 f.

UNISIST (1981). Princípios de indexação (1981). Revista Escola Biblioteconomia UFMG, v.10, n.1. p.83-94, mar.

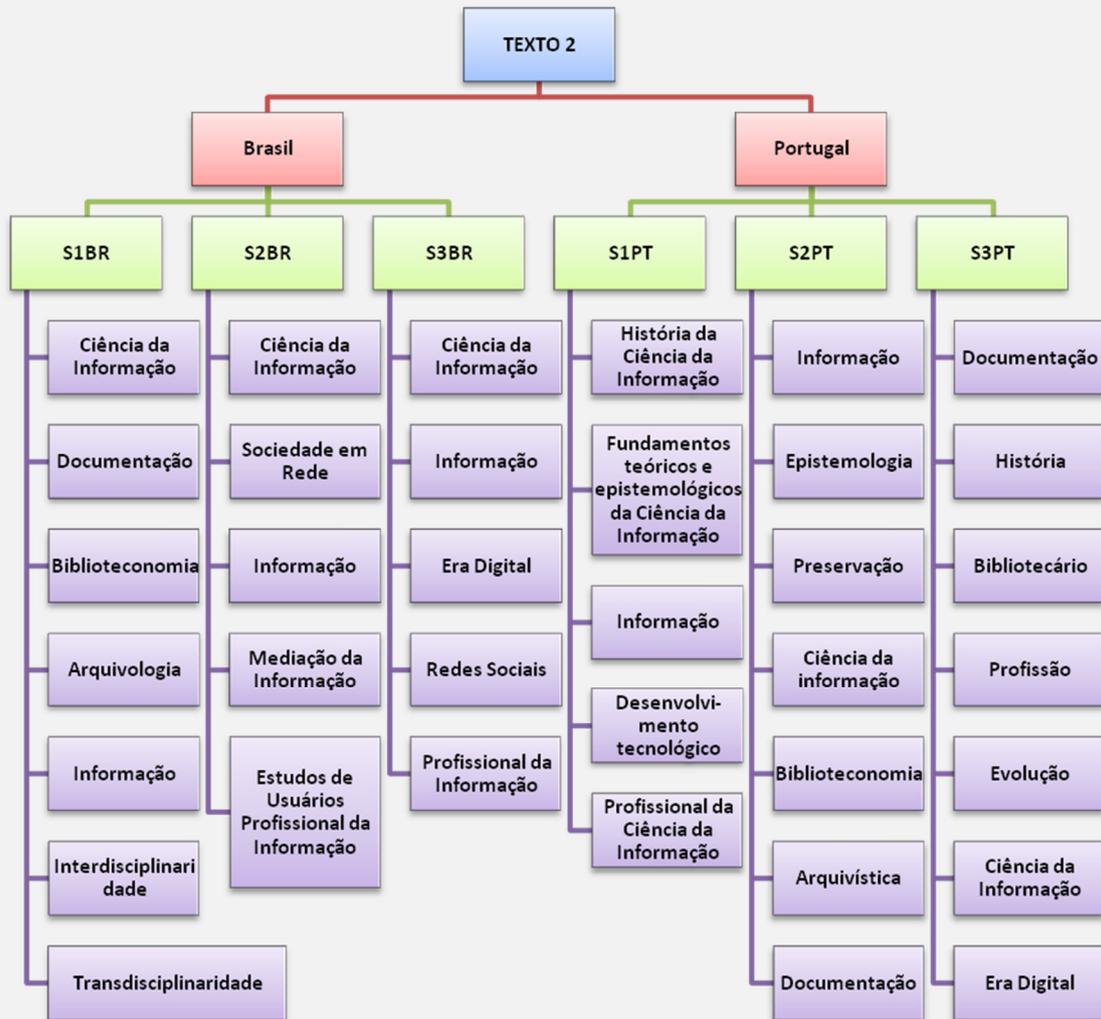
Universidade de São Paulo. (2015) Sistema Integrado de Bibliotecas. Vocabulário Controlado USP.
<http://www.sibi.usp.br/produtos/vocabulario-controlado-usp/> (2016-01-16).

APÊNDICES

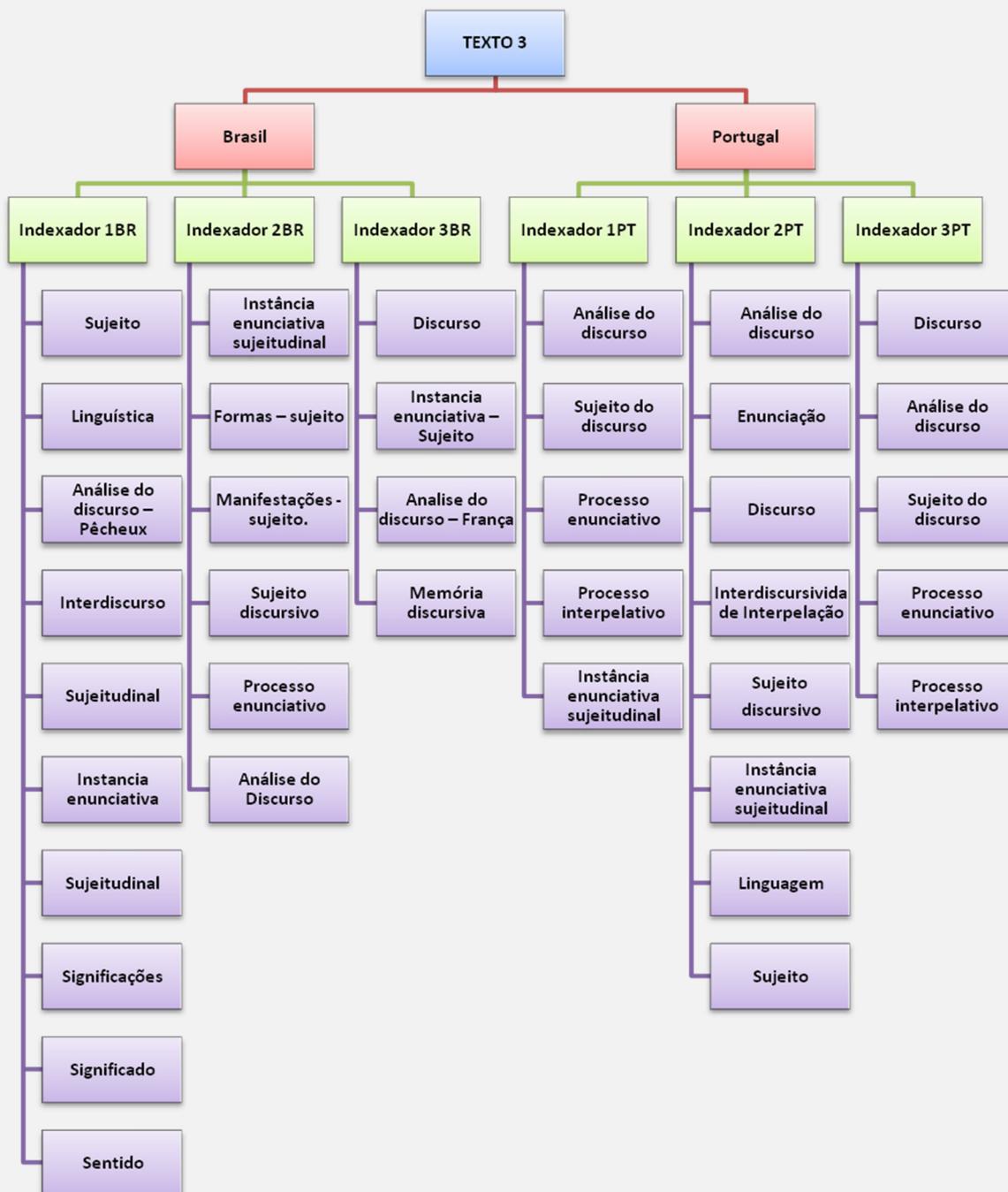
APÊNDICE I - MAPA 01 – LEITURA DO PRIMEIRO ARTIGO



APÊNDICE II - MAPA 02 – LEITURA DO SEGUNDO ARTIGO



APÊNDICE III - MAPA 03 – LEITURA DO TERCEIRO ARTIGO





TEXTO VERSUS IMAGENS? FOLKSONOMIAS E INDEXAÇÃO SOCIAL EM ARQUIVOS

Ana Margarida Dias da Silva¹, Leonor Calvão Borges²

¹Faculdade de Letras da Universidade de Coimbra, 0000-0003-1247-8346, anasilva@fl.uc.pt

²Faculdade de Letras da Universidade de Coimbra, 0000-0002-2316-9365, leonorcborges@gmail.com

RESUMO O objetivo deste trabalho é compreender se ou como é que os arquivos aproveitam as ferramentas colaborativas da Web 2.0 e a inteligência coletiva para a construção do conhecimento e, paralelamente, confrontar e analisar o recurso às *folksonomias* e à indexação social em documentação/informação textual e gráfica/imagens em arquivos. Usa-se a técnica da análise comparativa, dentro da abordagem da análise qualitativa dos vários projetos identificados, sedimentada na bibliografia científica, nacional e estrangeira, sobre o assunto. Constata-se uma escassa bibliografia sobre *folksonomias* em arquivos, e menor identificação de projetos/casos concretos. A colaboração na indexação de conteúdos é cada vez mais uma realidade, sobretudo porque permite um maior grau de exaustividade da descrição arquivística. Os utilizadores em ambiente participativo aderem massivamente ao pedido de informações sobre as imagens, permitindo melhorar a informação disponibilizada pelas instituições e aumentar a descrição dos seus acervos no catálogo institucional. No entanto, conclui-se que existe controlo quer dos internautas quer na criação dos pontos de acesso, e não a livre atribuição de etiquetas, no caso da documentação/informação textual.

PALAVRAS-CHAVE *Folksonomias, Indexação Social, Ciência da Informação, Representação da Informação, Descrição Arquivística*

ABSTRACT The main goal of this paper is to understand if or how the archives take advantage of the collaborative tools of Web 2.0 and of the collective intelligence for the construction of knowledge. At the same time, the paper aims to confront and analyze the use of folksonomies and social indexation in documentation / textual information and graphics / images archives. The technique of comparative analysis is used, within the approach of the qualitative analysis of the several identified projects, based in the scientific bibliography, national and foreign, on the subject. There is a scarce bibliography on folksonomies in archives, and less identification of concrete projects / cases. Collaboration in content indexing is increasingly a reality, especially since it allows for a greater degree of completeness of the archival description. The users in a participative environment massively adhere to the request for information about the images, allowing to improve the information made available by the institutions and to increase the description of their collections in the institutional catalog. However, it is concluded that there is control of both Internet users and the creation of access points, not the free allocation of labels, in the case of documentation / textual information.

KEYWORDS *Folksonomies, Social Indexation, Information Science, Information Representation, Archival Description*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A disponibilização massiva de conteúdos *on-line*, de caráter institucional ou através de partilha nas plataformas sociais, exigem novos modelos de indexação e recuperação da informação. É neste contexto que surgem as *folksonomias* e a indexação social, potenciadas pelo advento e generalização das ferramentas colaborativas da *web 2.0*, onde o utilizador passivo passou, simultaneamente, a consumidor e produtor de informação (Silva, 2013).

As *folksonomias*, neologismo atribuído a Thomas Vander Wal, parecem estabelecer-se como forma de resposta a essa necessidade de estruturação e recuperação da informação de conteúdos disponibilizados pela *web 2.0* feita pelos próprios utilizadores (Peters, 2009: 153; Caldas e Moreira, 2012).

As *folksonomias* caracterizam-se por “elementos como colaboração, interatividade, linguagem e sociabilidade em rede” onde se “explora, cada vez mais, a linguagem natural e a participação dos sujeitos informacionais” (De Assis & Moura, 2013: 86) e são “the result of personal free tagging of information and objects (anything with a URL) for one's own retrieval. The tagging is done in a social environment (usually shared and open to others). Folksonomy is created from the act of tagging by the person consuming the information.” (Vander Wal, 2007). Esta definição aponta, como características, essenciais, a motivação pessoal para posterior acesso e recuperação da informação, a realização em ambiente digital e social, e a produção de conteúdos pelo consumidor da informação.

Numa frase, Barros diz que “a folksonomia, [é uma] prática que permite aos usuários de websites classificar os conteúdos disponíveis na internet.” (Barros, 2011: 17). A que se pode acrescentar que “L’indexation du même site par plusieurs personnes profite de l’effet de masse pour constituer un tronc commun de mots clés qui sera une sorte de consensus acceptable par une majorité d’utilisateurs. Ce type d’indexation nous permet de parler de folksonomie.” (Francis & Quesnel, 2007: 60).

A indexação social refere-se “à ação de representação da informação, executada pelos próprios usuários do sistema, para fins de compartilhamento e recuperação de informação” (Guedes & Dias, 2004: 42).

Ponto comum a todas estas definições é a tríade utilizador, etiqueta (*tag*) e conteúdo. Veja-se o papel de cada um na construção do conceito de folksonomia.

O *folk* (povo, gente) passou a ser central como “sujeito social pragmático, uma vez que constroi suas relações pela via da linguagem e do compartilhamento de significados. Tal fenómeno marca a passagem de um usuário passivo em busca de recursos que atendam às suas necessidades de informação para um sujeito ativo e dinamizador dos fluxos informacionais. Essas alterações podem ser visualizadas e analisadas em ambientes em que ocorrem folksonomias.” (De Assis & Moura, 2013: 86). Como se vê “A peculiaridade da folksonomia está no fato dos usuários produzirem seus próprios descritores, utilizando a linguagem natural.” (Barros, 2011: 17). Sem a colaboração dos utilizadores e internautas não existem as folksonomias.

As *tags* ou etiquetas são “uma forma de indexação, em que as próprias pessoas, no caso os usuários da informação, classificam documentos e objetos informacionais” (Santos, 2013: 93) e “Ainda que receba diferentes nomenclaturas (indexação; catalogação de assunto; tagging; etiquetagem; etc.) o ato de representar um documento através de um conceito é de mesma natureza nos diferentes contextos – físico e digital.” (Guedes & Dias, 2004: 42). A atribuição de etiquetas a conteúdos Web é feita com termos escolhidos pelo utilizador comum (Gracioso, 2010: 140) e “É uma forma de indexação livre, em

linguagem natural onde não há controle de vocabulários, regras ou políticas de indexação.” (Barros, 2011: 17). As etiquetas podem ser utilizadas por um ou mais utilizadores e “Es precisamente esta dimensión social la que aporta mayor utilidad a las folksonomías en el área de la recuperación de información.” (Yedid, 2013: 16). De facto, a atribuição de etiquetas aos conteúdos e a construção de pontos de acesso por um conjunto alargado de participantes em ambiente visa a recuperação da informação e um dos argumentos “para utilizar folksonomías es que resultan “mejor que nada”, puesto que introducen un elemento para mejorar la capacidad de recuperación en un contexto en el que no es viable la aplicación de lenguajes controlados”(Yunta, 2009: 834).

As folksonomias surgem, então, como “sistemas de classificação distribuídos, criados por usuários individuais” (Santos, 2013: 96) e embora se distinga das taxonomias, ambas cumprem a função de descrição de conteúdos (Gracioso, 2010: 151-152).

Assim, da súmula dos trabalhos lidos retiramos, como sentido do termo folksonomia e indexação social, a utilização de dispositivos e aplicações da Web 2.0 que permitem ao utilizador colocar *tags*, retiradas do próprio vocabulário e da linguagem natural, que vão formar um conjunto de pontos de acesso que descrevem um elemento de informação, trabalho que é realizado pela comunidade de internautas, sob diferentes motivações mas que têm sempre como objetivo final a recuperação da informação.

Apesar de existirem diferentes designações para o conceito acima definido, o conceito de *folksonomia* parece ser o termo mais utilizado em diversos domínios científicos (Barros, 2011: 21; Guedes & Dias, 2004: 39; Lasić-Lazić et al., 2014: 685; Rapetti, 2007: 19-20), em detrimento de indexação social, por exemplo.

Os principais motivos que levam os investigadores a estudar as folksonomias “visam compreender essa modalidade de organização da informação em ambientes digitais e como essa configuração, que integra redes de conceitos, pessoas e conteúdos, pode contribuir para o desenvolvimento de metodologias e instrumentos de representação e recuperação da informação nos mais diversos contextos.” (De Assis & Moura, 2013: 87). É também nosso objetivo, com este trabalho, compreender se ou como é que os arquivos aproveitam as ferramentas colaborativas da Web 2.0 e a inteligência coletiva para a construção do conhecimento e, paralelamente, confrontar e analisar o recurso às *folksonomias* e à indexação social em documentação/informação textual e gráfica/imagens em arquivos.

Este trabalho exploratório procura ser um contributo para a identificação e conhecimento de uma realidade pouco estudada: as *folksonomias* em arquivos.

METODOLOGIA

Para realizar o trabalho procedeu-se à recolha de bibliografia científica nacional e estrangeira sobre os conceitos de folksonomias e indexação social, indexação de imagens, a criação de pontos de acesso em documentação/informação textual, e sobre o estabelecimento de novas conexões com os utilizadores das instituições de memória através da *web 2.0*.

Esta recolha permitiu caracterizar aqueles conceitos e, simultaneamente, identificar e recolher exemplos de trabalhos científicos sobre esta temática, necessários (também) para a discussão dos resultados.

A revisão da literatura permitiu-nos uma primeira conclusão a favor dos trabalhos que abordam a aplicação de folksonomias em arquivos de imagens das instituições de memória, onde a utilização do Flickr é já uma realidade vastamente documentada. O mesmo não se verifica em relação à aplicação das folksonomias em documentos textuais arquivísticos, uma vez que, muito embora exista já vasta literatura científica sobre o assunto, esta se insere, sobretudo, na área da Ciência da Informação e da Biblioteconomia, sendo escassos os trabalhos que abordam as folksonomias em arquivos. A mesma observação foi feita por Lasić-Lazić *et al.*: “As far as different communities goes, the new user centred approach in organizing knowledge produced a number of studies from the field of libraries and museums, where folksonomies are examined as a tool to enhance access to digitized collections and library catalogues. On the other hand, lack of research connected with archives could be noted, where folksonomies weren’t recognized as a viable approach.” (Lasić-Lazić *et al.*, 2014: 690).

De seguida, usaram-se duas linhas de análise, geograficamente circunscritas à Europa e América do Norte, a saber:

- 1) Um estudo exploratório em arquivos públicos, com recolha de dados *on-line* que permitisse saber, na prática, quais as características e especificidades das *folksonomias* e da indexação social em arquivos textuais;
- 2) Um estudo de caso a partir da plataforma *Flickr*, e dentro dela do projeto *Flickr Commons*, iniciado pela Biblioteca do Congresso em 2007, em parceria com a própria equipa de desenvolvimento do *Flickr*, com o objetivo de compreender o contributo que as *folksonomias* trazem para a identificação de imagens.

Em ambos os casos, identificam-se as instituições promotoras, os seus projetos e as estratégias de comunicação de conteúdos utilizadas. Usa-se a técnica da análise comparativa, dentro da abordagem da análise qualitativa dos vários projetos identificados, sedimentada na bibliografia científica, nacional e estrangeira, sobre o assunto.

RESULTADOS

As pesquisas realizadas *on-line*, executadas com o objetivo de localizar projetos colaborativos em arquivos textuais, não se verificaram muito frutíferas. Consta-se uma escassa bibliografia sobre *folksonomias* em arquivos, e menor identificação de projetos/casos concretos. O maior número de exemplos refere-se à utilização do *Flickr* para a identificação e descrição de fotografias, como veremos abaixo.

Paradigmático e exceção à regra para ser o caso francês onde foi possível identificar inúmeros projetos colaborativos em arquivos departamentais públicos franceses (Moirez, 2012; Bouyé, 2012; Silva, 2017). A França surge como um país europeu de democratização no acesso e disponibilização de arquivos *on-line*, característica que vem desde a Revolução Francesa e a criação dos *Archives Nationales* (que visavam a concentração dos arquivos e o acesso generalizado a todo e qualquer cidadão), ao mesmo tempo que promove e estimula a participação dos cidadãos no enriquecimento das descrições dos seus arquivos.

Entre os arquivos departamentais franceses, Pauline Moirez identificou cerca de 20 projetos de indexação colaborativa (Moirez, 2012: 188) e Edouard Bouyé indicou a intenção da realização de 16 novos projetos de indexação colaborativa para o ano 2012 (Bouyé, 2012: 9). No seu estudo, Édouard

Bouyé divide os arquivos em três grupos segundo a forma de participação dos utilizadores na indexação: aqueles que dão liberdade total ao internauta, sem necessidade de inscrição (ex: Cantal, Corrèze, Eure-et-Loir, Loire-Atlantique, Nièvre, Var); aqueles que dão liberdade total ao internauta, após uma inscrição obrigatória e abertura de conta na página do arquivo (ex: Martinique, Hautes-Alpes, Rhône, Yvelines); e, um terceiro grupo, onde o internauta tem um enquadramento que exige inscrição, passar em testes de leitura paleográfica e trabalhar em documentação previamente escolhida pelos serviços de arquivo (ex: Rennes, Aube, Mayenne, Puy-de-Dôme) (Bouyé, 2012: 3).

As características dos documentos de arquivo, sobretudo o seu carácter único, condicionam o seu tratamento, e não possibilitam a troca de informação entre instituições (Moirez, 2012: 190). As tipologias documentais mais disponibilizadas são, sobretudo, registos paroquiais (batismo ou nascimento, casamentos e óbitos), matrículas de militares, recenseamentos, em suma, documentação com nomes de pessoas e de lugares. Aquilo que se pede ao utilizador é que crie um ponto de acesso unívoco (nome, apelido, local, idade, etc.), uma representação fiel do conteúdo informacional, não havendo lugar para a ambiguidade, a homonímia, a sinonímia e a polissemia (Yedid, 2013: 18), tantas vezes assinaladas pelos autores como desvantagens das folksonomias. Os pontos de acesso criados nestes arquivos aproximam-se do conceito de Quintarelli de “etiquetas estreitas” que “são caracterizadas por poucas pessoas utilizando uma ou mais etiquetas, o que significa compartilhar vocabulários próprios e não termos tão populares como nas folksonomias largas.” (Barros, 2011: 26).

Não se pretende que o indexador-colaborador-utilizador atribua um termo da linguagem natural, mas sim procura-se que leia e transcreva, com precisão, os termos a descrever, havendo a possibilidade de correção por parte do organismo ou de outros utilizadores (ex: arquivos do Oise). O resultado final do conjunto dos pontos de acesso criados são índices alfabéticos (onomásticos e/ou geográficos), o que coloca este trabalho ao nível da representação da informação e não da representação do conhecimento (Brascher & Café, 2008).

A maior parte da documentação custodiada em arquivos é descrita ao nível da série, pois o grau de exaustividade e especificidade necessários para a descrição ao nível da peça não se coadunam nem com as capacidades humanas nem com as possibilidades financeiras das instituições, pela morosidade que isso implicaria. Isso provoca uma quantidade significativa de massa documental não tratada e “C’est précisément pour répondre à cette demande que des opérations de crowdsourcing sont mises en place : identification, description et indexation au niveau de la pièce, voire transcription des contenus pour permettre une recherche en plein texte.” (Moirez, 2012: 190).

Algumas das características das *folksonomias*, como a ambiguidade, homonímia, sinonímia e polissemia não se verificam nos projetos de indexação em arquivos textuais visto que aquilo que é pedido é uma representação fiel e exata do conteúdo. Os pontos de acesso gerados pelos utilizadores têm de ser unívocos porque são nomes (de pessoas, de lugares, de atos), e existe uma forte componente paleográfica e de leitura de manuscritos, que não se encontra em outros suportes (fotografias, mapas, livros, música, vídeo) e, portanto, menos abrangente no número de pessoas que podem contribuir.

Já no caso das imagens, o principal contributo tem sido a quantidade de informação especializada sobre as mesmas, que permite melhorar significativamente as próprias descrições, com o aparecimento de *power taggers* (utilizadores que contribuem com cerca de 3 000 *tags* (Springer, *et al.*, 2009, p. 19) ou a quantidade de *links* para artigos especializados sobre as fotografias, que as instituições acrescentam ao seu catálogo.

É ainda de destacar o número significativo de utilizadores que estes projetos atingem e a sua capacidade de crescimento, cujo perfil é bastante diferenciado do do público tradicional das instituições de memória, como referem diferentes estudos (Almarcha & Fernández & Villena, 2014; Caldas, 2012; Chan, 2010; Colquhoun, 2013; Edmunson-Morton, 2009; Kalfatovic, 2009; Rafferty & Hilderley, 2007; Rodrigues, 2010; Rorissa, 2010; Springer, 2010; Willey, 2011).

O projeto *Flickr Commons* surge no âmbito de um novo entendimento da missão dos arquivos e bibliotecas como colaborativos ou participativos: “Help us catalog the world’s public photo archives”, onde se interage com e pelo público (Theimer, 2011 e 2014).

A escolha de fotografias ou coleções para partilhar também se revelou uma decisão crítica de sucesso, com opção maioritária de proporcionar o acesso via álbuns temáticos e a existência de conjuntos de fotografias mais populares que outros. As instituições foram assim obrigadas a repensar as suas estratégias de divulgação e a desenvolver projetos de fidelização (novos álbuns, apelo direto ao público, exposição das fotografias mais comentadas).

As fotografias, têm a vantagem de ser interessantes para uma vasta e variada audiência que pode fornecer informação útil.

CONCLUSÃO

Tomando como ponto de partida a visão simplista de que as *folksonomias* são a conjugação de utilizador, etiqueta e conteúdo, então pode-se dizer que os projetos analisados em arquivos cumprem esta tríplice observação: é feito um pedido expresso e claro à participação dos utilizadores internautas para auxiliar na tarefa de atribuição de pontos de acesso (etiquetas) ao conteúdo dos objetos digitais disponibilizados para posterior facilidade na recuperação da informação. Neste ponto, os arquivos não se distanciam muito daquilo que é feito nas bibliotecas ou nos museus, por exemplo. No entanto, se nestes domínios culturais, nomeadamente no caso dos museus, a utilização da Web visa o aumento dos utilizadores e visitantes presenciais, os arquivos centram a sua atenção no valor do documento sendo “possible de mettre en place une véritable dissémination massive des fonds d’archives, et d’initier un travail scientifique collaboratif sur celles-ci, à distance” (Moirez, 2012: 191). De facto, aquilo que se verifica é que os arquivos utilizam uma abordagem diferente quer na estratégia de representação dos conteúdos quer na relação com os utilizadores, e é isso que se conclui se se fizer o contraponto entre o conceito de *folksonomia* e a realidade dos projetos acima analisados.

A chamada à participação na indexação dos conteúdos justifica-se, por um lado, pelo volume dos fundos por tratar e, por outro, pela morosidade do processo, que se tornará mais rápido com a ajuda de todos. No entanto, a ausência de um maior número de projetos de indexação ou de descrição colaborativa parece estar relacionada com a falta confiança dos profissionais nas capacidades dos internautas na criação de pontos de acesso. É por isso que diversos arquivos fazem um apelo direcionado a um grupo muito específico, aos genealogistas (os arquivos de Niève chamam-lhes “généanautas”) que, enquanto utilizadores “egocêntricos” (Canãda, 2006), querem saber mais sobre as suas origens e a sua família, e acabam por partilhar informação com todos. Acaba por ser, igualmente, um apelo pessoal pois cada utilizador querera aceder à informação relativa aos seus antepassados e caso cada um contribua com essa identificação a informação ficará acessível para toda a comunidade de investigadores, na perspetiva

de “Cohen [que] acredita que os indivíduos só compartilham se receberem algo em troca.” (Santos, 2013: 99).

As características dos documentos arquivos, únicos e com limitações de acesso e comunicabilidade devido a questões legais, também se manifestam na especificidade dos projetos analisados. A extração de conteúdos para indexação por assuntos pode originar um número ilimitado de etiquetas ou pontos de acesso para um mesmo objeto informacional. No caso dos documentos textuais não é esse o objetivo, sobretudo naqueles disponibilizados em projetos colaborativos de *folksonomias* e indexação social (registros paroquiais, registos civis, listas de recenseamentos da população, entre outros) porque aqui só interessa um ponto de acesso, único e exato: o nome da pessoa, o local de nascimento, a sua idade, por exemplo. É essa a especificidade dos documentos textuais relativamente às imagens (e a outro tipo de informação): não se pretende um máximo de etiquetas mas sim a máxima precisão no ponto de acesso.

Por outro lado, a análise de resultados publicados e estudos académicos permite-nos constatar que a adesão a plataformas sociais de disponibilização de fotografias tem tido reflexos muito positivos e objetivos na estratégia de comunicação e difusão das instituições de memória, revelando uma tendência na disponibilização deste tipo de acervos, um pouco por todo o mundo ocidental.

A existência de projetos de disponibilização de imagens em movimento com objetivos similares e idêntica participação leva-nos a concluir pela adesão e utilização privilegiada da *web 2.0* e uso de *folksonomias* em arquivos de imagens (fixas ou em movimento).

O *Flickr Commons*, enquanto projeto agregador de conteúdos dessas mesmas instituições tem aumentado significativamente o número de participantes, numa estratégia que une arquivos, bibliotecas, museus e sociedades históricas, de uma forma geral.

Os estudos dedicados à análise de etiquetas referem que a grande maioria tem utilidade prática para a instituição, permitindo aumentar a descrição dos seus acervos no catálogo institucional. A sugestão de que as *folksonomias*, apesar das suas limitações, podem ser utilizadas como complemento da indexação tradicional, parece ser comumente aceite.

A colaboração na indexação de conteúdos é cada vez mais uma realidade, sobretudo porque permite um maior grau de exaustividade da descrição arquivística. Os utilizadores em ambiente participativo aderem massivamente ao pedido de informações sobre as imagens, permitindo melhorar a informação disponibilizada pelas instituições e aumentar a descrição dos seus acervos no catálogo institucional. No entanto, conclui-se que existe controlo quer dos internautas quer na criação dos pontos de acesso, e não a livre atribuição de etiquetas, no caso da documentação/informação textual.

REFERÊNCIAS BIBLIOGRÁFICAS

Almarcha, E. & Fernández, O., & Villena, R. (2014). La utilización de Flickr para la diffusion de colecciones fotográficas institucionales. In *Girona: Archivos e Industrias Culturales*.

Barros, L. M. de S. (2011). *A Folksonomia como prática de classificação colaborativa para a recuperação da informação*. Universidade Federal do Rio de Janeiro. Retrieved from <http://ridi.ibict.br/handle/123456789/737>

- Bouyé, É. (2012). Le web collaboratif dans les services d'archives publics : un pari sur l'intelligence et la motivation des publics. In *Rencontre annuelle des services d'archives départementales (RASAD)*, (pp. 1–12). Bordeaux. Retrieved from http://archives.cantal.fr/download.cgi?filename=accounts/mnesys_ad15/datas/cms/Bouye_RASAD_Gazette.pdf
- Bray, P., et al. (2011). Rethinking Evaluation Metrics in Light of Flickr Commons. In J. Trant and D. Bearman (eds). *Museums and the Web 2011: Proceedings*. Toronto: Archives & Museum Informatics. Retrieved from http://www.museumsandtheweb.com/mw2011/papers/rethinking_evaluation_metrics_in_light_of_flic
- Brascher, M., & Café, L. (2008). Organização da Informação ou Organização do Conhecimento? In *IX ENANCIB* (pp. 1–14). São Paulo. Retrieved from [http://skat.ihmc.us/rid=1KR7TM7S9-S3HDKP-5STP/BRASCHER_CAF%C3%89\(2008\)-1835.pdf](http://skat.ihmc.us/rid=1KR7TM7S9-S3HDKP-5STP/BRASCHER_CAF%C3%89(2008)-1835.pdf)
- Caldas, W. F. & Moreira, M. P. (2012). *Folksonomia e classificação de etiquetas: estudo de caso Flickr*. Retrieved from: <http://repositorios.questoesemrede.uff.br/repositorios/handle/123456789/62>
- Cañada, J. (2006) *Tipologias y estilos en el etiquetado social*. Retrieved from <http://www.terremoto.net/tipologias-y-estilos-en-el-etiquetado-social/>
- Chan, S. (2010). *Why Flickr Commons? (and why Wikimedia Commons is very different)*. *Fresh + New(er): discussion of issues around digital media and museums*. Retrieved from: <http://www.freshandnew.org/2010/01/why-flickr-commons-and-why-wikimedia-commons-is-very-different/>
- Colquhoun, B. (2013). Making Sense of Historic Photographic Collections on Flickr The Commons: Institutional and User Perspectives. In *Museums and the Web 2013*, N. Proctor, N. & R. Cherry (eds). Silver Spring, MD: Museums and the Web. Retrieved from: <http://mw2013.museumsandtheweb.com/paper/making-sense-of-historic-photographic-collections-on-flickr-the-commons-institutional-and-user-perspectives/>
- De Assis, J., & Moura, M. A. (2013). Folksonomia: a linguagem das tags. *Encontros Bibli: Revista Eletrônica de Biblioteconomia E Ciência Da Informação*, 18(36), 85–106. <http://doi.org/10.5007/1518-2924.2013v18n36p85>
- Edmunson-Morton, T. (2009). Talking and Tagging: Using CONTENTdm and Flickr in The Oregon State University Archives. Retrieved from: <http://interactivearchivist.archivists.org/case-studies/flickr-at-osu/>
- Francis, E., & Quesnel, O. (2007). Indexation Collaborative et folksonomies. *Documentaliste-Sciences de L'information*, 44(1), 58–63. Retrieved from <https://www.cairn.info/revue-documentaliste-sciences-de-l-information-2007-1-page-58.htm>
- Freixa-Font, P. (2011). Patrimonio fotográfico y web 2.0: la experiencia Flickr The Commons. *El profesional de la información*, 20 (4), 432-438. doi: 10.3145/epi.2011.jul.10
- Gracioso, L. D. S. (2010). Parâmetros teóricos para elaboração de instrumentos pragmáticos de representação e organização da informação na Web: considerações preliminares sobre uma possível

proposta metodológica. *InCID: R. Ci. Inf. E Doc*, 1(1), 138–158. Retrieved from <http://www.revistas.usp.br/incid/article/view/42310>

Guedes, R. D. M., & Dias, E. J. W. D. (2004). INDEXAÇÃO SOCIAL: ABORDAGEM CONCEITUAL. *Word Journal Of The International Linguistic Association*, 15(1), 39–53.

Guimarães, R. C. (2012). *Uso da folksonomia e da etiquetagem na indexação de imagens* (Dissertação de Doutorado não publicada). Faculdade de Ciência da Informação da Universidade de Brasília. Retrieved from: http://bdm.unb.br/bitstream/10483/4268/1/2012_RafaelCostaGuimaraes.pdf

Lasić-Lazić, J., Špiranec, S., & Ivanjko, T. (2014). Tag-Resource-User: A Review of Approaches in Studying Folksonomies. *Qualitative and Quantitative Methods in Libraries (QQML)*, 3, 683–692. Retrieved from http://www.qqml.net/papers/September_2014_Issue/3312QQML_Journal_2014_LasicLazicSpiranecIvanjko_Sept_683-692.pdf

Kalfatovic, M. R., *et al.* (2009). Smithsonian Team Flickr: a library, archives, 3 and museums collaboration in web 2.0 space. Retrieved from: <https://siarchives.si.edu/sites/default/files/pdfs/Kalfatovic%20et%20al%20Flickr%20Paper%20Author%20Proofs.pdf>

Moirez, P. (2012). Archives participatives. *Bibliothèques 2.0 à L'heure Des Médias Sociaux*, 187–197. Retrieved from http://archivesic.ccsd.cnrs.fr/sic_00725420

Peters, I. (2009). *Folksonomies: indexing and retrieval in Web 2.0*. Berlin: De Gruyter. Quintarelli, E. (2005) Folksonomies: power to the people. In: INCONTRO ISKO ITALIA – UNIMIB, Milão, 2005 Papers. Milan: Università di Milano. Disponível em: <http://www.iskoi.org/doc/folksonomies.htm>

Rapetti, L. (2007). *Folksonomia: Uso e Organização da Informação na Web*. Universidade Federal do Rio Grande do Sul. Retrieved from <https://www.lume.ufrgs.br/bitstream/handle/10183/18728/000667018.pdf?sequence=1>

Rafferty, P. & Hilderley, R. (2007). Flickr and Democratic Indexing: dialogic approaches to indexing In *Aslib Proceedings: New Information Perspectives*, 59 (4/5), 397-410.

Ribeiro, F. (1996). *Indexação e Controlo de Autoridade em Arquivos*. (C. M. do Porto, Ed.). Porto. Retrieved from <http://hdl.handle.net/10216/10721>

Rodrigues, A. A.A. (2010). *Folksonomia: Análise de etiquetagem de imagens no Flickr* (Tese de Mestrado não publicada). Escola de Ciência da Informação da Universidade Federal de Minas Gerais. Retrieved from: http://www.bibliotecadigital.ufmg.br/dspace/bitstream/handle/1843/ECID-88EJT5/folksonomia_an_lise_de_etiquetagem_de_imagens_no_flicr_dis.pdf?sequence=1

Rorissa, A. (2010). A comparative study of Flickr tags and index terms in a general image collection. *J. Am. Soc. Inf. Sci.*, 61, 2230–2242. doi: 10.1002/asi.21401.

Samouelian, M. E. (2009). Embracing Web 2.0: Archives and the Newest Generation of Web Applications. *The American Archivist*, 72 (Spring/Summer).

Santos, H. P. (2013). Etiquetagem e folksonomia: o usuário e sua motivação para organizar e compartilhar informação na Web 2.0. *Perspectivas Em Ciência Da Informação*, 18(2), 91–104. <http://doi.org/10.1590/S1413-99362013000200007>

Shatford-Layne, S. (1994). Some Issues in the Indexing of Images. In *Journal of the American Society for Information Science*, 45 (8), 583-588.

Shatford-Layne, S. (2002). Subject Access to Art Images. In Baca, M. (Ed.), *Introduction to Art Image Access: Issues, Tools, Standards, and Strategies*. NY: Getty Publications. Retrieved from: http://www.getty.edu/research/publications/electronic_publications/intro_aia/

Silva, A. M. D. da (2017). “Folksonomies in archives: controlled collaboration for specific documents” [Em linha] *Ariadne* 77 (Junho). Retrieved from <http://www.ariadne.ac.uk/issue77/margaridadiasdasilva>

Silva, A. M. D. da. (2013). *O Uso da Internet e da Web 2.0 na difusão e acesso à informação arquivística: o caso dos arquivos municipais portugueses*. Universidade Nova de Lisboa. Retrieved from <http://run.unl.pt/handle/10362/12014>

Springer, M., et al. (2009). *For the Common Good: The Library of Congress Flickr Pilot Project*. Library of Congress. Retrieved from: http://www.loc.gov/rr/print/flickr_report_final.pdf

Strehl, L. (2011). As folksonomias entre os conceitos e os pontos de acesso : as funções de descritores , citações e marcadores nos sistemas de recuperação da informação. *Perspectivas Em Ciência Da Informação*, 16(2), 101–114. Retrieved from <http://www.scielo.br/pdf/pci/v16n2/07.pdf>

Theimer, K. (Ed.) (2011). *A different kind of web: new connections between archives and our users*. Chicago: Society of American Archivists.

Theimer, K. (2014). The Future of Archives is Participatory: Archives as Platform, or A New Mission for Archives. Retrieved from: <http://www.archivesnext.com/?p=3700>

Vander Wall, T. (2007). Folksonomy. Retrieved from:<http://vanderwal.net/folksonomy.html>

Willey, E. (2011). A cautious partnership: The growing acceptance of folksonomy as a complement to indexing digital images and catalogs. *Faculty and Staff Publications – Milner Library*. Paper 57.

Yedid, N. (2013). Introducción a las folksonomías: definición, características y diferencias con los modelos tradicionales de indización [Introduction to folksonomies: Definition, Characteristics and Differences with Traditional Indexing Models]. *Información, Cultura y Sociedad*, 29, 13-26.

Yunta, L. R. (2009). Nuevas perspectivas para la difusión y organización del conocimiento. *Ix Congreso Isko-España*, 2, 832–845. Retrieved from http://eprints.rclis.org/12571/1/Comunicacion_Luis_RYunta_ISKO2009.pdf



PROPOSTA DE UM MÉTODO PARA O DESENVOLVIMENTO DE OBJETOS DE APRENDIZAGEM BASEADO NA SEMIÓTICA E NO *DESIGN* DA INFORMAÇÃO

Marcia Cristina dos Reis¹, Edberto Ferneda²

¹Universidade Estadual Paulista, Instituto Federal do Paraná, 0000-0001-6199-0309, marcia.reis@ifpr.edu.br

²Universidade Estadual Paulista 0000-0002-8808-1217, ferneda@marilia.unesp.br

RESUMO Os objetos de aprendizagem estão sendo cada vez mais utilizados para subsidiar atividades de ensino, em cursos presenciais e à distância. São importantes para o contexto educacional, pois possibilitam a elaboração de novas estratégias pedagógicas, facilitam a disseminação de conhecimento, ampliam a abrangência da educação *online* e são essenciais para a preservação e socialização da informação. Em contrapartida, para que isto ocorra, é necessário que eles sejam desenvolvidos de forma padronizada, que possuam uma interface eficiente de interação com o usuário, que sejam compreensíveis e que possam ser recuperados de forma ágil e fácil. Nesse contexto, o presente artigo tem como objetivo apresentar a proposta de um método para o desenvolvimento de objetos de aprendizagem, tendo como base a semiótica e as diretrizes do *design* da informação. Para isto, adotou-se pesquisa exploratória e bibliográfica, baseada em autores relevantes das áreas da Educação e da Ciência da Informação, mais especificamente em Objetos de Aprendizagem, Semiótica e *Design* da Informação. Como principais resultados, disponibilizou-se um método que, embora ainda não tenha sido testado, poderá contribuir significativamente para a melhoria da qualidade dos objetos de aprendizagem e, conseqüentemente, facilitar o seu processo de desenvolvimento e incentivar a sua utilização.

PALAVRAS-CHAVE *Ciência da Informação, Design da Informação, Objetos de Aprendizagem, Semiótica, Tecnologias Educacionais.*

ABSTRACT Learning objects are increasingly being used to subsidize teaching activities, in face-to-face and distance learning courses. They are important for the educational context, since they allow the elaboration of new pedagogical strategies, facilitate the dissemination of knowledge, broaden the scope of online education and are essential for the preservation and socialization of information. On the other hand, for this to happen, they must be developed in a standardized way, that have an efficient user interaction interface, that are comprehensible and can be retrieved in an agile and easy way. In this context, the present article aims to present the proposal of a method for the development of learning objects, based on the semiotics and guidelines of information design. Para isto, we adopted exploratory and bibliographic research, based on relevant authors from the areas of Education and Information Science, more specifically in Learning Objects, Semiotics and Information Design. As main results, a method was available that, although not yet tested, can contribute significantly to the improvement of the quality of learning objects and, consequently, facilitate its development process and encourage its use.

KEYWORDS *Information Science, Information Design, Learning Objects, Semiotics, Educational Technologies.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A tecnologia da informação é fator determinante para o sucesso de universidades, centros de ensino, escolas e instituições que visam desenvolver novas técnicas e estratégias para melhorar o processo de ensino e aprendizagem. Perrenoud (2000, p.65) defende que a “utilização destas ferramentas permite que sejam criadas situações de aprendizagens ricas, complexas, diversificadas, não fazendo com que todo trabalho repouse sobre o professor”. Neste cenário tecnológico, os objetos de aprendizagem estão sendo amplamente utilizados por inúmeras instituições como ferramenta de apoio à educação e disseminação de conhecimento (Mehlhorn et al., 2011).

No entanto, apesar das suas inúmeras contribuições, a literatura ainda não apresenta um método de desenvolvimento único e consolidado para este tipo de recurso. O que se observa são aplicações elaboradas em áreas específicas, de acordo com critérios definidos por cada autor, sem nenhum padrão de similaridade entre as etapas envolvidas no processo. Essa ausência de um método sistematizado gera dificuldades de entendimento das variáveis envolvidas e ausência de avaliações mais rígidas do procedimento, uma vez que, com a impossibilidade de realizações de testes, não se pode afirmar que um método é mais eficiente que outro para a construção de materiais educacionais.

Como forma de minimizar os problemas observados, o presente artigo tem como objetivo propor um método para o desenvolvimento de objetos de aprendizagem, com foco na Semiótica e no *Design* da Informação, buscando assegurar a efetividade da comunicação a partir da simplificação de processos de percepção, leitura, compreensão, memorização e uso da informação apresentada (Frascara, 2011).

A metodologia contou com pesquisa bibliográfica, de caráter exploratório, fundamentada em autores relevantes da literatura da Educação e da Ciência da Informação, mais especificamente em Objetos de Aprendizagem e *Design* da Informação.

O estudo é relevante para a Ciência da Informação, pois discute questões relacionadas à Semiótica e ao *Design* da Informação e sua contribuição no desenvolvimento de objetos de aprendizagem mais eficazes e interativos. Oliveira (2015) afirma que, após ampla investigação, não encontrou trabalhos publicados em língua portuguesa que investiguem o *Design* de Informação na área da Ciência da Informação, indicando um campo ainda pouco explorado e com perspectivas para o desenvolvimento científico.

No âmbito da Educação, o estudo se destaca pela possibilidade de ampliar os conceitos sobre os objetos de aprendizagem, que são recursos pedagógicos valiosos para a melhoria do processo educacional, uma vez que subsidiam diferentes práticas de ensino e representam uma abordagem mais eficiente para despertar o interesse do aluno e desenvolver habilidades intelectuais de pesquisa e investigação (Behar et al., 2009; Scortegagna, Barrère, & Barbosa, 2012).

Nesta perspectiva, pretendeu-se apresentar os conceitos fundamentais acerca dos objetos e aprendizagem, da Semiótica e do *Design* da Informação. Posteriormente, apresentou-se a proposta inicial de um método para o desenvolvimento de objetos de aprendizagem.

BASES TEÓRICAS DOS OBJETOS DE APRENDIZAGEM, DA SEMIÓTICA E DO DESIGN DA INFORMAÇÃO

A inserção de novos métodos, técnicas e tecnologias fazem parte do cenário educacional contemporâneo e permitem a utilização de uma combinação variável de recursos pedagógicos, que podem ser utilizados em diversas modalidades, incluindo cursos presenciais e à distância.

Os recursos estimulam os alunos a desenvolverem habilidades intelectuais de pesquisa e investigação, pois o conteúdo não lhes é dado pronto. Isso os instiga a estarem mais concentrados e interessados em aprender. Estimulam a buscar informações sobre um assunto e relacioná-las com aquelas adquiridas em outros momentos. E, ainda promovem cooperação entre os alunos (Scortegagna, Barrére, & Barbosa, 2012, p. 1967).

Em razão das diferentes concepções dos pesquisadores, não existe na literatura um entendimento único ou uma nomenclatura específica para o conceito de objetos de aprendizagem, que aparecem como: objeto de aprendizagem virtual, objeto jogo, objetos educacionais, objetos instrucionais, objetos inteligentes, materiais de aprendizagem *on-line*, dentre outras (Ferlin, 2009).

Em relação às definições, uma das mais citadas é da IEEE (LOM, 2002), que afirma que um objeto de aprendizagem é “qualquer entidade, digital ou não, que possa ser usada para o ensino, educação ou treinamento”.

Hoffmann et al. (2007) e Silva (2011) apresentam alguns exemplos de objetos de aprendizagem: apostilas, mapas, jogos, áudio, vídeo, animação, simulação, páginas web, *software* educacional, figuras, mapas mentais e todos os demais recursos que possam ser usados para compor uma aula, uma disciplina, um curso ou simplesmente facilitar e promover a aprendizagem em um contexto específico.

Para que os objetos de aprendizagem possam atender os propósitos para os quais foram desenvolvidos e contribuir para o ensino e aprendizagem, é importante que sejam incluídos conceitos semióticos e de *design* da informação no processo de desenvolvimento destes materiais, como forma de auxiliar a equipe no processo de comunicação e significação, bem como na organização do conteúdo que se pretende comunicar.

Segundo Santaella (2015), a semiótica é uma das disciplinas que fazem parte da ampla arquitetura filosófica de Peirce, e está fundamentada na Fenomenologia, uma quase-ciência que investiga os modos como as coisas são aprendidas à medida que aparecem à mente das pessoas: um cheiro, um ruído de chuva, uma formação de nuvens, ou mesmo algo mais complexo, como um conceito abstrato provocado por uma lembrança.

De acordo com Peirce, os signos podem ser classificados em três tricotomias, a partir das relações que ele estabelece: a primeira, que relaciona o signo a ele mesmo (suas propriedades internas, seu poder para significar), a segunda, que se refere ao signo em relação ao seu objeto (aquilo que ele indica, se refere ou representa) e a terceira, que diz respeito ao signo em relação ao interpretante (efeitos que ele é capaz de produzir na mente dos seus receptores) (Perez, 2004; Silva & Almeida, 2006).

A primeira tricotomia considera o signo nele mesmo, ou seja, seu modo de ser ou sua aparência, classificando-o em três espécies: quali-signo, sin-signo e legi-signo. A segunda analisa o signo em relação ao objeto, produzindo um ícone, índice ou símbolo. E a última tricotomia refere-se à relação do signo com o seu interpretante, gerando três possibilidades: o interpretante imediato, dinâmico e final.

Em relação ao *design* da informação, a literatura ainda é incipiente e apresenta muitas lacunas teóricas e conceituais, destacando-se pela carência relacionada à nomenclatura, definições, objetivos, metodologias e áreas convergentes. Além disso, não existe um consenso bem definido sobre os seus limites e as áreas de atuação, sendo tema de confusão e incertezas entre os profissionais envolvidos. Entre as inúmeras abordagens, o termo pode ser utilizado para contemplar uma pluralidade de conceitos tais como: projetar, desenhar, esboçar, planejar, modelar, entre outros (Oliveira, 2015).

Design da informação é uma área do design gráfico que objetiva equacionar os aspectos sintáticos, semânticos e pragmáticos que envolvem os sistemas de informação por meio da contextualização, planejamento, produção e interface gráfica da informação junto ao público alvo. O princípio básico é otimizar o processo de aquisição de informação efetivado nos sistemas de comunicação analógicos e digitais (Sbdi, 2013).

O *Design* da Informação é considerado interdisciplinar, uma vez que dialoga com diversas outras disciplinas, buscando facilitar a comunicação por meio da apresentação da informação e otimizando o seu processo de aquisição pelos usuários (Mijksenaar, 1997). Está presente nas áreas de Linguagem, Artes e Estética, Comunicação, Cognição e Informação e vem sendo utilizado em diversas vertentes de conhecimento, tais como: fotografia, cartografia, *design* gráfico, industrial, arquitetura e outros. Possui contribuições de diversos autores com a finalidade de apresentar diretrizes, processos, métodos e técnicas que possam auxiliar na tarefa de elaborar mensagens eficientes, satisfatórias, destacando diversos autores: Engelhardt (2002), Lipton (2007), Lohr (2003), Lupton e Phillips (2008), Norman (1998), Pettersson (2002, 2012) e Redig (2004).

PROPOSTA DE UM MÉTODO PARA O DESENVOLVIMENTO DE OBJETOS DE APRENDIZAGEM

O método proposto para o desenvolvimento de objetos de aprendizagem é dividido em duas etapas principais, sendo a primeira fundamentada na Semiótica e a segunda no *Design* da Informação, conforme pode ser observado no quadro abaixo.

Quadro 1. Método Proposto para o Desenvolvimento de Objetos de Aprendizagem

Etapa/Abordagem	Fases
ETAPA 1 - SEMIÓTICA <i>Processo de Comunicação e Significação</i> (O que Comunicar?)	a. Identificação dos Valores Centrais
	a.1 Construir um Personagem
	a.2 Identificar uma Voz Visual
	a.3 Associar com as Classes Sígnicas
	b. Geração de Alternativas
ETAPA 2 - DESIGN DA INFORMAÇÃO <i>Planejamento e Organização da Informação</i> (Como Comunicar?)	c. Avaliação da Comunicação do Objeto de Aprendizagem
	d. Aplicação dos Princípios e Diretrizes de Pettersson
	d.1 Aplicar os Princípios Funcionais
	d.2 Aplicar os Princípios Estéticos
	d.3 Aplicar os Princípios Administrativos
	d.4 Aplicar os Princípios Cognitivos
e. Avaliação do Objeto de Aprendizagem	

Fonte: adaptado de Niemeyer (2013) e Pettersson (2012)

Dois elementos são importantes no método proposto: o gerador (responsável pelo objeto de aprendizagem) e o interpretador (estudante). O primeiro é responsável pela escolha da mensagem que se deseja comunicar, das estratégias e dos recursos mais indicados, podendo ser o professor, o pedagogo, o pesquisador, o detentor do conhecimento ou uma equipe pedagógica, formada por diversos profissionais interessados na elaboração do material educacional. O interpretador é o destinatário (público alvo) que se deseja atingir.

ETAPA 1 – PROCESSO DE COMUNICAÇÃO E SIGNIFICAÇÃO

IDENTIFICAÇÃO DOS VALORES CENTRAIS

A primeira etapa do processo refere-se à identificação dos valores centrais referentes ao objeto de aprendizagem que será desenvolvido. A atividade deve ser iniciada com um *brainstorming*, no qual serão levantados os conceitos e as características do material para o qual será elaborada a comunicação. Uma primeira possibilidade de controle do processo surge quando os resultados dessa seção são formulados como um *re-briefing* e analisados com o professor ou equipe pedagógica. Algumas questões podem orientar esta etapa:

- a) Qual é o problema ou situação para resolução?
- b) Qual é o contexto do material educacional?
- c) Que desafios e expectativas deverão ser implementadas?
- d) Quais são as competências necessárias para atingir os objetivos da comunicação?
- e) Que recursos (humanos, tecnológicos e informacionais) são necessários?
- f) Que informações interdisciplinares são requeridas?

Após o levantamento das primeiras informações sobre o objeto de aprendizagem, a etapa de identificação de valores centrais é composta por três partes: construir um personagem, identificar uma voz visual e associar com as classes sógnicas. Segundo Niemeyer (2013), a primeira parte relacionada à construção de um personagem, tem como objetivo identificar as qualidades centrais implícitas no objetivo estratégico do objeto de aprendizagem para o qual será elaborado o processo de comunicação. Trata-se de três ou quatro palavras que expressem o sentimento, a base da sua existência e que, se forem abandonadas, tornam sem sentido aquilo que pretende alcançar. Neste ponto, é necessário ter muita clareza sobre quais são as características fundamentais do público alvo (faixa etária, conhecimentos necessários, nível cultural, padrão econômico, idioma, preferência por recursos de aprendizagem, preferência de padrões estéticos, entre outros).

A partir do estudo das características do público alvo, é importante avaliar como o objeto de aprendizagem poderá funcionar em relação aos tipos de prazeres:

- a) Fisiológicos: referem-se ao corpo, às percepções sensoriais, às primeiras impressões;
- b) Sociais: tratam do relacionamento entre pessoas próximas, do *status* socioeconômico, do *status* cultural, de fatores inclusivos ou excludentes;
- c) Psicológicos: abordam a satisfação de necessidades emocionais e efetivas;
- d) Ideológicos: envolvem valores morais, geracionais, de nacionalidade e crença.

A segunda parte refere-se à busca de uma voz visual e deve ser elaborada em equipe, que pode ser composta por todos os interessados no projeto. Tem como objetivo compreender os atributos essenciais do objeto de aprendizagem e como eles se expressam perceptualmente para os integrantes do grupo. Neste ponto, os participantes devem elaborar, de forma intuitiva e espontânea, uma listagem de todas as palavras que possam expressar valores relacionados ao material educacional. Em seguida, os termos devem ser analisados cuidadosamente e, após sucessivas eliminações, a equipe deve encontrar três ou quatro palavras que sintetizem todo o conceito e resumam a ideia do objeto de aprendizagem a ser desenvolvido. Posteriormente, é preciso identificar os elementos predominantes como: a) características das cores; b) nível de luminosidade dominante; c) se as cores são saturadas, paracromáticas ou acromáticas, claras ou escuras; d) se as cores são análogas, contrastantes, complementares; e) se os acordes são constantes ou dissonantes; e) se há repetição de motivos, formando padrões; f) os tipos de texturas presentes; g) a dominância das espessuras e das formas de linhas e de áreas. Com a finalização desta etapa, o gerador terá em mãos um *briefing* do material para a construção do enunciado que se propõe, dando assim, a “voz” formal à “personalidade” do objeto de aprendizagem. Cabe a ele decidir quais serão apresentados de forma imagética ou verbal, de acordo com o conceito formal que será construído para representar o material.

A terceira parte de identificação dos valores centrais consiste em associar nas classes sígnicas todos os componentes dos enunciados, tendo em vista as suas relações possíveis com o objeto (icônica, indicial e simbólica). Essa associação deve ser realizada conforme o quadro abaixo:

Quadro 2. Associação nas Classes Sígnicas

	Componente 1 do enunciado	Componente 2 do enunciado	Componente 3 do enunciado
ICÔNICA Categoria 1 da relação com o Objeto	Representação gráfica do componente 1, segunda a Categoria 1 da relação com o Objeto	Representação gráfica do componente 2, segunda a Categoria 1 da relação com o Objeto	Representação gráfica do componente 3, segunda a Categoria 1 da relação com o Objeto
INDICIAL Categoria 2 da relação com o Objeto	Representação gráfica do componente 1, segunda a Categoria 2 da relação com o Objeto	Representação gráfica do componente 2, segunda a Categoria 2 da relação com o Objeto	Representação gráfica do componente 3, segunda a Categoria 2 da relação com o Objeto
SIMBÓLICA Categoria 3 da relação com o Objeto	Representação gráfica do componente 1, segunda a Categoria 3 da relação com o Objeto	Representação gráfica do componente 2, segunda a Categoria 3 da relação com o Objeto	Representação gráfica do componente 3, segunda a Categoria 3 da relação com o Objeto

Fonte: Niemeyer (2013, p. 70)

Segundo Niemeyer (2013), cada célula do quadro corresponde a um princípio de representação para o componente do enunciado. Cada item do quadro é analisado por seus elementos imagéticos, isto é, decompostos em Quali-signos, Sin-signos e Legi-signos. É importante destacar que neste estágio o número de possibilidades se multiplica, uma vez que cada unidade é, muitas vezes, composta de uma série de signos dos três tipos. O processo de seleção é muito intenso nesta fase, pois existem diversos elementos que podem ser ligados, formando uma série de associações que irá construir um conceito formal.

GERAÇÃO DE ALTERNATIVAS

A fase de Geração de Alternativas tem como foco subsidiar a elaboração dos esboços, a partir dos resultados obtidos na etapa anterior. É importante que outro *brainstorming* seja realizado, em que o enunciado de cada esboço seja discutido, permitindo o confronto de diferentes propostas e a concretização das abordagens preliminares. A geração de alternativas está baseada na teoria semiótica, com a diferenciação dos vários tipos de signos nas diferentes dimensões.

As alternativas resultantes são submetidas a uma avaliação empírica por meio de testes comparativos, para se fazer a seleção. Com isso, pretende-se verificar se a solução adotada veicula os conceitos inicialmente definidos. O resultado final cumpre, assim, o objetivo da metodologia adotada: dar expressão formal ao propósito comunicacional (NIEMEYER, 2013, p.72).

A decisão final e a realização do projeto são feitas com base sobre este último controle e seleção. Ao utilizar os elementos de linguagem, em princípio denotativos do repertório do Interpretador, o gerador considera como definida a sua dimensão pragmática. Assim, os enunciados elaborados conforme esta estratégia, são considerados como dotados de características que propiciem o tipo de comunicação desejada.

AValiação DA COMUNICAÇÃO DO OBJETO DE APRENDIZAGEM

Esta fase consiste na avaliação da comunicação elaborada, disponibilizando ao gerador as seguintes atividades:

- a) definição dos requisitos a serem testados;
- b) possibilidade de isolar os aspectos a serem testados;
- c) análise da consistência dos requisitos a serem testados;
- d) possibilidade de reavaliação dos testes;
- e) avaliação de imagens semelhantes e concorrentes; e
- f) experimentação com modelos conceituais.

Após o desenvolvimento das tarefas previstas nesta fase, que contempla as avaliações e os ajustes necessários no processo de comunicação e significação do objeto de aprendizagem, encerra-se o ciclo de atividades fundamentadas no campo da Semiótica. A seguir, será apresentada a segunda etapa do método proposto neste estudo, juntamente com as respectivas fases que o compõe.

ETAPA 2 – PLANEJAMENTO E ORGANIZAÇÃO DA INFORMAÇÃO

APLICAÇÃO DOS PRINCÍPIOS E DIRETRIZES DE PETERSSON

Pettersson (2012) apresenta quatro grupos de princípios guiados pelo *design*, que contribuem para o desenvolvimento de mensagens eficientes em materiais de aprendizagem. De acordo com o autor, estes princípios são universais, mas devem ser adaptados ao contexto social do público alvo, para que haja uma representação mais adequada. Nesta fase do método, deve-se aplicar ao objeto de aprendizagem os princípios funcionais, estéticos, administrativos e cognitivos.

Em relação aos princípios funcionais, Pettersson (2012) orienta que é importante atender os seguintes requisitos:

- a) Problema: definir adequadamente a situação para resolução, levantando aspectos dos emissores, dos receptores, das representações e do contexto. Este passo pode ser subsidiado pelas informações sobre o público alvo levantadas na primeira etapa do método;
- b) Estrutura: organizar adequadamente a estrutura do conteúdo e a hierarquia da informação;
- c) Clareza: as informações presentes no material educacional devem ser transparentes, legíveis e sem ambiguidade;
- d) Simplicidade: as informações devem ser adaptadas adequadamente para os leitores, estando relacionadas à percepção, processamento e memória;
- e) Ênfase: utilizar elementos específicos para dar ênfase à informação, criando contrastes claros para atrair, dirigir e reter a atenção do público alvo;
- f) Unidade: combinar as informações disponibilizadas no material com coesão e coerência. Usar também *layouts* e tipografias consistentes.

Pettersson (2012) apresenta dois princípios estéticos que são importantes para a elaboração de objetos de aprendizagem:

- a) Harmonia: desenvolver normas para *templates* de *design* e encontrar equilíbrio entre seus elementos (*e.g.* tipografias e cores).
- b) Proporção: encontrar preferências dos usuários pelas mais variadas proporções estéticas, que podem ser também levantadas na primeira etapa do método.

Além dos princípios funcionais e estéticos, é importante que o material educacional atenda também alguns requisitos administrativos que, de acordo com Pettersson (2012), são:

- a) Acesso: aplicar normas internacionais que visem o correto armazenamento, a facilidade de acesso e a segurança do objeto de aprendizagem no sistema de gerenciamento;
- b) Custo: considerar os custos para o *design* na produção do material, sua distribuição e seu armazenamento;
- c) Ética: avaliar se o conteúdo apresentado respeita as regras de ética, direitos autorais e *copyright*;
- d) Qualidade: revisar as informações disponibilizadas no material em relação à sua credibilidade e ao uso de terminologias.

Pettersson (2012) destaca ainda a importância de quatro outras diretrizes, que fazem parte da categoria dos princípios cognitivos:

- a) Atenção: utilizar recursos que chamem a atenção do público alvo, através de diretrizes para textos, símbolos, *layout* e cores;
- b) Percepção: facilitar e orientar a percepção através de elementos como clausura, continuidade, proximidade, similaridade e contraste;
- c) Processamento Mental: facilitar o processamento mental através da adequada correspondência de palavras, frases, parágrafos, textos, figuras, *layout* e realidade;

buscando entender o conceito da representação das palavras e das subculturas da audiência;

- d) Memória: apresentar somente um número limitado de elementos informacionais ao mesmo tempo, promovendo conteúdos significativos e conectando ilustrações e textos.

É importante destacar que Pettersson (2012) apresenta um total de 150 diretrizes ligadas aos princípios básicos de *Design* da Informação. No entanto, foram selecionadas apenas aquelas que são indispensáveis para o planejamento e a organização da informação em objetos de aprendizagem e que possam contribuir para o desenvolvimento de mensagens eficazes e satisfatórias.

AVALIAÇÃO DO OBJETO DE APRENDIZAGEM

A última fase do método proposto refere-se à avaliação do objeto de aprendizagem desenvolvido. Nesta etapa, além das análises realizadas pelo gerador durante o processo, sugere-se que um grupo de usuários pertencentes ao público alvo possa avaliar o material desenvolvido e verificar se os princípios propostos por Pettersson (2012) foram atendidos e implementados. Os testes realizados podem potencializar a qualidade do material desenvolvido e permitir a correção de possíveis erros.

Deste modo, conclui-se o método proposto para o desenvolvimento de objetos de aprendizagem, fundamentado nos princípios e diretrizes da Semiótica e do *Design* da Informação. Neste percurso, o gerador deverá compreender as complexas relações que se estabelecem no processo de ensino e aprendizagem e, a partir disso, produzir materiais que possibilitem uma melhor interação com o usuário. É importante destacar, no entanto, que as etapas disponibilizadas neste estudo não são regras rígidas ou direcionamentos precisos, mas apenas elementos que podem subsidiar a elaboração de um processo comunicacional e de organização da informação, contribuindo diretamente para a concepção e produção de mensagens mais eficazes e significativas.

CONSIDERAÇÕES FINAIS

O desenvolvimento de objetos de aprendizagem, direcionados principalmente para o compartilhamento e a reutilização de recursos, vem se consolidando nos últimos anos, devido aos inúmeros benefícios evidenciados no contexto educacional. Os pesquisadores da área têm direcionado esforços para projetar metodologias e ferramentas que consigam minimizar os problemas observados no desenvolvimento e catalogação de objetos de aprendizagem, bem como na recuperação destes recursos. Neste contexto, o estudo teve como objetivo central propor um método para o desenvolvimento de objetos de aprendizagem, baseado nos conceitos semióticos e nos princípios e diretrizes do *Design* da Informação.

O método proposto disponibiliza um conjunto de etapas lógicas e sistematizadas para desenvolver o processo de comunicação e significação dos objetos de aprendizagem, bem como realizar o planejamento e a organização da informação a ser apresentada. Independentemente do meio de transmissão escolhido, um material com uma comunicação bem elaborada e um *design* de informação adequado em sua mensagem, terá condições de atender os requisitos do usuário e as demais vertentes necessárias: esteticamente, economicamente e ergonomicamente. Tal abordagem tenta preencher uma lacuna existente na literatura que, apesar de discutir uma pluralidade de métodos, nenhum deles é apresentado de forma consolidada, o que dificulta a sua reprodutibilidade.

Neste sentido, como sugestão de trabalhos futuros, seria de grande importância que o método fosse testado e aprimorado no contexto do desenvolvimento dos objetos de aprendizagem. Em outra vertente, as diretrizes propostas neste estudo apresentam potencial ainda para serem utilizadas na avaliação de materiais educacionais, uma vez que muitos apresentam problemas referentes à qualidade, tais como: ausência de contextualização, conteúdo superficial, mensagem confusa, falta de integração de conteúdos, dentre outros.

É importante destacar ainda que este estudo não tem a intenção de apresentar ao gerador regras rígidas ou fórmulas claras e concretas para o desenvolvimento de materiais educacionais. O que se propõe, na verdade, são alguns princípios e diretrizes que, se forem incluídos no processo, poderão contribuir para que os objetos de aprendizagem desenvolvidos tenham mais qualidade e apresentem resultados mais eficientes para o contexto educacional.

REFERÊNCIAS BIBLIOGRÁFICAS

Behar, P. A. et al. (2009). Objetos de aprendizagem para educação a distância. In P. A. Behar (Org.), *Modelos pedagógicos em educação a distância* (pp. 66-92). Porto Alegre, Brasil, Artmed.

Engelhardt, Y. (2002). *The language of graphics*. Amsterdam: ILLC.

Ferlin, J. (2009). *Repositório de objetos de aprendizagem para a área de Informática*. Trabalho de Conclusão de Curso, Universidade Estadual de Santa Catarina (UDESC), Florianópolis, Brasil.

Frascara, J. (2011). *¿Qué es el diseño de información?* Buenos Aires: Ediciones Infinito.

Gomes, F. M. (2014). O uso de objetos de aprendizagem na educação: Recursos digitais interativos em repositórios gratuitos. *Simpósio internacional de educação a distância – Encontro de pesquisadores em educação a distância SIED* (pp. 1-12), São Carlos, São Paulo, Brasil.

Hoffmann, A. V. et al. (2007). *Objetos de aprendizagem para a TV pendrive: Conhecendo e produzindo*. (3a ed). Curitiba: Secretaria da Educação.

Ipton, R. (2007). *The practical guide to information design*. Hoboken, NJ: John Wiley & Sons, Inc.

Lohr, L. L. (2003). *Creating graphics for learning and performance*. Lessons in Visual Literacy. Upper Saddle River, New Jersey: Pearson Education, Inc.

Lom – Learning Object Metadata. (2002). *Draft standard for learning object metadata*: IEEE 1484.12.1-2002. Learning Technology Standards Committee. Recuperado em 2 novembro, 2016, de http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf

Lupton, E., & Phillips, J. C. (2007). *Novos fundamentos do design*. Tradução de Cristian Borges. São Paulo: Cosac Naify.

Mehlhorn, S., Parrott, S., Mehlhorn, J., Burcham, T., Roberts, J., & Smartt, P. (2011). Use of digital learning objects to improve student problem solving skills. *Annual Meeting*, Corpus Christi, Texas, 5-8.

- Mijksenaar, P. (1997). *Visual function: An introduction to information design*. 010 Publishers: Rotterdam.
- Niemeyer, L. (2013). *Elementos de semiótica aplicados ao design*. Rio de Janeiro: 2AB.
- Norman, D. A. (1998). *The design of everyday things*. London: The MIT Press.
- Nöth, W. (2005). *Panorama da semiótica: De Platão a Peirce*. São Paulo, Annablume.
- Oliveira, J. A. D. B. (2015). *A Ciência da Informação e o Design da Informação: Perspectivas interdisciplinares*. Dissertação de Mestrado, Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Filosofia e Ciências, Marília, SP, Brasil.
- Oliveira, R. R. S. (2014). *O design da informação em redes sociais educacionais brasileiras: Uma análise a partir da Teoria da Atividade*. Dissertação de Mestrado, Universidade Federal de Pernambuco, Centro de Artes e Comunicação, Recife, PE, Brasil.
- Peirce, C. S. (2000). *Semiótica*. (3a ed.). São Paulo: Perspectiva.
- Perez, C. (2004). *Signos da marca: Expressividade e sensorialidade*. São Paulo: Editora Thomson Learning.
- Perrenoud, P. (2000). *Dez novas competências para ensinar*. Porto Alegre: Artmed.
- Pettersson, R. (2002). *Information design: An introduction*. Tullinge: John Benjamins Publication.
- Pettersson, R. (2012). *It Depends: principles and guidelines*. Tullinge: Institute for Infology.
- Quintão, F. S., & Triska, R. (2013). Design de informação em interfaces digitais: Origens, definições e fundamentos. *Revista Brasileira de Design da Informação*, São Paulo, 10(2), 105-118.
- Redig, J. (2004). Não há cidadania sem informação, nem informação sem design. *Revista Brasileira de Design da Informação*, São Paulo, 1(1), 58-66.
- Santaella, L. (2005). *Matrizes da linguagem e do pensamento: Sonora, visual, verbal: Aplicações na hipermídia*. São Paulo: Iluminuras.
- Santaella, L. (2007). *O que é semiótica*. São Paulo: Brasiliense.
- Santaella, L. (2015). *Semiótica aplicada*. São Paulo: Cengage Learning.
- Sbdi - Sociedade Brasileira de Design de Informação. (2013). Brasil. Recuperado em 12 novembro, 2016, de <http://www.sbdi.org.br>.
- Scortegagna, L., Barrère, E., & Barbosa, G. (2012). Objetos de aprendizagem para ensino de matemática: Reflexões. *Anais da Reunião Latino Americana de Educação Matemática* (pp. 1967-1973), Campinas, SP, Brasil, 26.
- Silva, J. A. da, & Almeida, C. C. (2006). Análise de embalagens e linguagem regional: aplicação da semiótica peirceana no estudo de rótulos de embalagens de erva-mate. *UNIrevista*, 1(3), 1-14.

Silva, J. M. C. (2011). *Análise técnica e pedagógica de metadados para objetos de aprendizagem*. Tese de Doutorado, Centro Interdisciplinar de Novas Tecnologias, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil.



GESTÃO TERMINOLÓGICA: SUBSÍDIOS PARA A CONSTRUÇÃO DO VOCABULÁRIO CONTROLADO DA CARNAÚBA

Jacqueline A. Souza¹, Monica Marques Carvalho Gallotti², Jailma Santos³

¹Professora Assistente II da Universidade Federal do Rio Grande do Norte, 0000-0001-5677-8282, jackebci@gmail.com

²Professora Adjunta IV da Universidade Federal do Rio Grande do Norte, 0000-0002-3044-2032, monica_mcg@hotmail.com

³Bibliotecária do Instituto Federal de Educação, Ciência e Tecnologia do Ceará, 0000-0001-8569-896X, jailma.snt@gmail.com

RESUMO O contexto atual complexo traz em si a necessidade de proposição de estratégias pontuais de organização da informação e do conhecimento em variadas áreas. Diante disso, o Nutseca - Núcleo Temático da Seca e do Semiárido situado na Universidade Federal do Rio Grande do Norte - vem ao longo de 35 anos de existência reunindo mais de 100 anos de informação e coleções de documentos sobre a temática seca e assuntos relacionados. Dentre estas coleções está a uma coleção temática sobre Carnaúba. Este trabalho, em forma de pôster, visa detalhar as ações de um projeto de extensão denominado “Vocabulário Controlado da Carnaúba” financiado pelo Banco do Nordeste do Brasil. Com vistas a organizar e representar essa coleção, foi desenvolvido um projeto cujo objetivo foi: 1) compor um vocabulário controlado da carnaúba, aplicando os princípios da Linguística de Corpus (LC) e ferramentas de gestão terminológica. Especificamente visou: 2) determinar quais premissas da LC são fundamentais para a elaboração de linguagens documentárias, 3) relatar no âmbito do ensino de OI uma prática interdisciplinar. Assim, a pesquisa justifica-se pela necessidade de desenvolver instrumentos para representação e tratamento temático da coleção e é relevante pela prática interdisciplinar que converge teorias, técnicas e metodologia da Linguística de Corpus, Biblioteconomia e Terminologia. Para execução da pesquisa, foram aplicados os pressupostos teórico-metodológicos da Linguística de Corpus no tocante a manipulação dos dados linguísticos e, conseqüentemente, a ferramenta computacional E-Termos. Pelo exposto, foram identificadas as dificuldades e vantagens encontradas com a aplicação dos pressupostos metodológicos da Linguística de Corpus, bem como um protótipo do vocabulário controlado da Carnaúba.

PALAVRAS-CHAVE *Organização da Informação. Vocabulário controlado. Linguística de Corpus.*

ABSTRACT The current complex context fosters the need to propose specific information and knowledge organization strategies in various areas. Thus, this work in form of a poster, aims to describe experiences related to an extension project of a controlled vocabulary in the subject of a *Carnaúba*. The project “Carnauba Controlled Vocabulary” received financial support from *Banco do Nordeste do Brasil* and took place at the *Núcleo Temático da Seca e do Semiárido*, a documentation center that deals with drought at the Federal University of Rio Grande do Norte. The center exists for over 35 years and has gathered more than 100 years of information and collections of documents on the subject. In order to organize and represent this collection, the project was developed and it proposed: 1) to elaborate a controlled vocabulary of *Carnaúba*, applying the principles of Corpus Linguistics (CL) and use terminology management tools. Specifically, it aimed at: 2) determining which premises of the CL are fundamental for the elaboration of documentary languages, 3) Describe an IO interdisciplinary practice. Thus, the research is justified by the need to develop instruments

aiming thematic treatment and representation of collections, relevant as an interdisciplinary practice that converges theories, techniques and methodology of Corpus Linguistics, Librarianship and Terminology. For the execution of the research, the theoretical-methodological assumptions of Corpus Linguistics were applied in relation to the manipulation of linguistic data and, consequently, the computational tool *E-Termos*. It was possible to identify the difficulties and advantages encountered with the application of the methodological assumptions of Corpus Linguistics, as well as a prototype of the *Carnaúba* controlled vocabulary.

KEY-WORDS *Information Organization. Controlled Vocabulary. Corpus Linguistics.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Caracterizado pela baixa umidade e pelo pouco volume pluviométrico, o semiárido nordestino representa 18,3% do território brasileiro e no Nordeste é formado por nove estados: Maranhão, Piauí, Paraíba, Pernambuco, Alagoas, Sergipe, Bahia e Rio Grande do Norte. Neste cenário, há uma riqueza de plantas adaptadas ao clima seco e que são economicamente exploradas para a produção de óleos como o Catolé, Marmeleiro, Oiticica; de látex como o Pinhão Maniçoba e de ceras, como a Carnaúba. (Suassuna, 2011).

Nomeadamente, a carnaúba (*Copernicia prunifera*) ou carnaíba, carnaubeira, trata-se de uma palmeira também conhecida como a árvore da vida, tem importância econômica, social e ambiental. Acerca da economia, Carvalho & Gomes (2009) ressaltam que consiste no conjunto de atividades que utilizam as folhas, o caule, o talo, a fibra, o fruto e as raízes dessa palmeira para a fabricação de inúmeros produtos artesanais e industriais, destacando a extração do pó de carnaúba, a partir das folhas, por sua importância na geração de emprego e renda complementar no campo e por ser a matéria-prima da cera, que possui inúmeras aplicações no ramo industrial, sendo um importante elemento para redução da pobreza na região nordestina, entretanto sua exploração sofre influências do clima seco.

Diante desta realidade acima mencionada e dada a importância destes recursos faz-se necessário a elaboração de estratégias em torno da organização de informações provenientes deste contexto. Tal necessidade se dá devido ao fato de que na atualidade existe uma dispersão de informação impulsionado pelo avanço da Tecnologia da Informação e Comunicação. Organizar a informação acumulada em estoques informacionais tem se tornado um desafio e ao mesmo tempo uma solução. Nesse sentido corroboramos com o pensamento de Rowley quando afirma que “organizar a informação para melhor compreendê-la e recuperá-la é imprescindível, para isso são necessários padrões de estruturação pré-definidos, visando facilitar, agilizar e otimizar a chegada da informação ao usuário final, resultando em sua satisfação ajudando-o em seus processos decisórios (Rowley, 2008).

Diante disso, o Nutseca - Núcleo Temático da Seca e do Semiárido - situado na Universidade Federal do Rio Grande do Norte, o qual contém coleções de documentos sobre a temática seca e assuntos relacionados. Assim, este trabalho, em forma de pôster, visa detalhar as ações de um projeto de extensão que recebeu financiamento do Banco do Nordeste do Brasil e visou a criação de um vocabulário controlado sobre Carnaúba. Com vistas a organizar e representar essa coleção, foi desenvolvido um projeto cujo objetivo foi: 1) compor um vocabulário controlado da carnaúba, aplicando os princípios da Linguística de Corpus (LC) e ferramentas de gestão terminológica. Especificamente visou: 2) verificar

quais premissas da LC são fundamentais para a elaboração de linguagens documentais, 3) relatar no âmbito do ensino de OI uma prática interdisciplinar.

O desenvolvimento desta pesquisa procurou conciliar as necessidades do Nutseca para indexar sua coleção e promover maior visibilidade e acesso ao conteúdo informacional sobre a seca e o semiárido que, dado seu impacto e importância social e econômica, proporcionalmente pode-se considerar que ainda academicamente não tem o seu potencial explorado, desde pouca produção na literatura científica como sistematização para acesso a fontes de informação. No âmbito do Departamento de Ciência da Informação da UFRN, tocante ao ensino na área de Organização da Informação, também proporcionou desenvolver uma atividade prática, evidenciando a interdisciplinaridade.

NÚCLEO TEMÁTICO DA SECA E DO SEMIÁRIDO

O Núcleo Temático da Seca e do Semiárido é um órgão suplementar ligado ao Centro de Ciências Sociais Aplicadas (CCSA) da Universidade Federal do Rio Grande do Norte (UFRN). É caracterizado como um centro de documentação especializado em Seca e assuntos relacionados. Seus principais objetivos são os de: propor estratégias em torno da captura, organização, tratamento e difusão da informação sobre a temática seca e semiárido bem como apoiar os programas de ensino, pesquisa e extensão em suas necessidades de acesso à informação especializada em seca e semiárido e dinamizar a produção científica concernente a seca e semiárido. (CARVALHO, 1998).

Seu acervo é constituído livros, periódicos, fotografias, artefatos tridimensionais, literatura cinzenta, anuários, índices, mapas, entre outros documentos compondo uma rica “ecologia” documental e massa crítica na área da seca. Seu acervo é tido como único e um dos maiores do Brasil. Este, está dividido em coleções específicas tais como: A Universidade e a Questão Nordestina, Seca e Semiárido, Carnaúba e a do Vale do Assú ou Projeto Baixo-Assú, construído numa visão epistemológica interdisciplinar. Esta vasta documentação é organizada com vistas a difusão da informação por meio de portais de informação e websites institucionais.¹

VOCABULÁRIO CONTROLADO E A LINGUÍSTICA DE CORPUS

Enriquecendo as pesquisas sobre a construção de ontologias e vocabulários controlados, Silva, Almeida & Souza (2013) consideraram que os procedimentos mais adequados são os sugeridos pelo manual disponível no sítio da Biblioteconomia, Informação e Tecnologia da Informação (BITI), autoria de Campos, Gomes & Motta, (2004):

- 1) Planejamento
- 2) Composição do vocabulário
- 3) Organização dos conceitos
- 4) Apresentação final
- 5) Avaliação

¹ Ver <https://ccsa.ufrn.br/nutseca/NUT-Seca/Nut%20Seca.html> e http://www.siabiclientes.com/nutseca/Telas/w_busca_rapida.php

Com base nestas etapas para construção, para atender aos objetivos propostos neste trabalho, discorrer-se-á sobre Linguística de Corpus. Neste sentido, convém apresentar o seu conceito, assim como a definição de corpus e suas características.

A Linguística de Corpus ocupa-se da coleta e exploração de corpora, ou conjuntos de dados linguísticos textuais que foram coletados criteriosamente com o propósito de servirem para a pesquisa de uma língua ou variedade linguística. Como tal, dedica-se à exploração da linguagem através de evidências empíricas, extraídas por meio de computador. (Berber Sardinha, 2000, p. 325).

Condicionada a tecnologia, a LC circunscreve-se a uma abordagem probabilística da língua, cujo armazenamento, acesso, manuseio e exploração ocorre via computador. Suas aplicações versam acerca da compilação dos corpora², desenvolvimento de ferramentas computacionais para análise linguística, concordanciadores, para descrição de linguagem, identificação de padrões, é útil para ensino-aprendizagem de línguas, processamento de língua natural, confecção de dicionários terminológicos e lexicográficos, para análises morfossintáticas, discursivas, semânticas, gêneros discursivos e traduções.

Neste sentido, Oliveira (2009) ressalta que a LC representa uma nova abordagem filosófica para os estudos da linguagem e que não pode ser definida apenas como uma metodologia emergente, mas uma nova maneira de desenvolver pesquisas. Como uma ciência empírica fornece dados sobre a linguagem em uso para descrevê-la de forma adequada, que não se restringe apenas a compilação e coleta de dados, mas para geração de novos conhecimentos, novas descrições das línguas, entre outras atividades que envolvam a experiência humana tal como é construída na linguagem, o que é concernente a Organização e Representação da Informação.

Central a LC, a definição de corpus que incorpora suas principais características e que expõe como pode ser útil no desenvolvimento de produtos informacionais é:

Um conjunto de dados linguísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso linguístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise' (Sanchez, 1995, *apud* Berber Sardinha, 2000, p. 338)

Acerca desta definição, Berber Sardinha (2000) pontuou aspectos relevantes destacados a seguir:

- **A origem:** Os dados devem ser autênticos, ou seja, o corpus deve ser composto de textos em linguagem natural. Os textos não podem já ter sido produzidos para fins de pesquisa linguística. Também, deve-se observar se o texto é produzido por falantes nativos.
- **A composição:** O conteúdo do corpus deve ser criteriosamente escolhido. Os princípios da escolha dos textos devem seguir, acima de tudo, as condições de naturalidade e autenticidade. Ou seja, o conteúdo do corpus deve ser selecionado a fim de garantir que o corpus tenha uma certa característica. Por exemplo, se o desejo é construir um corpus de português brasileiro escrito que represente a língua portuguesa, tal qual ela é escrita no Brasil, em sua totalidade, a coleta deve ser guiada por um

² O plural de corpus é corpora.

conjunto de critérios que garanta, entre outras coisas, que o maior número possível de tipos textuais existentes no português brasileiro esteja representado, que haja uma quantidade aceitável de cada tipo de texto e que a seleção dos textos seja aleatória, a fim de que não se contamine a coleta com variáveis indesejáveis.

- **A formatação:** Os dados do corpus devem ser legíveis por computador.
- **A representatividade:** O corpus deve ser representativo de uma língua ou variedade, ou seja, uma vez que a linguagem é de caráter probabilístico, há possibilidade de estabelecer uma relação entre traços que são mais comuns e menos comuns em determinado contexto.
- **A extensão:** o corpus deve ser vasto e, neste sentido, comporta 3 dimensões: o número de palavras, sendo que quanto maior o número de palavras, maior será a chance do corpus conter palavras de baixa frequência, as quais formam a maior parte das palavras de uma língua; número de textos, pois quanto mais texto, maior será a garantia de tipo textual, gênero e registro; e o número de gêneros ou tipos textuais, que permitirá maior abrangência do espectro genérico da língua.

Com base na definição de corpus e nos pontos relevantes a ser considerado na composição do corpus, podemos verificar que estes princípios são importantes na elaboração de linguagens documentárias. Por exemplo: uma vez que a linguagem documentária deve abarcar a maior variedade de descritores e não-descritores para que cumpra sua função.

METODOLOGIA

Em termos metodológicos este trabalho inicia por uma pesquisa bibliográfica acerca dos temas da Organização da Informação, construção de vocabulário controlado e Linguística de Corpus. Tal abordagem foi necessária para se traçar as premissas conceituais a fim de fundamentar o trabalho em tela. Em seguida foi feito um relato de experiência em consonância com os postulados de Minayo (1994) e Connelly & Clandinin (1995) cuja concepção geral prevê que registrar e relatar experiências científicas é fundamental para o avanço da ciência. Para execução da pesquisa, foram aplicados os pressupostos teórico-metodológicos da Linguística de Corpus no tocante a manipulação dos dados linguísticos e, conseqüentemente, a ferramenta computacional e-Termos³, acrônimo de Termos eletrônicos, um ambiente computacional colaborativo web de acesso livre e gratuito dedicado a gestão terminológica.

³ A ferramenta computacional é oriunda do Projeto e-Termos, um projeto acadêmico desenvolvido em parceria entre a Embrapa Informática Agropecuária (CNPTIA), Universidade de São Paulo (USP Campus de São Carlos, SP) e Universidade Federal de São Carlos (UFSCar), representados pelos laboratórios de pesquisa LabInfo (Laboratório de Organização e Tratamento da Informação Eletrônica), NILC (Núcleo Interinstitucional de Linguística Computacional) e o GETerm (Grupo de Estudos e Pesquisas em Terminologia), localizados nas três instituições, respectivamente. Informações disponíveis em: <https://www.etermos.cnptia.embrapa.br/index.php>

RESULTADOS

O Vocabulário Controlado da Carnaúba pretende auxiliar a indexação das fontes digitalizadas, implicando um tratamento temático do acervo do Nutseca. Dada a configuração do projeto, em sua primeira fase, os resultados obtidos são a configuração da sua macroestrutura onde se observam o domínio temático estruturado através da hierarquização e seus desdobramentos, como também as relações lógicas que se baseiam na similaridade das características entre conceitos, formando, assim, uma relação de gênero / espécie. O vocabulário é composto por um conjunto com média de 400 termos que representam conceitos relacionados à Cultura da Carnaúba, os quais estão dispostos em uma estrutura em ordem alfabética, destacando-se as relações de equivalência.

Tocante as etapas para construção e gestão dos termos, foram implementadas todas as etapas sugeridas pela LC. A compilação do corpus deu-se com textos oriundos da web e com a própria coleção do Nutseca, com 296 textos autênticos e 6 gêneros distintos. Considerando a abordagem probabilística da LC, foram observados os aspectos quantitativos e qualitativos, com um corpus de 1.288.461 palavras e bom índice de riqueza vocabular.

Ainda sobre gestão terminológica, verificou-se que as etapas sugeridas para elaboração de vocabulário controlado, convergem em muitos pontos com as etapas implementadas no uso da ferramenta e-Termos.

CONCLUSÕES

O Relato de experiências no terreno das práticas de Organização da Informação no geral em específico da construção de vocabulários controlados é uma mais valia pois permitem a socialização de procedimentos e metodologias adotadas agregando valor à base de conhecimento existente no assunto. Baseando-nos no que foi apresentado, concluiu-se que a Linguística de Corpus fornece aportes teórico e metodológico na gestão terminológica para a elaboração de linguagens documentais. Ressalta-se a importância de critérios para compilação do corpus e destaca-se o uso das ferramentas computacionais, sobretudo por promoverem a colaboratividade entre diferentes atores envolvidos no processo, como bibliotecários e os especialistas de domínios específicos que, por exemplo, são consultados para validar os termos que compõem os vocabulários e esclarecer dúvidas, dentre outros aspectos.

REFERÊNCIAS BIBLIOGRÁFICAS

Barité, M. (2011). Sistemas de Organización del Conocimiento: una tipología actualizada; In: Sistemas de Organização do Conhecimento: uma tipologia atualizada. *Informação & Informação*, 16(2), 122-139. Doi: 10.5433/1981-8920.2011v16nesp.p122

Campos, M. L. D. A., Gomes, H. E., & Motta, D. F. (2006). *Elaboração de tesouro documentário: Tutorial*.

Carlan, E., & Medeiros, M. B. B. (2011). Sistemas de Organização do Conhecimento na visão da Ciência da Informação. *Revista Ibero-Americana de Ciência da Informação*, 4(2). Disponível em: <http://repositorio.unb.br/handle/10482/12867>

Carvalho, J. D., & Gomes, J. (2009). Pobreza, emprego e renda na economia da Carnaúba. *Revista Econômica do Nordeste*, 40(02), 361-378.

Carvalho, R. P. F. de. (1999). Núcleo temático da seca/UFRN: uso do acervo informacional. *Informação & Sociedade: Estudos*, 9(2).

Freire, I. M., Carvalho, L. M., Carvalho, M. M., & Aranha, T. D. Q. (2008). Ampliando o acesso livre a informação: a digitalização do acervo do Núcleo Temático da Seca. *Informação & Sociedade*. João Pessoa, v.18, n.2, p. 137-142, maio/ago. Disponível em: <http://www.repositorio.ufrn.br:8080/jspui/handle/1/3125>

Lima, J.L.O. & Alvares, L. (2012). *Organização e representação da informação e do conhecimento*. In: L. Alvares (org), *A organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações* p. 21-47. São Paulo.

Oliveira, L. P. (2009). Linguística de Corpus: teoria, interfaces e aplicações. *Matraga-Revista do Programa de Pós-Graduação em Letras da UERJ*, 16(24). doi: 10.12957/matruga

Rowley, J. E., & Hartley, R. J. (Eds.). (2008). *Organizing knowledge: an introduction to managing access to information*. Ashgate Publishing, Ltd.

Sardinha, A. P. B. (2001). Uso de corpora na formação de tradutores. *Delta*, 19, 43-70.

Sardinha, T. B. (2000). Corpus linguistics: history and problematization. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, 16(2), 323-367. Disponível em: <http://bit.ly/2sx8hld>

Schiessl, M., & Shintaku, M. (2012). *Sistemas de organização do conhecimento*. In: Alvares, L. (org), *A organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações* (49-118). São Paulo.

Silva, D. L. D., Souza, R. R., & Almeida, M. B. (2013). *Uma comparação de metodologias para construção de ontologias e vocabulários controlados*. Disponível em: <http://hdl.handle.net/123456789/965>

Silva, D. L. D., Souza, R. R., & Ramos, J. A. D. A. (2010). Representação da informação e do conhecimento em sistemas de informações: contribuições de campos interdisciplinares. Disponível em: <http://hdl.handle.net/123456789/914>

Standars, N. I. S. O. (2005). *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*.

Suassuna, J. (2011). Semi-árido: proposta de convivência com a seca. *Cadernos de Estudos Sociais*, 23(1-2).



OS PROCESSOS DE NEGÓCIO, A GESTÃO DE DOCUMENTOS E OS FLUXOS DOCUMENTAIS: ALGUMAS PERSPECTIVAS E RELAÇÕES

Livia Marangon Duffles Teixeira¹, Elisângela Cristina Aganette²

¹Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte – MG, <http://lattes.cnpq.br/5511349023940518>, liviamarangon@gmail.com

²Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte – MG, <http://lattes.cnpq.br/6224307970906388>, elisangelaaganette@gmail.com

RESUMO O presente artigo propõe uma discussão entre as temáticas processos de negócios, gestão de documentos e fluxo documental. O objetivo do diálogo está baseado na necessidade de i) esclarecer a relação entre o processo de negócios, o gerenciamento de documentos e o fluxo documental, e ii) explicar a relevância de considerar os documentos como portadores de conteúdo dinâmico, envolvendo seu próprio fluxo de aprovação e ciclo de vida muitas vezes não completamente ligados ao processo em que os mesmos documentos participam. Para tal, será elaborado um bem fundamentado arcabouço teórico que subsidiará a análise comparativa prática de dois modelos de processos de negócio, onde apenas um deles é contemplado pelos estudos documentais do processo. A construção teórica será capaz de alinhar diferentes áreas do conhecimento à Ciência da Informação, além de enriquecer a sua produção em relação à gestão de informação em relatos de casos. De forma prática, será capaz de demonstrar o impacto da ausência de estudos documentais em projetos de modelagem de processos e contribuir para a prática do profissional da informação.

PALAVRAS CHAVE *gestão de documentos, gestão de informações, gestão de processos de negócios, fluxo documental, procedimentos arquivísticos.*

ABSTRACT This article proposes a discussion between thematic business processes, document management and document flow. The purpose of the dialogue is based on the need to (i) clarify the relationship between the business process, document management and document flow, and (ii) explain the relevance of considering documents as having dynamic content involving their own flow of information. Approval and life cycle often not completely linked to the process in which the same documents participate. For this, a well-founded theoretical framework will be elaborated that will subsidize the practical comparative analysis of two models of business processes, where only one of them is contemplated by the documentary studies of the process. The theoretical construction will be able to align different areas of knowledge to Information Science, in addition to enriching its production in relation to information management in case reports. In a practical way, it will be able to demonstrate the impact of the absence of documentary studies in process modeling projects and contribute to the practice of the information professional.

KEYWORDS *Document management, information management, business process management, document flow, archival procedures.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Um processo de negócio consiste em um conjunto de atividades que se desdobram em uma instituição com o propósito de alcançar determinado resultado. Para ser considerado um processo, o mesmo necessita ter características essenciais em sua composição. Uma dessas características essenciais se refere à geração e à agregação de documentos que dão suporte às atividades presentes no respectivo processo, tanto no ambiente físico, quanto no eletrônico. Assim, pode-se inferir que gerenciar documentos e informações é reconhecidamente um pré-requisito necessário em projetos de gestão documental e de mapeamento e modelagem de processos. As práticas arquivísticas configuram-se como primeiros procedimentos executados em projetos de gestão documental e resultam em um inventário de tipologias documentais sobre um processo e ou atividade. As referidas práticas arquivistas podem se desdobrar em instrumentos como o plano de classificação e a tabela de temporalidade e destinação documental.

Atualmente, modelar processos de negócios tem sido demandado para alinhar as rotinas das organizações às práticas mais eficazes e eficientes de gestão, uma vez que a realidade econômica tem trazido para as organizações a necessidade de um novo planejamento para sua manutenção no mercado. Como forma de adequação ao novo contexto, a gestão de processos, *Business Process Management* (BPM), tem sido entendida como uma forma de otimizar a organização e seus recursos de forma estratégica. A representação dos processos de negócios da organização conduz à geração de modelos representativos por meio de diagramas e metodologias, como a *Business Process Model and Notation* (BPMN), que oferece uma linguagem comum entre todos os envolvidos na respectiva atividade de modelagem: *stakeholders*, analistas, administradores etc.

O que se observa nos modelos de notação gerados e que se configura um problema no momento de se propor melhorias para o processo como um todo é a identificação do item documental como um “artefato”. Essa característica culmina em um modelo que não contempla a correta identificação arquivística do documento, não mapeia o seu fluxo documental no processo e seu respectivo ciclo de vida. Ressalta-se portanto a necessidade da continuidade da modelagem dos fluxos e trâmites documentais vinculados aos processos de negócio, para se obter um processo documental otimizado e alinhado ao processo de negócio, resultando em um modelo de gestão de documentos e informações com um nível de maturidade adequado. Diante do exposto, o objetivo do presente artigo consiste em relacionar a metodologia de *mapeamento e modelagem de processos de negócio* à *gestão de documentos e à etapa de fluxo documental*, pelos pontos de vista teóricos e práticos, fundamentados por meio da análise e desempenho prático advindos de projetos de gestão por processos em instituições reais.

Processo, de acordo com Cruz (2011), refere-se a um conjunto de elementos que guiam grupos ou indivíduos entre o início e fim de um trabalho específico, ou ainda, um processo é um conjunto de atividades que se desenrolam dentro de uma corporação com o objetivo de atingir metas previamente definidas. Em um contexto organizacional, os processos, tanto aqueles referentes às atividades administrativas, por exemplo, a contabilidade, quanto às atividades de infraestrutura, tais como a limpeza, e as atividades-fim, a saber, em um hospital, um atendimento médico ou em uma usina hidrelétrica, a geração de energia precisam ser geridos de forma que mantenham seus critérios de qualidade e atendam aos requisitos dos seus clientes, sejam esses internos ou externos à organização (NBR ISO 9000:2015).

Krogstie *et al* (2008) apresentam cinco grandes grupos de utilização da modelagem de processos de negócios, a saber: i) criação de sentido e a comunicação humana; ii) análise assistida por computador; iii) gestão da qualidade; iv) implementação e ativação do modelo para integrá-lo em um sistema de informação; v) utilização do modelo como contexto para um projeto de desenvolvimento de sistema. Outrossim, uma gestão por processos permite ao tomador de decisão ter uma visão holística da organização e, com base no conhecimento sobre os mesmos, é possível focar na melhoria das atividades produtivas, vislumbrar o registro de conhecimento relacionado e ainda trabalhar para a sua automação.

Assim, uma solução de BPM, de acordo com Recker *et al* (2006), consiste no conhecimento e registro, de ponta a ponta, de todos os recursos que estão envolvidos no processo, com vistas ao negócio da organização. Considerando que uma das características essenciais dos processos é a presença de documentos que dão suporte às atividades e decisões, estejam esses tanto no ambiente físico quanto no eletrônico, evidencia-se que uma destas “pontas”, raramente contemplada na modelagem de processos, consiste na etapa de mapeamento do fluxo de documentos presentes nos processos de negócio. Vale ressaltar que o mapeamento de processos de negócio realizado por meio da linguagem padrão BPMN (OMG, 2011) identifica o tipo documental como um “objeto de dados, sem nenhum efeito sobre o processo” (Aagesen e Krogstie, 2015, p.227) - o artefato, desconsiderando seu próprio fluxo, seu ciclo de vida e os seus preceitos arquivísticos.

A arquivística é a área do conhecimento cujo objeto de pesquisa é o documento com característica probatória, contemplando seu ciclo de vida desde sua gênese à destinação final (Duranti, 1994). De acordo com Aganette, Teixeira e Almeida (2015) os preceitos arquivísticos aplicam seu potencial nas instituições por meio de seus aspectos, a saber: i) teóricos, como a teoria das três idades; ii) metodológicos, na criação dos instrumentos plano de classificação documental e tabela de temporalidade e destinação de documentos; iii) legais, vinculação probatória respaldada em lei; e iv) da prática arquivística em si, na operação da gestão. Somente com a implantação de metodologias de gestão de documentos nas organizações é possível identificar as tipologias documentais desde o momento da sua criação, mapear seu fluxo e trâmite no processo de negócio e gerir seu prazo de guarda para o contexto específico até o seu arquivamento ou eliminação. Conhecer os tipos e conteúdos permite participar na evolução de suportes (de documentos físicos para eletrônicos), na evolução das aprovações (de assinaturas físicas para eletrônicas ou aprovações em *workflows* em sistemas), racionalizar a guarda e trazer outros tipos de melhorias para as organizações.

Na atualidade, a rotina informacional das organizações tem transitado nos contextos físico e eletrônico, o que tem sido um dificultador para as organizações. Gerenciar e conciliar um legado de documentos em meio físico armazenados em arquivos com características probatórias e legais, com sistemas de informações e documentos em meio eletrônico tem desafiado pesquisadores, profissionais da informação e profissionais que atuam no mercado de tecnologia da informação. A gestão de documentos e informações configura-se, reconhecidamente, um pré-requisito para projetos de mapeamento e modelagem de processos. Dessa forma, não há outro caminho a não ser o de alinhar processos de negócios, procedimentos arquivísticos, estratégias corporativas, *compliance*, cultura organizacional, arquitetura tecnológica corporativa, patrocínio de gestores, gestão de mudanças, normatização e demais áreas do conhecimento envolvidas, como a Ciência da Informação. No entanto, não se trata de uma solução simples e rápida de se implementar em um momento de constantes transformações, mas é a mais segura para se criar uma estrutura real e executável, e que deve ser considerada em projetos de automação dessa natureza.

Diante do diálogo teórico e pelo ponto de vista das áreas do conhecimento envolvidas, a análise de um caso real por seguinte evidencia o problema apontado e sugere que a proposta indicada em relação à inserção de uma etapa de análise e mapeamento dos fluxos documentais solucione os *gaps* identificados. E ainda, por meio do conhecimento gerado, estabelece-se uma análise fundamentada do caso em relação aos projetos de gestão de processos de negócio em instituições de forma prática e real.

Os resultados do presente artigo refletem em contribuições práticas e teóricas pertinentes e valiosas para o campo da Ciência da Informação, diretamente vinculadas à gestão de documentos, à modelagem de processos, e ainda, à especificação de requisitos para desenvolvimento e modelagem de sistemas.

METODOLOGIA

Hammer (2013) apresenta a origem da gestão de processos sob a ótica de seus dois antecessores: o *controle estatístico de processos*, que originou o movimento da qualidade e a *reengenharia de processos de negócios*, sendo que esse último trouxe uma definição aprimorada de processo e desenho do processo. Essa condução evolutiva pode ser observada acompanhando as mudanças econômicas, produtivas e sociais, sob as quais as organizações estão contextualizadas. O cenário econômico da atualidade tem demandado, das organizações, mais eficiência e eficácia na condução e execução dos seus processos. Ou seja, aumentar a produção, executando as atividades da forma correta e ao mesmo tempo demandando menos recursos (pessoas, materiais e finanças) é um objetivo que somente por meio da gestão de processos pode-se alcançar.

A ausência de estudo dos documentos que participam dos processos de negócios acaba por pouco melhorar a forma como os processos são executados. O que se observa na prática são processos documentais fluindo da mesma forma como antes, uma vez que esses documentos são portadores de conteúdo dinâmico e dos registros de aprovações (como carimbos e assinaturas). Assim, continua-se gerando, imprimindo e assinando formulários físicos, acumulando documentos e registros e, por consequência, arquivando volumes cada vez maiores de papéis e mantendo os antigos problemas de recuperação de documentos e informações.

Além disso, os tomadores de decisões, muitas vezes impulsionados por ondas de inovações tecnológicas, optam pela automação como solução aos seus problemas em processos, culminando o momento de euforia inicial em grande frustração pelo não alcance dos resultados inicialmente esperados. Outrossim, novos projetos de automação de processos de negócios têm surgido e trazido consigo desafios de romper com resquícios de projetos anteriores mal sucedidos, que tinham como foco principal a tecnologia.

Do ponto de vista acadêmico, em especial da Ciência da Informação, observa-se pouca produção acadêmica acompanhando o desenvolvimento exponencial do mercado de soluções tecnológicas para gestão de documentos e informações. Poucos pesquisadores têm dedicado algumas de suas pesquisas de forma a gerar conhecimento estruturado e sistematizado capaz de subsidiar tecnicamente o assunto, como evidenciado por Bax *et al* (2016). Outrossim, trazer para o contexto científico a confluência dos assuntos configura uma contribuição para a construção de conhecimento por meio da elaboração de uma base teórica científica que sustente práticas demandadas pelo mercado.

Assim sendo, a metodologia de pesquisa utilizada para se atingir os objetivos do presente artigo consistiu em, no campo teórico, consolidar o tema “processo” e “gestão de processos” e seus relacionamentos de forma fundamentada. Na mesma perspectiva, trazer os conhecimentos de práticas

arquivísticas, da gestão de documentos ativos e fluxos documentais. Assim, estabelecendo, consolidando e relacionando sob a perspectiva da Ciência da Informação as temáticas envolvidas, é constituído um arcabouço teórico, justificando-as como área do conhecimento impactante para o referido estudo e para a Ciência da Informação.

A metodologia de pesquisa, no ponto de vista prático, consiste em comparar um projeto de modelagem de processos na prática em que se inseriu a análise documental como etapa do projeto e outro projeto de modelagem de processos na prática em que NÃO se inseriu a análise documental como etapa do projeto. Assim, existe a expectativa de criação de subsídio para a justificativa de que, para a finalidade de obter modelos de nível de qualidade superior, é preciso projetar o fluxo documental juntamente com o processo de negócios, além de evidenciar a importância da atuação do profissional da informação neste mercado de trabalho.

Para tal, dois modelos de processos foram extraídos de casos reais, sendo um em que se contemplou a análise documental como etapa do projeto de modelagem e outro não considerou a referida etapa. Para esse estudo consideraram-se tanto para o levantamento quanto para a análise, os produtos finais dos referidos projetos e seus respectivos resultados práticos.

MODELAGEM DE PROCESSOS EM QUE SE CONSIDEROU A ANÁLISE DOCUMENTAL COMO ETAPA DO PROJETO

O processo a ser analisado refere-se a uma transação contábil, modelado em um Projeto de Pesquisa & Desenvolvimento (P&D). O Projeto, de característica multidisciplinar, foi realizado através de parceria entre Universidade Federal de Minas Gerais (UFMG) e Companhia Energética de Minas Gerais (CEMIG), com financiamento da Agência Nacional de Energia Elétrica (ANEEL). As autoras do artigo lideraram e conduziram os estudos documentais, arquivísticos e as configurações de ambos na tecnologia.

O diagnóstico inicial, realizado por meio de entrevistas e pela geração do modelo “*as is*”, foi capaz de evidenciar elevado número de impressões e cópias xerográficas, além da redundância de formatos (físico original, físico cópia, cópia digitalizada, microfilme). Essa duplicidade de formatos e suportes não resultava em um processo de recuperação de documentos e informações ágil, eficiente e eficaz, além de agregar altos custos associados (impressão, cópias, armazenagem e consulta).

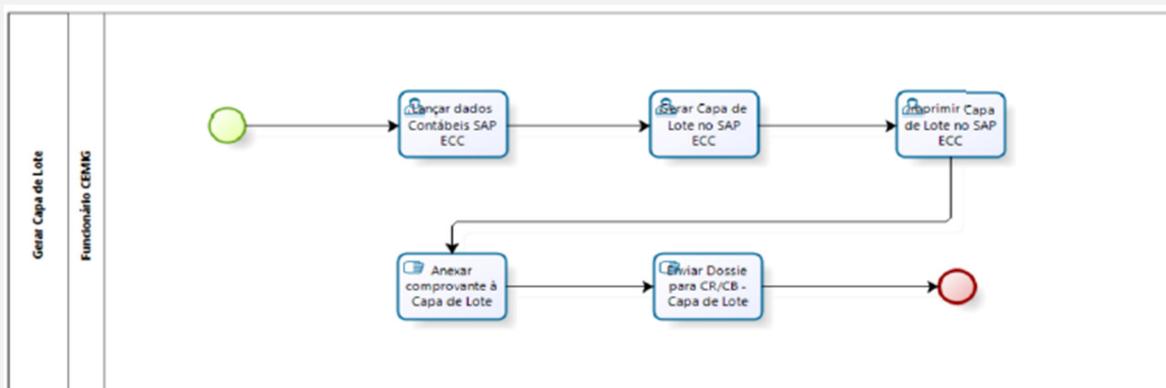


Figura 1. Modelagem “*as is*” do processo

Para que pudesse se alcançar a melhoria do processo de negócio apresentado, foram realizadas reuniões e oficinas com os seus usuários-chave, com o propósito de gerar de um modelo “to be” do mesmo.

A análise do processo documental associado ao processo de negócio configurou-se como o ponto-chave do Projeto. Através do estudo dos tipos de dossiês documentais no sistema eletrônico de gestão financeira utilizado na Empresa, vinculado ao estudo dos itens documentais que constituem esses dossiês por práticas arquivísticas, permitiu-se otimizar a gestão do processo. As otimizações foram refletidas na redução de 126 para 37 tipos de dossiês financeiros. Desses 37 vigentes, apontou-se que 11 deles dispensam outros tipos documentais anexados. O levantamento de tipos de documentos foi de 92 iniciais para 93 após o estudo. Na elaboração dos instrumentos de gestão arquivística, identificou-se que aproximadamente 75% dos tipos de documentos são nativos digitais.

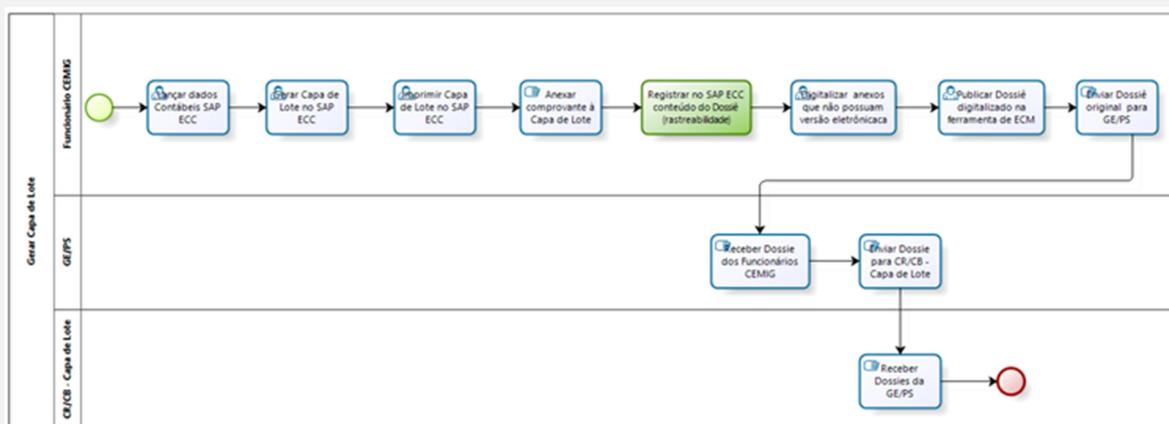


Figura 2. Modelagem “to be” do processo

Com o estudo do processo de negócio e o estudo do processo documental relacionado, a tecnologia foi capaz de absorver a necessidade de impressão ao manter o formato digital original dos documentos (como, por exemplo, a nota fiscal eletrônica), permitiu capturar a imagem digitalizada do original físico e associar através dos metadados de forma automatizada ao dossiê a que pertence além de indicar seu correto arquivamento físico, disponibilizar atendimento de rotina ou auditorias internas e principalmente as externas através da consulta direta à imagem através da ferramenta *web*. Ainda associado às melhorias do processo, foi possível calcular que, em cinco anos, a economia é significativamente relevante (estimam-se milhões de reais) ao deixar de imprimir o desnecessário e gerar cópias, ao eliminar o armazenamento físico em redundância ou por tempo excessivo, ao reduzir custos para as consultas físicas e na otimização de equipe envolvida em todo o processo.

MODELAGEM DE PROCESSOS EM QUE NÃO SE CONSIDEROU A ANÁLISE DOCUMENTAL COMO ETAPA DO PROJETO

O processo a ser analisado refere-se a uma Requisição Externa de Compras, modelado em um Projeto de Mapeamento e Modelagem de Processos. O Projeto, de caráter consultivo, foi realizado pelas autoras deste artigo, em uma instituição pública de direito público, sem fins lucrativos, a ICismep - Instituição de Cooperação Intermunicipal do Médio Paraopeba. A referida instituição gerencia o dinheiro público direcionado à saúde de forma cooperativa, mediante otimização dos processos de negócio. A sede da instituição localiza-se no Município de Betim, em Minas Gerais, com desenvolvimento de atividades de saúde em vários outros Municípios consorciados.

O diagnóstico inicial fora realizado por meio de entrevistas com os funcionários de cada área e pela geração do modelo “*as is*”, neste caso apenas na versão descritiva de processos, ou seja, não gerou-se o desenho dos fluxos, evidenciando-se a necessidade urgente de reestruturação dos processos de negócio.

Dessa forma, para a geração do modelo “*to be*” dos processos de negócio mapeados, realizaram-se reuniões com os usuários-chave de cada área escopo. A versão “*to be*” dos referidos processos permitiu a revisão e reestruturação dos mesmos, a identificação e correção de *gaps* na rotina de atividades, bem como possibilitou a reestruturação dos processos que não estavam em consonância com o planejamento estratégico da Instituição. O trabalho permitiu ainda a revisão do organograma corporativo, a otimização e potencialização dos recursos intelectuais atuantes, o que culminou na redução do quadro de pessoal.

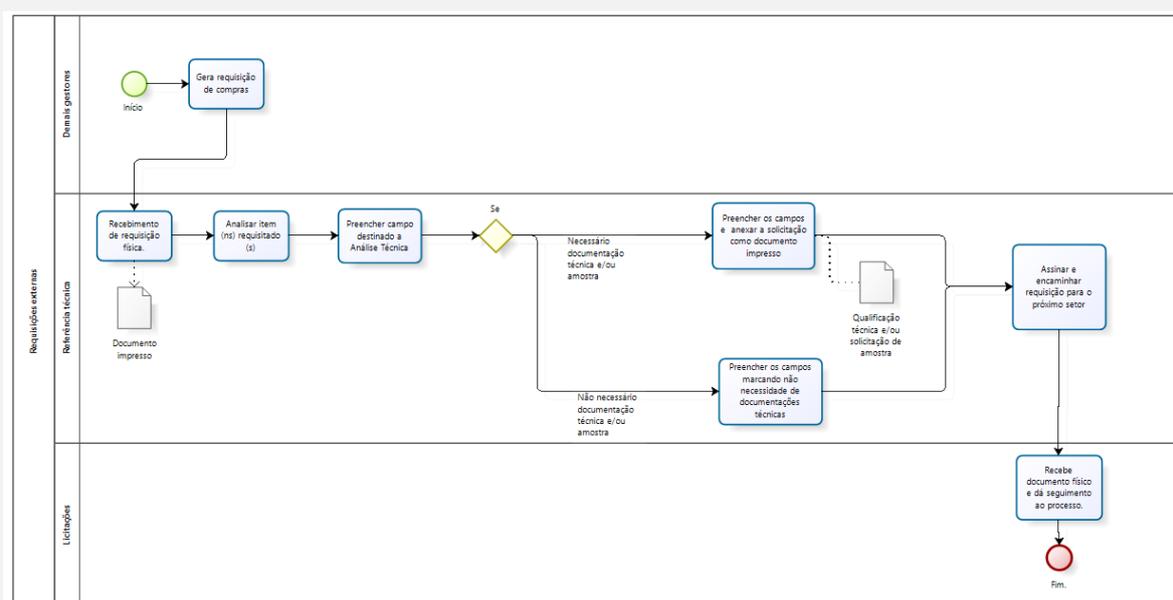


Figura 3. Modelagem “*to be*” do processo: Requisição Externa de Compras

No entanto, como mencionado anteriormente, a etapa de análise documental associada ao processo de negócio não foi considerada no momento do mapeamento e modelagem dos processos de negócio, uma vez que a referida etapa não fazia parte do escopo contratado para execução do Projeto. Esse fato se configurou como um ponto falho, uma vez que não houve o devido estudo dos tipos de dossiês documentais e seus respectivos fluxos. Diante do contexto, cuja proliferação da produção de documentos e conseqüente acúmulo desordenado aponta como uma problemática também comum a outras instituições, a ausência da gestão dos fluxos documentais ocasiona bastante insegurança quanto ao uso dos documentos como insumo para tomadas de decisões, além de deixar exposta a Instituição em relação a atendimento de fiscalizações, auditorias e outras demandas probatórias.

RESULTADOS

Espera-se, com a conclusão deste trabalho colaborativo, contribuir tanto pelo ponto de vista teórico quanto do ponto de vista prático para o campo da Ciência da Informação.

Teoricamente, confere-se uma caracterização bem fundamentada do impacto dos estudos documentais alinhados à modelagem de processos, inserindo novos conhecimentos à Ciência da Informação e, ao mesmo tempo, levando a Ciência da Informação para outras áreas do conhecimento.

De forma prática, o trabalho trouxe subsídios para a análise e desempenho de projetos de gestão de processos por meio da abordagem de gestão de documentos e informações, especificamente por meio da inserção da análise documental como etapa imprescindível em projetos de modelagem de processos. Essa mudança e padronização torna-se cada vez mais necessária, uma vez que a informação, por ser considerada insumo de valor e importância cada vez maior, faz-se necessária em qualquer contexto organizacional, seja como um recurso interno e ou externo. Assim, torna-se evidente que, quanto mais fidedigna, segura, acessível e com um fluxo contínuo estas informações estarem inseridas, maior facilidade de gestão dos processos de negócio e conseqüentemente maior competitividade, as organizações terão. Para isso, é preciso que se adote a prática da análise documental, para um correto fluxo informacional aliada a modelagem de processos e às tecnologias para minimizar deficiências nos processos estratégicos advindos da falta de informações, enquanto essas se acumulam de forma desordenada.

Rezende e Abreu (2003) afirmam a necessidade de um cuidado especial na elaboração e produção de documentos, criando-se um padrão normativo para a produção documental no âmbito organizacional, adequado às necessidades de fluxos informacionais da instituição. Padrão normativo esse, crucial para uniformização dos fluxos documentais nos sistemas de gestão documental, facilitando o compartilhamento, uso e reuso de informações nos diversos setores e processos de negócio da instituição e uniformizando os procedimentos referentes à gestão documental. Diante do exposto, considera-se ainda que os estudos documentais permitem uma reflexão sobre a real necessidade dos tipos de documentos no processo, sobre a possibilidade de automação do documentos e do seu ciclo de aprovação (como formulários eletrônicos associados a *workflow*).

Outrossim, é notória a necessidade de definir modelos, métodos e técnicas que possibilitem às instituições gerenciar os fluxos documentais de forma ágil e segura de forma associada aos processos de negócios.

CONCLUSÕES

A gestão por processos é a metodologia através da qual as organizações se mantêm ativas diante a concorrência que, ao contrário de 30 anos atrás, não é apenas local. Seja por sistema de gestão da qualidade (Norma ABNT ISO), seja pelo MEG (Modelo de Excelência de Gestão), através da Acreditação Hospitalar ou outras metodologias em qualquer tipo de instituição, os processos estão sempre sendo submetidos a ciclos de conformidade e melhorias. Uma evolução natural desses ciclos, ao se alcançar a maturidade, pode ser a automação desses processos, onde se buscam formas de tratativas eletrônicas/digitais de tornar as organizações mais eficientes. Essa evolução nem sempre reflete em menos burocracia ou dependência de atividades manuais (como vistos, assinaturas e carimbos), pois o ciclo de vida dos documentos, via de regra, não acompanha o mapeamento do ciclo de vida dos processos.

Todas as organizações necessitam de algum tipo de informação ou conhecimento para dar continuidade aos processos, sejam essas simples atividades diárias ou mais complexas tomadas de decisão. Dessa

forma, existe a demanda constante por informações organizadas e padronizadas de forma a proporcionar uma recuperação simples e fácil. Esse é o “negócio” do profissional da informação.

Não existe caminho alternativo à modelagem de processos a fim de que se obtenham conhecimento e visão crítica da organização para então culminar em melhorias. A visão hierárquica, a visão funcional, as políticas, a cultura, as regras de negócios, os sistemas legados, as tecnologias vigentes, os documentos, os metadados, os arquivos... é necessário identificar e estudar as instâncias que impactam e modificam a forma como as pessoas (o principal recurso) trabalham para que os projetos não fracassem. É necessário representar essas instâncias na modelagem do processo de forma real, conferindo a elas a sua funcionalidade. As instâncias precisam cumprir seu papel, ou seja, a sua função, e mesmo que seja automatizado o processo e a forma de se executar seja transformada, deve ser garantida a segurança, a rastreabilidade e a conformidade do processo como um todo.

REFERÊNCIAS BIBLIOGRÁFICAS

- Aagesen, G.; Krogstie, J. (2013). Análise e estrutura de processos de negócios utilizando BPMN. In: Brocke, J. V.; Rosemann, M. *Manual de BPM: gestão de processo de negócio*. Porto Alegre: Bookman. cap.6
- Aagesen, G.; Krogstie, J. (2015). BPM 2.0 for Modeling Business Process. In: Brocke, J. V.; Rosemann, M. *Handbook on Business Process Management 1*. 2.ed. Berlim/Heidelberg/Nova York: Springer-Verlag, p. 219-250.
- Aganette, E. C.; Teixeira, L. M. D.; Almeida, M. B. (2015). A prática arquivística e a gestão do conteúdo organizacional: relato de experiência em empresa do setor energético. *Tendências da Pesquisa Brasileira em Ciência da Informação*, v. 8, n. 1.
- Aganette, E. C.; Teixeira, L. M. D.; Aganette, K. J. P. *Modelagem de processos em empresa do setor de saúde pública: i-CISMEP, um relato de caso*. Artigo no prelo
- Associação Brasileira de Normas Técnicas. (2015). *NBR ISO 9000:2015 – Sistema de gestão da qualidade – fundamentos e vocabulário*.
- Bax, M. P.; Ferreira, L. G.; Teixeira, L. M. D.; Teixeira, A. M. F. (2016). *Apresentação: Gestão de conteúdo corporativo: uma análise conceitual*. In: X EDICIC, Belo Horizonte.
- Bax, M. P.; Ferreira, L. G.; Teixeira, L. M. D. (2017). *Gestão de conteúdo corporativo: apontamentos teóricos e práticos*. Artigo aceito para publicação.
- Cruz, T. (2011). *Sistemas, organização e métodos: estudo integrado das novas tecnologias de informação*. São Paulo: Atlas.
- Cruz, T. (2015). *Sistemas, métodos & processos: administrando organizações por meio de processos de negócios*. Rio de Janeiro: Atlas.
- Duranti, L. (1994). Registros documentais contemporâneos como provas de ação. *Estudos Históricos*, v. 7, n. 13, p. 49-64.
- Fundação Nacional da qualidade. (2011). *Crêterios de excelência*. 19. ed. São Paulo: FNQ.

Hammer, M. (2013) O que é gestão de processos de negócios?. In: Brocke, J. V.; Rosemann, M. *Manual de BPM: gestão de processo de negócio*. Porto Alegre: Bookman, cap.1.

Krogstie, J.; Dalberg, V.; Jensen, S. M. (2008). Process modeling value framework, enterprise information systems. In: Manolopoulos, Y.; Filipe, J.; Constantopoulos, P.; Cordeiro, J. (eds.) *Lecture notes in business information processing*. Heidelberg: Springer-Verlag, v.3.

Oliveira, R. B. C. de. (2009). *Uma metodologia de modelagem de processos de negócio orientada a gestão da informação e do conhecimento*. 2009. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte.

OMG. (2011). *BPMN 2.0 Specification: business process model and notation (BPMN) - version 2.0, technical report*. Recuperado a partir de: <<http://www.omg.org/spec/BPMN/2.0/examples/PDF>>.

P&D CEMIG GT507. *Modelagem dos processos de negócios do P&D GT507*. (2014) Belo Horizonte: Documentos do P&D GT507.

Recker, J., Indulska, M.; Rosemann, M.; Green, P. (2006). How good is BPMN really? Insights from theory and practice. In: *Anais...14 ECIS - European Conference on Information Systems*, p. 1-12. Recuperado a partir de: <<http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1075&context=ecis2006>>.

Rezende, D. A.; Abreu, A. F. (2000). *Tecnologia da informação*. São Paulo: Atlas.

Zaidan, F. H. (2015). *Aportes da arquitetura corporativa para o ambiente dos sistemas informatizados de gestão arquivística de documentos: aplicação em companhia de energia elétrica*. Tese (Doutorado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2015. Recuperado a partir de: <<http://www.bibliotecadigital.ufmg.br/dspace/handle/1843/BUBD-A6KGZG>>.



RELAÇÕES SEMÂNTICAS EM FOLKSONOMIA NO DOMÍNIO DA DIABETES

Jacqueline A. Souza

*Faculdade de Letras da Universidade do Porto/ Univerisidade Federal do Rio Grande do Norte,
0000-0001-5677-8282, jackebci@gmail.com*

RESUMO A folksonomia, apesar de apresentar limitações decorrentes do uso da língua natural e a falta de controle do vocabulário, detém um conhecimento útil para aprimorar o tratamento da informação em contexto digital. Como uma das vantagens, por meio de conceitos e relacionamentos semânticos implícitos, oferece contributos para a estruturação semântica. Ademais, investigar sobre relações semânticas, comumente compreendidas como hierárquicas, de equivalência e associativas, permite apurar as representações e configurações conceituais sobre diversos domínios de conhecimento, de forma aprimorar processos que visam à recuperação da informação. Neste sentido, o objetivo deste trabalho, desenvolvido no âmbito do doutoramento, é discutir acerca das relações semânticas implícitas nas folksonomias e identificar um núcleo de relações semânticas, sobretudo relações associativas, no domínio da diabetes. Para consecução desta investigação de caráter exploratório, adotou-se como abordagem metodológica a análise de domínio e a análise conceitual, assim como a utilização do vocabulário estruturado e trilingue DeCS como referência para a análise. Para coleta das tags, totalizando 600 indexações livres, foram explorados os sistemas baseados em folksonomia "LibraryThing" e "Cituelike". Finalmente, expor a variedade de relações semânticas e suas propriedades provenientes das folksonomias contribuiu para compreender operações intelectuais, consequentemente delinear uma área.

PALAVRAS-CHAVE *Organização da informação, Folksonomia, Semântica, Relações associativas.*

ABSTRACT Although folksonomy presents limitations due to the use of natural language and lack of vocabulary control, it has useful knowledge for improving the treatment of information in a digital context. As one of the advantages, through implicit semantic concepts and relationships, is in offering contributions to the semantic structuring. In addition, investigating semantic relations, commonly understood as hierarchical, equivalence and associative, allows the determination of conceptual representations and configurations on several knowledge domains, in order to improve processes that aim at information retrieval. In this sense, the objective of this work, developed within the scope of the doctorate, is to discuss the semantic relations implicit in folksonomies and to identify a nucleus of semantic relations, mainly associative relations, in the area of diabetes. In order to achieve this exploratory research, domain analysis and conceptual analysis were adopted as the methodological approach, as well as the use of the structured and trilingual vocabulary DeCS as a reference for the analysis. To collect the tags, totaling 600 free indexes, the systems based on folksonomy "LibraryThing" and "Cituelike" were explored. Finally, exposing the variety of semantic relations and their properties from folksonomies contributed to understanding intellectual operations, hence outlining an area.

KEYWORDS *Information organization, Folksonomy, Semantic, associative relationships.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

O diabetes é uma doença metabólica crônica caracterizada por níveis elevados de glicose no sangue (ou açúcar no sangue) que, ao longo do tempo, causa danos graves ao sistema vascular, renal, oftalmológico. A mais comum é a diabetes tipo 2, adquirida geralmente por adultos, e caracterizada pela resistência à insulina ou a sua produção insuficiente. O diabetes tipo 1, conhecida também como diabetes juvenil, é uma condição crônica em que o pâncreas produz pouca ou nenhuma insulina por si só. Hoje constitui um dos maiores desafios da saúde pública, sendo considerada uma epidemia e está relacionada a um conjunto de fatores de ordem genética, comportamental, social (OMS, 2017).

A sua prevalência continua crescendo, devido ao envelhecimento populacional, desenvolvimento econômico e urbanização que desencadearam mudanças importante no estilo de vida das pessoas, marcado pela presença de sedentarismo e obesidade. Trata-se de uma das doenças crônicas mais importantes e impactantes para o sistema de saúde pública, devido ao elevado grau de morbimortalidade e aos altos custos para o controle metabólico e tratamento de suas complicações microvasculares. Estima-se que a diabetes tipo 1, que atinge menos de 5% da população mundial com diabetes, tem vindo a aumentar ligeiramente em idades cada vez mais jovens. Entretanto, é o diabetes tipo 2 que apresenta a maior prevalência, em mais de 400 milhões de pessoas, segundo a OMS (2017).

Neste seguimento, a OMS desenvolve um programa para prevenir o diabetes, assim como busca minimizar as complicações e maximizar a qualidade de vida dos doentes. O objetivo é melhorar a saúde por meio de estímulos e apoio à adoção de medidas efetivas para vigilância, prevenção, e controle de diabetes e suas complicações, sobretudo em países de baixa e média renda. Com base neste objetivo, focaliza supervisionar e desenvolver normas internacionais, promover e contribuir para vigilância do diabetes, suas complicações e fatores de risco, aumentar a conscientização sobre este problema global de saúde pública e atuar na prevenção e controle em populações vulneráveis.

Neste cenário complexo, cujo impacto econômico e social é inquestionável, em busca de estreitar ações no âmbito da informação para e sobre a saúde, circunscrevem-se os contributos da Organização da Informação (OI), enquanto “processo que envolve a descrição física e de conteúdo dos objetos informacionais, no qual o produto desse processo descritivo é a representação da informação, as quais são construídas por meio de linguagens elaboradas especificamente para atender aos objetivos da OI” (Medeiros & Café, 2013).

Concernente ao desenvolvimento de novos esquemas de representação, ressalta-se que a Web 2.0 possibilitou a personalização tecnológica, abrindo novas possibilidades para o desenvolvimento de produtos e serviços agregados, assim como ampliou a presença do utilizador comum na coordenação de ações de produção, organização e difusão de informações voltadas a públicos segmentados pela natureza das redes sociais aos quais pertencem e, é neste contexto que as folksonomias contribuíram enormemente para a popularização de novas perspectivas de classificação de documentos digitais e ampliaram as possibilidades de compartilhamento de novas significações para termos e conceitos socialmente estabelecidos e debatidos em ambientes virtuais. (Moura, 2009).

Complementarmente, pesquisar sobre as relações semânticas que compõem parte de um universo do conhecimento, sobretudo no domínio da saúde, permite apontar a redução das diferenças entre as representações e configurações conceituais apresentadas por sistemas de informação e as representações mentais do conhecimento especializado de seus utilizadores. A perspectiva de ampliação da dimensão

semântica de sistemas artificiais pode ainda contribuir para uma maior aproximação entre atores humanos e artificiais, otimizando a interação e o diálogo entre ambos. As relações semânticas estabelecidas entre as unidades de conhecimento, os conceitos, auxiliam nos processos de raciocínio e na tomada de decisão, e os sistemas de informação têm suas potencialidades ampliadas (Biolchini, 2001).

Neste sentido, o objetivo deste trabalho é discutir acerca das relações semânticas implícitas nas folksonomias e identificar um núcleo de relações semânticas, sobretudo relações associativas, no domínio da diabetes.

FOLKSONOMIA NA ÁREA DA SAÚDE

Folksonomia (ou *folksonomy*) refere-se à junção dos termos *folk* (povo, pessoas) com *taxonomy* (classificação) e o conceito mais difundido é o elaborado por Vander Wal (2007):

the result of personal free tagging of information and objects (anything with a URL) for one's own retrieval. The tagging is done in a social environment (usually shared and open to others). Folksonomy is created from the act of tagging by the person consuming the information (Vander Wal, 2007).

Constituída pela tríade: utilizador, conteúdo e tag, suas limitações e desvantagens referem-se à baixa precisão no momento da busca, uma vez que uma mesma tag pode assumir diferentes significados, termos imprecisos, diferenças linguísticas, presença de erros ortográficos, manifestação dos fenômenos da língua natural (polissemia, ambiguidade, sinonímia). Porém, entre suas vantagens, destaca-se o baixo custo para indexar grandes coleções, a serendipidade e a garantia do uso de descritores.

Peters (2009) expõem a diferente abordagem das folksonomias tocante a classificação e estrutura de recursos informacionais em ambientes digitais:

Instead of choosing a classification criterion and filling it with resources, it is now the resources that are allocated the criteria. Folksonomies turn the classification system from a criteria-centric into a resource-centric approach. This means that multiple storage no longer refers to the resources but to the multiple allocations of the "folders", "drawers" or "shelves" that are the tags of the folksonomy. Pinned to the information resources are as many tags as are necessary to adequately describe and retrieve them. Thus tags enable the most diverse criteria to be allocated to the resources and in this way guarantee a much broader access to them, which, due to the collaborative construction of the folksonomy, is also independent of the guardian. In the digital world, however, this approach always requires an indexing and retrieval system to render the folksonomy-based classification system manageable. The user may have created numerous access paths to the information resources, but a system will be needed to aggregate the tags and so provide links to the desired resources. So in order to structure and classify resources, folksonomies sidestep onto a meta-level, which represents the resource via (a whole lot of) tags. (Peters, 2009, p.3)

À vista disto, a autora ressalta que a abordagem focada no termo, no uso de vocabulário controlado e realizada por profissionais treinados, conseqüentemente a recuperação da informação independe de sua representação via linguagem.

Certamente, desde o seu advento, a folksonomia vem fomentado muitas questões acerca dos aspectos sobre colaboratividade, compartilhamento, inteligência coletiva, sua utilidade e comparação em relação

às técnicas tradicionais de representação da informação, sobre a qualidade, tipo de linguagem, termos ou descritores que são utilizados, seu nível de especificidade. Não obstante a essas discussões, a título de elucidação, convêm apresentar brevemente como a folksonomia tem sido aplicada na área da saúde. Assim, seguem descritos alguns casos:

- O PatientsLikeMe é uma comunidade de redes sociais online e seu público são pacientes com 3 diagnósticos: esclerose lateral amiotrófica, esclerose múltipla e doença de Parkinson. Os membros da comunidade podem descrever seus sintomas online em linguagem natural, resultando em tags populares disponíveis para análise clínica e para procurar por outros utilizadores e pacientes. A iniciativa comprovou os desafios da comunicação online dos pacientes, não apenas dos profissionais. Também, reconheceu o potencial da folksonomia para reduzir barreiras entre consumidores, pacientes e especialistas (Frost & Massagli, 2008).
- O artigo produzido por Linehan, Doughty, Lawson, Kirman, Olivier & Moynihan (2010) descreve o projeto e a avaliação inicial do Tag-liatelle, um aplicativo de marcação colaborativa para incentivar a alimentação saudável. Os utilizadores fotografaram suas próprias refeições e carregam essas fotos em um sítio, onde outros utilizadores fazem a etiquetagem do conteúdo anonimamente. Os resultados sugeriram que a marcação do conteúdo referente a alimentação é uma atividade muito popular e que há possibilidade de extrair informações nutricionais por meio das tags geradas.
- A Agência de Pesquisa e Qualidade da Saúde dos Estados Unidos (*U.S. Agency for Healthcare Research and Quality*) criou um site público para divulgar informações sobre suas iniciativas em tecnologia da informação em saúde. No último projeto de melhoria de qualidade contínua, o NRC (National Resource Center) usou os *logs*¹ de busca do site para extrair frases de pesquisa geradas pelo utilizador. As frases foram então comparadas ao vocabulário controlado do site em relação à linguagem, gramática e precisão de busca. Os resultados da comparação demonstram que os dados do registro de pesquisa podem ser uma maneira econômica de melhorar os vocabulários controlados, bem como a recuperação da informação. As frases de pesquisa inseridas pelo utilizador também encontraram muitas semelhanças com tags (Dixon & McGowan (2010).
- Kristine Adland e Lykke (2012) relatam a pesquisa que o objetivo foi explorar como a marcação social poderia ser útil em um site de informações para pacientes com câncer e seus familiares. Verificou-se o potencial das tags para descrever e fornecer acesso ao conteúdo do site, assim como elas podem melhorar a interação e a comunicação entre utilizadores leigos e especialistas de domínio em um site de informações dentro dos cuidados de saúde.

Certamente, as características da terminologia da área da saúde podem dificultar a criação de taxonomias, entretanto, os exemplos mencionados reforçam e maximizam a aplicabilidade das folksonomias na área da saúde.

¹ Log de dados é uma expressão utilizada para descrever o processo de registro de eventos relevantes em um sistema computacional.

RELAÇÕES SEMÂNTICAS: CONCEITO E PROPRIEDADES

Sob o prisma da semântica lexical, uma vertente dos estudos semânticos que trata do significado cognitivo que envolve a relação entre a língua e os constructos mentais que representam ou estão codificados no conhecimento semântico do falante (Cançado, 2013), compreende-se por relação semântica, a relação significativa entre conceitos, os quais não existem de forma independente uns dos outros, isoladamente (Khoo & Na, 2006; Storey, 1993). Assim, Khoo e Nah (2006) apresentam a seguinte definição:

Semantic relations are meaningful associations between two or more concepts, entities, or sets of entities. They can be viewed as directional links between the concepts/entities that participate in the relation. The concepts/entities are an integral part of the relation as a relation cannot exist by itself. Associations between concepts/entities can be categorized into different types, abstracted, conceptualized and distinguished from other associations, and can thus be assigned meaning. The meaning or type of an association can sometimes but not always be derived from the meanings of the concepts involved. Psychologists and philosophers have attempted to identify the main types of relations and their features. [...] Semantic relations can refer to relations between concepts in the mind (called conceptual relations), or relations between words (lexical relations) or text segments. [...] semantic relations can occur at higher levels of text-between phrases, clauses, sentences, and larger text segments, as well as between documents and sets of documents. (Khoo & Nah, p.4, 2006).

Este excerto, esclarece que as relações semânticas podem ser analisadas sobre distintas perspectivas e em várias áreas do conhecimento. Os autores, fundamentados em Chaffin e Heernann (1988) e Chaffin (1992), ao discutirem se as relações semânticas são conceitos, enfatizam que elas possuem as principais características dos conceitos, destacando quatro pontos, a saber:

- 1) As relações podem ser analisadas em elementos ou características mais básicas
- 2) Uma nova relação pode ser uma combinação de outras relações;
- 3) As relações têm estrutura graduada (algumas instâncias da relação, representada por pares de palavras, são mais típicas de uma relação particular do que outras);
- 4) As relações variam na facilidade com que podem ser expressas.

Relativamente a tipologia, comumente as relações semânticas podem ser denominadas e divididas em sintagmáticas e paradigmáticas (Storey, 1993; ISO 25964; Stock, 2010; Peters, 2009). As relações sintagmáticas são os relacionamentos que existem apenas porque os conceitos ocorrem juntos no contexto de um documento particular. Quando dois ou mais termos de indexação são atribuídos ao mesmo documento, este ato efetivamente registra uma relação sintagmática entre eles (ISO 25964). Enquanto as paradigmáticas incluem as relações hierárquicas, de equivalência e associativas.

As relações hierárquicas expressam a subordinação entre dois termos em função de sua significação. Ou seja, quando um termo genérico engloba conceitualmente o outro, no qual a significação do termo genérico sempre engloba o termo específico. Há também as relações partitivas, as quais se ocorrem entre o todo e suas partes. Comumente, refere-se a um objeto ou organização social, política, administrativa (ANSI/NISO Z 39.19, 2005; Simões, 2008).

A relação de equivalência, com a finalidade de controlar a sinonímia, ocorre quando há dois ou mais termos que expressam o mesmo conceito e que são traduzidos por um mesmo descritor e este tipo de sinonímia aparece em termos que apresentam o mesmo significado (Simões, 2008).

Relativamente às relações associativas, elas se estabelecem entre os termos próximos entre si, os quais representam conceitos que podem ocorrer mutuamente por associação de ideias, uma vez que não são equivalentes, nem apresentam entre eles uma relação hierárquica. Neste sentido, esta é uma relação que pode ter uma natureza muito diversificada ao estar embasada na associação de ideias e, conforme Maldonado-Martínez (2010, p. 185) “*Se puede asociar entes com processos, processos com objetos, objetos com propriedades, etc. La associativa es um tipo de relacion muy aberta*”.

Todos os termos que constitui um sistema são passíveis de associação mental. Quando existe uma relação associativa entre um termo com um outro termo de grau superior hierárquico, esta é extensiva aos termos que de si dependem. Na prática, não é necessário estabelecer ligações entre todos os termos afins. Isso ocorrerá quando se observar uma sobreposição de sentido. Outro ponto relevante das relações associativas é que elas fornecem pistas sobre assuntos afins por meio dos termos relacionados, o que permite ao utilizador uma navegação horizontal, uma pesquisa intra ou inter-categorial. Estes relacionamentos são complementares aos hierárquicos, uma vez que o utilizador ao entrar numa categoria hierarquizada, pode realizar pesquisas genéricas e específicas. (Simões, 2008).

As relações associativas são sempre recíprocas e assimétricas (Simões, 2008), entretanto, seu estabelecimento é complexo, uma vez que, em uma área de conhecimento, todos os termos possuem um tipo de relação, mas é necessário verificar e estabelecer de forma consistente a relação associativa, formulando regras e critérios para o seu estabelecimento, bem como verificar sua pertinência, pois deve-se evitar ligações desnecessárias entre conceitos relacionados.

Mostafa e Nova Cruz (2012) ao revisar as relações associativas no escopo dos conceitos de David Hume, filósofo setecentista, destacam que estas relações não permitem uma relação de termos hierárquica ou absoluta e que as mesmas se fazem a partir da experiência que sempre se dá em um tempo e um determinado lugar. Outrossim, alicerçadas em Hume, enfatizam que as associações podem ser naturais quando se referem às relações inatas à natureza humana ou, filosóficas quando dependem da ideia para serem estabelecidas e, quando se trata da associação de ideia em campos específicos do conhecimento, as considerações hierárquicas não dão conta de resolver problemas impostos pelas experiências do vivido.

Continuamente, as relações semânticas podem ser observadas por vários prismas, incluindo nível textual ou nível lógico. Considerando a perspectiva da semântica lexical, baseado no linguista Murphy, Khoo e Nah (2006) identificam as propriedades gerais das relações semânticas:

- **Produtividade:** novas relações podem ser criadas facilmente;
- **Binaridade:** algumas relações, por exemplo a antonímia, são binárias no sentido de uma palavra pode ter apenas um antônimo verdadeiro, enquanto outras relações, como a sinonímia, podem relacionar-se a um conjunto de palavras;
- **Variabilidade:** as relações entre as palavras variam com o sentido conforme o contexto de uso;
- **Prototípica e canônica:** alguns pares de palavras são melhores exemplos de uma relação do que outros, e alguns pares de palavras tem status especial;
- **Semisemanticidade:** propriedades não emocionais, como categoria gramatical, coocorrência no texto, semelhança morfológica, que pode afetar a relação entre duas palavras;
- **Incontabilidade:** as relações semânticas são amplas e não podem ser contadas;

- **Previsibilidade:** as relações semânticas seguem certas regras e padrões gerais;
- **Universalidade:** os mesmos tipos de relações semânticas são usados em qualquer linguagem e em diferentes línguas.

O conhecimento acerca das propriedades das relações semânticas, auxiliam na sua compreensão. As tipologias e suas denominações podem variar conforme a abordagem, por exemplo em lógica ocorrem as denominações inclusão, posse, em semântica lexical utiliza-se meronímia, hiperonímia, antônimo, em Ciência da Informação, equivalência, hierárquica e associativa. Entretanto, não são conflitantes.

METODOLOGIA

Para consecução desta investigação de caráter exploratório, adotou-se como abordagem metodológica a análise de domínio e a análise conceitual. A análise de domínio por abarcar a delimitação e compreensão do conjunto de informações de um dado contexto, por meio do entendimento de padrões de comunicação e da relevância. E, a análise conceitual, oriunda da Teoria do Conceito, por envolver a identificação, determinação, estruturação e validação dos conceitos e relações semânticas apropriadas para a representação do conhecimento (Campos, 2009). Especificamente, na da Teoria do Conceito (Dahlberg, 1978), considera-se que o conceito é formado por três elementos: o referente (aquilo que se pretende conceituar), as características (soma dos enunciados verdadeiros sobre o referente) e a forma verbal – formando o triângulo do conceito.

Para auxiliar na compreensão dos conceitos e suas características, prioritariamente, utilizou-se como referência o vocabulário estruturado e trilingue DeCS (Descritores em Ciências da Saúde), criado pela Bireme (Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde), desenvolvido a partir do MeSH (*Medical Subject Headings da U. S. National Librar of Medicine (NLM)*). Secundariamente, quando surgiram dúvidas, foram realizadas pesquisas na PubMed.

Para coletar as tags, os sistemas baseados em folksonomia (SBF) utilizados foram: o CiteULike, que faz a indexação livre de trabalhos acadêmicos e o LibraryThing de livros. A revisão bibliográfica sobre as relações semânticas, sobretudo as associativas, foram utilizadas para auxiliar a identificação das relações semânticas.

RESULTADOS

Evidenciou-se como a variedade de relações semânticas se impõem como objetos centrais nas tarefas cognitivas, para compreensão de discursos, inferências e resolução de problemas. Foram obtidas para a análise 600 indexações livre, sendo que para este trabalho, foram selecionados 20 conceitos.

Considerando que a estrutura das relações são as mesmas, identificou-se que neste domínio, o padrão de relações semânticas, para além das relações hierárquicas, incidem nas relações da diabetes com outras doenças, o que revela a complexidade do domínio, uma vez que se vincula e apropria-se de uma diversidade de vocábulos. Verificou-se a predominância das relações associativas de causa e efeito, assim como procedimentos. Tocante as propriedades, observam-se a variabilidade das relações.

CONCLUSÕES

O tratamento e a recuperação da informação podem ser melhorados com a intensificação da exploração das informações semânticas intrínsecas aos recursos informacionais, de forma a expandir a compreensão e padrões de significado em textos, imagens e outras mídias. Adjacente a isso, é evidente que a aplicabilidade das folksonomias é bastante diversificada e vai além da verificação e validação do vocabulário do utilizador. Apesar das desvantagens que apresentam no que tange a riqueza vocabular, erros ortográficos e ambiguidade das tags, pesquisas com abordagem interdisciplinar, com destaque às áreas de Ciência da Informação, Computação, Linguística Computacional, comprovam o potencial da folksonomia para representação, recuperação e acesso a informação, assim como aprimoram a comunicação entre diferentes atores em um domínio de especialidade. Também, relativamente a área da saúde, verifica-se o seu potencial para a produção de instrumentos e plataformas digitais que visam melhorar a qualidade de vida de pacientes, seja para monitoramento, supervisão, sistemas de recomendação, decorrente do caráter colaborativo e do compartilhamento de informação entre muitos.

Relativamente às relações semânticas, apesar de sua variedade, identificar os padrões de ocorrência poderá, em trabalhos futuros, servir como subsídios para o desenvolvimento de pesquisas no âmbito do Processamento de Linguagem Natural. Ainda assim, neste contexto, a relevância de analisar as relações semânticas encontra-se sobretudo nas relações associativas, dada suas especificidades, critérios e subjetividade para o seu estabelecimento.

REFERÊNCIAS BIBLIOGRÁFICAS

- Bamidis, P. D., Kaldoudi, E., & Pattichis, C. (2009). From taxonomies to folksonomies: a roadmap from formal to informal modeling of medical concepts and objects. In *Information Technology and Applications in Biomedicine*, 2009. ITAB 2009. 9th International Conference on (pp. 1-4). IEEE.
- Chaffin, R., & Herrmann, D. J. (1988). Effects of relation similarity on part-whole decisions. *The Journal of general psychology*, 115(2), 131-139.
- Chaffin, R. (1992). The concept of a semantic relation. *Frames, Fields and Contrasts*, 253-288.
- Biolchini, J. C. D. A. (2001). Semântica e cognição em bases de conhecimento: do vocabulário controlado à ontologia. *Datagramazero*, 2 (3).
- Campos, J. A. G., & Souza, R. R. (2009). Análise conceitual sobre as relações semânticas em Ciência da Informação: contribuições para o desenvolvimento de ontologias. *Perspectivas em Ciência da Informação*, 14(3), 243-243.
- Cançado, M. (2013). Semântica Lexical: uma entrevista com Márcia Cançado.
- Dahlberg, I. Teoria do conceito. (1978). *Ciência da Informação*, Rio de Janeiro, v. 7, n. 2, p. 101- 107.
- Dixon, B. E., & McGowan, J. J. (2010). Enhancing a taxonomy for health information technology: an exploratory study of user input towards folksonomy.

- Frost, J. H., & Massagli, M. P. (2008). Social uses of personal health information within PatientsLikeMe, an online patient community: what can happen when patients have access to one another's data. *Journal of medical Internet research*, 10(3).
- Khoo, Christopher SG; NA, Jin-Cheon. (2006). *Semantic relations in information science*.
- Kristine Ådland, M., & Lykke, M. (2012). Chapter 5 Social Tagging in Support of Cancer Patients' Information Interaction. In *Social information research* (pp. 101-128). Emerald Group Publishing Limited.
- Lee, K., Kim, H., Shin, H., & Kim, H. J. (2009, May). Folksoviz: A semantic relation-based folksonomy visualization using the wikipedia corpus. In *Software Engineering, Artificial Intelligences, Networking and Parallel/Distributed Computing, 2009. SNPDP'09. 10th ACIS International Conference on* (pp. 24-29). IEEE.
- Linehan, C., Doughty, M., Lawson, S., Kirman, B., Olivier, P., & Moynihan, P. (2010, April). Tagliatelle: social tagging to encourage healthier eating. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems* (pp. 3331-3336). ACM.
- Maldonado-Martínez, Á. (2010). Los tesauros como herramienta fundamental en la descripción y recuperación de documentos electrónicos
- Mantelo Cecilio, H. P., Oliveira de Arruda, G., Ferraz Teston, E., Lima Santos, A., & Silva Marcon, S. (2015). Comportamentos e comorbidades associados às complicações microvasculares do diabetes. *Acta Paulista de Enfermagem*, 28(2).
- Medeiros, M. B. B., & Café, L. M. A. (2013). Organização da informação ou organização do conhecimento?. Disponível em: <http://repositorios.questoesemrede.uff.br/repositorios/bitstream/handle/123456789/809/17.pdf?sequencia=1>
- Moura, M. A. (2009). Folksonomias, redes sociais e a formação para o tagging literacy: desafios para a organização da informação em ambientes colaborativos virtuais. *Informação & Informação*, 14 (1), p.25-45.
- Mostafa, S. P., & NOVA CRUZ, D. V. D. (2012). David Hume e as relações associativas: da teoria ao tesouro.
- Peters, I. (2009). *Folksonomies. Indexing and retrieval in Web 2.0*. Walter de Gruyter.
- Stock, W. G. (2010). Concepts and semantic relations in information science. *Journal of the Association for Information Science and Technology*, 61(10), 1951-1969.
- Storey, V. C. (1993). Understanding semantic relationships. *The VLDB Journal—The International Journal on Very Large Data Bases*, 2(4), 455-488.



ANÁLISE BIBLIOMÉTRICA DA PRODUÇÃO CIENTÍFICA DO GRUPO DE PESQUISA MHTX

Gercina Ângela de Lima¹, Benildes Coura Moreira dos Santos Maculan², Filipi Miranda Soares³

¹Professora da Escola de Ciência da Informação (ECI), Universidade Federal de Minas Gerais (UFMG), Brasil, 0000-0003-0735-3856, glima@ufmg.br

²Professora da Escola de Ciência da Informação (ECI), Universidade Federal de Minas Gerais (UFMG), Brasil, 0000-0003-4303-9071, benildes@gmail.com

³Graduando em Biblioteconomia pela ECI, UFMG, Brasil, bolsista de Iniciação Científica da FAPEMIG, 0000-0002-0674-7960, filipivgp2011@gmail.com

RESUMO Apresenta o mapeamento da produção de teses e dissertações do Grupo de Pesquisa Modelagem Conceitual para Organização Hipertextual de Documentos - MHTX, realizando, a partir de uma abordagem quantitativa, com a coleta de dados, uma análise qualitativa dos dados levantados. Teve como objetivos identificar os autores mais citados na produção acadêmica do Grupo e identificar a base teórico-metodológica que vem sendo utilizada nas pesquisas de seus membros. Buscou-se saber se houve ampliação dos conteúdos para novos campos do conhecimento desenhando assim, a interdisciplinaridade, uma vez que se pressupõe ser característica do Grupo fazer interface com as áreas da Ciência da Informação, Biblioteconomia, Ciência da Computação, Ciências Cognitivas e Linguística. A pesquisa é caracterizada como documental descritiva e de abordagem quantitativa e qualitativa, na qual foram coletados os dados das produções científicas dos integrantes do Grupo de Pesquisa MHTX, referentes às teses e dissertações defendidas no período entre junho de 2004 (origem do Grupo) a junho de 2015, tendo como amostra cinco teses e onze dissertações. Os dados foram coletados a partir das referências bibliográficas, coletados utilizando planilhas do *Microsoft Office Excel* a partir dos seguintes parâmetros: autores; título; origem do texto; data; tipologia documental; editora ou periódico; instituições de pesquisas mais citadas; temporalidade; palavras chaves. Os resultados demonstraram que existe uma coerência teórica metodológica seguida nas pesquisas, e que tem sido ampliada conforme as pesquisas avançam fronteiras interdisciplinares.

PALAVRAS-CHAVE *Bibliometria, Grupo de Pesquisa MHTX, Produção científica, Teses e dissertações.*

ABSTRACT This paper presents the mapping of the theses and dissertations production from the Conceptual Modeling Research Group of the Hypertextual Documents Organization - MHTX. It was possible to do a qualitative analysis of the data collected through a quantitative approach with the data collection. It aimed to identify the most cited authors in the academic production of the Group and to identify the theoretical-methodological basis that has been used in the research of its members. It was intended to know if there was an expansion of the contents for new fields of knowledge using the interdisciplinarity, since it is assumed that it is a characteristic of the Group to do an interface with the areas of Information Science, Library Science, Computer Science, Cognitive Sciences, and Linguistics. The research is characterized as a descriptive documentary with a quantitative and a qualitative approach, in which the data of the scientific productions of the members of the MHTX Research Group were collected, referring to theses and dissertations defended in the period from June 2004 (origin of the Group) to June of 2015, taking as a sample five theses and eleven dissertations. The data were collected from the bibliographic references, using Microsoft Office Excel spreadsheets from the following parameters: authors; title; origin of the text; date; documentary typology; publisher or journal; the most cited research institutions; temporality; keywords.

The results showed that there is a theoretical methodological coherence in a sequence in the researches, and that it has been expanded, according to the research advances interdisciplinary frontiers.

KEYWORDS *Bibliometry, MHTX Research Group, Scientific production, Theses and Dissertations.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

1 INTRODUÇÃO

Os grupos de pesquisas são compostos por pesquisadores que se organizam em torno de uma linha de pesquisa, de uma área específica do conhecimento, buscando desenvolver investigações científicas para a produção e publicação de conhecimento. Vickery (1972, p. 33) define a investigação científica como toda e “qualquer investigação metódica, desenvolvida para fornecer informações e conhecimentos que foram utilizados para solucionar um problema”. Portanto, as investigações dos grupos de pesquisa são responsáveis pela validação de teorias e metodologias de pesquisa em uma área do saber. Nesse sentido, Alves (1987, p. 149) afirma que a publicação científica é suporte para o “processo de comunicação da produção científica e cultural, transforma-se em força motriz, na medida em que é recuperada e divulgada, impulsionando o desenvolvimento intelectual e realimentando o ciclo de geração de conhecimento”. Por isso, conhecer o “estado da arte” da produção científica dos grupos de pesquisa se tornou importante e relevante para a ciência. A análise dessa produção permite determinar as principais tendências desses estudos e delinear estratégias que possam fortalecer ainda mais os grupos de pesquisas e as áreas de conhecimento às quais eles estão vinculados.

Portanto, o objetivo principal deste estudo foi realizar uma coleta de dados para apoiar uma análise qualitativa da produção científica dos integrantes do Grupo de Pesquisa MHTX. Sendo assim, este estudo é parte de uma pesquisa mais ampla, composta por duas fases: (1) a primeira fase, que se trata da coleta de dados, com a aplicação de uma abordagem quantitativa, e, a (2) segunda fase, foi realizada a análise qualitativa dos dados levantados que visa a identificar as bases teórico-metodológicas que vêm sendo adotadas nas pesquisas que foram desmembradas a partir dos estudos apresentados na primeira tese defendida que deu origem ao Grupo de Pesquisa MHTX.

O ponto de partida teórico da pesquisa dessa primeira tese se fundamentou na Teoria da Análise Facetada, desenvolvida por Ranganathan (1967), que orientou a modelagem da estrutura do documento acadêmico. A aplicação dessa teoria “satisfaz a não linearidade do hipertexto, facilitando a criação de estrutura fluída e interativa como a do Mapa Hipertextual” (Lima, 2007, p. 39), pois permitiu uma análise na área de assunto do documento, com a identificação de conceitos que puderam ser agrupados de maneira analítica, permitindo ao usuário ver o assunto sob diferentes pontos de vista. Desde que foi implantado, em 2004, o protótipo MHTX serve como base para pesquisas de alunos da Pós-Graduação da ECI-UFMG, sob a orientação de sua coordenadora. Os resultados dessas pesquisas visam à simplificação dos processos de organização de informações, que são processos comumente complexos, visando a facilitar o acesso e a recuperação da informação contida nas versões impressas de teses e dissertações (Lima, 2007), agregando conhecimentos oriundos de outras áreas.

Nesse sentido, este estudo apresenta os resultados da coleta de dados e a análise qualitativa do levantamento bibliométrico que foi realizado para mapear a produção científica dos integrantes do Grupo de Pesquisa Modelagem Conceitual para Organização Hipertextual de Documentos - MHTX

(Grupo de Pesquisa MHTX), que é registrado no Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), desde 2004. A produção científica analisada é composta pelas teses e dissertações defendidas desde 2004, todas orientadas pela coordenadora do Grupo de Pesquisa MHTX. As teses e dissertações estão vinculadas ao Programa de Pós-Graduação em Ciência da Informação (PPGCI), da Escola de Ciência da Informação (ECI) da Universidade Federal de Minas Gerais (UFMG), dentro da linha de pesquisa Organização e Uso da Informação. Esses resultados têm como objetivo identificar as bases teórico-metodológicas que vêm sendo empregadas nos estudos dos pesquisadores do grupo, a partir da determinação dos autores mais citados nessa produção científica. Dessa forma, o estudo não pretendeu verificar o fator de impacto da publicação dos integrantes do grupo, mas averiguar quais autores e conteúdos esses pesquisadores vêm utilizando em suas pesquisas.

A técnica utilizada para analisar a produção científica dos integrantes do grupo de pesquisa foi a Bibliometria, que é uma “técnica quantitativa e estatística de medição dos índices de produção e disseminação do conhecimento científico” (Araújo, 2006, p. 12). Os estudos bibliométricos tiveram início como uma abordagem quantitativa, mas, aos poucos, a partir dos dados estatísticos, foi possível desenvolver análises mais complexas e qualitativas, ampliando o seu uso no meio acadêmico. Na literatura sobre bibliometria encontram-se “leis” que norteiam a análise da produção científica em diferentes áreas do conhecimento que, segundo Chen, Chong e Tong (1994), as principais são: *Bradford* (produtividade de periódicos), *Lotka* (produtividades de autores) e *Zipf* (frequência de ocorrência de palavras).

Os resultados demonstraram que existe uma coerência teórico-metodológico nas pesquisas do grupo, e que tem sido ampliada conforme as pesquisas avançam fronteiras interdisciplinares.

2 METODOLOGIA

Um pesquisador é formado a partir de sua inserção em contextos científicos, ao ter contato com lacunas e questões não elucidadas sobre problemas em uma área do conhecimento. Em sua busca por soluções, esse pesquisador se depara com grande diversidade de teorias e metodologias que poderão respaldar o caminho que será adotado em seu estudo. É uma característica importante e requerida no percurso de um pesquisador, que suas pesquisas reflitam uma linha teórica metodológica coesa, possibilitando analisar e evidenciar seu perfil acadêmico, e sua real contribuição para uma determinada área do conhecimento. Assim, neste trabalho, as questões que se pretendeu investigar foram: Os autores utilizados como base referencial na tese, que deu origem a este Grupo, continuaram sendo importante para as pesquisas subsequentes realizadas? Existe uma coerência teórica metodológica utilizada nas pesquisas realizadas pelos pesquisadores deste Grupo? A bibliografia analisada possibilitou verificar quais outros campos de conhecimentos foram sendo agregados nas pesquisas sequenciais realizadas neste período?

A pesquisa é caracterizada como documental, descritiva, com a abordagem quantitativa e qualitativa. Como amostra foram coletados os dados das produções científicas dos integrantes do Grupo de Pesquisa MHTX, referentes às teses e dissertações defendidas no período entre junho de 2004 (origem do Grupo) a junho de 2015, sendo cinco teses e onze dissertações.

Na coleta de dados foram compiladas as referências bibliográficas destes dezesseis documentos, totalizando-se 1551 referências. Desse total, foram descartadas 99 referências, após aplicar os seguintes critérios de exclusão:

- a) descartar as referências dos dicionários “não especializados”, como Aurélio, Houaiss e outros dicionários de língua e gramática;
- b) descartar referências de tesouros, glossários e enciclopédias não especializados;
- c) descartar referências de portais, bibliotecas digitais e outros sites sem cunho científicos;
- d) descartar referências de normas técnicas e verbetes, referentes a publicações tais como dicionários de línguas (ex.: português-inglês, de sinônimos), tesouros, portais online, e de bibliotecas digitais.

Os dados das referências bibliográficas foram analisados e coletados de forma manual (intelectual), utilizando-se o programa editor de planilhas *Microsoft Office Excel*, de acordo com os seguintes campos: nome de autores; título da publicação; indicação se o texto é original, traduzido, ou se é uma adaptação; no de publicação; tipologia documental; editora ou periódico, que publicou o texto e instituições de ensino e pesquisa mais citadas.

Além, desses campos mencionados anteriormente, foram adicionadas cinco colunas para possibilitar a inserção da atribuição de diferentes pesos para a citação recebida por um autor. Conforme Maltrás-Barba (2003), a análise de coautoria, para estabelecer redes de colaboração, pode ser realizada pelos métodos completo ou fracionado. Por não ser objetivo identificar as ligações relacionais entre os diferentes autores, mas conhecer aqueles autores que são mais citados pelos integrantes do grupo, decidiu-se por aplicar uma metodologia híbrida, com a qual se valorizam os trabalhos em colaboração, ao dar o peso de uma frequência para o primeiro autor, sem tirar deste a importância da responsabilidade pela publicação, fracionando pela metade o peso da frequência para os demais coautores. Dessa forma, criou-se parâmetros para avaliação das possíveis variáveis que fossem surgindo na análise dos dados, tanto em relação à posição ocupada na citação pelos autores, quanto nas variações dos nomes dos autores, para tentar solucionar problemas de inconsistência dos dados.

A análise qualitativa partiu dos dados quantitativos, tendo sido criados parâmetros para coletar dados sobre os autores mais citados. Teve como ponto de corte da lista, a citação do mesmo autor com até cinco citações no todo, independente de sua posição na ordem de autoria e coautoria, sendo citados em mais de duas citações.

Posteriormente, através das análises das referências citadas, foi possível analisar se foram mantidos somente os autores mais referenciados da lista de autores mais citados, independentes da posição da citação de autoria na referência bibliográfica.

O Gráfico 1, reporta os dados sobre os autores mais citados por atribuição de peso, como resultado da análise das autorias das 1452 referências, em relação ao valor absoluto de ocorrências, demonstrados na forma de gráfico, os principais autores utilizados como base referencial na teses e dissertações do Grupo.

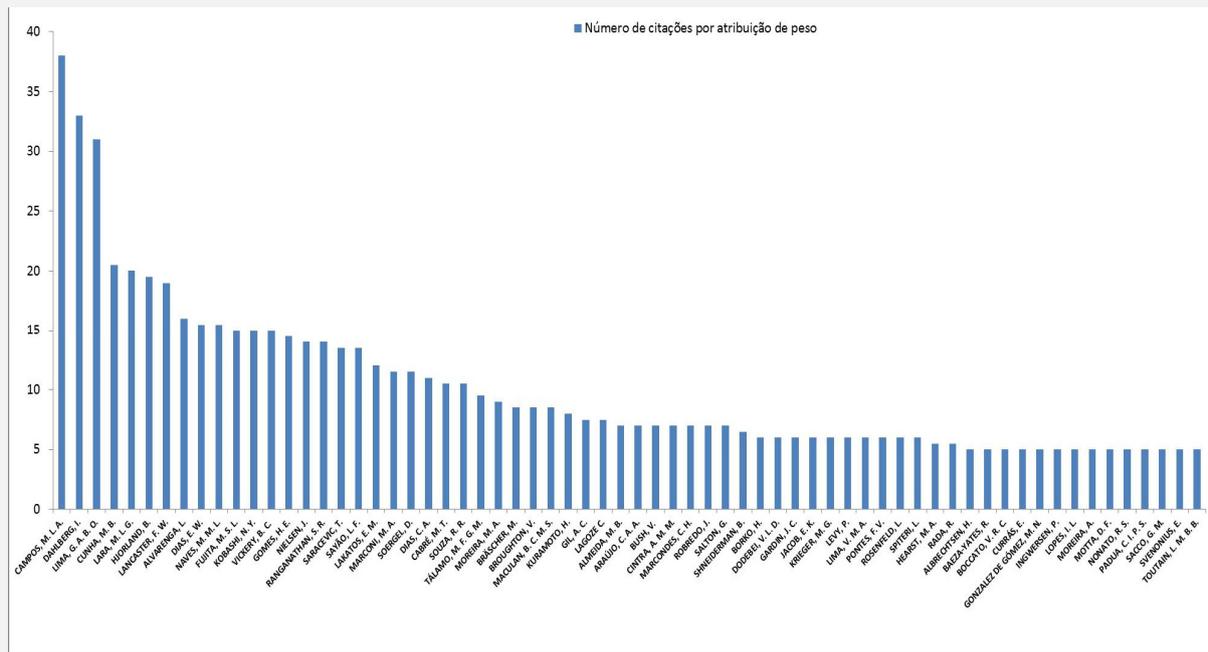


Gráfico 1. Autores mais citados por atribuição de peso

Referência: Resultado da pesquisa, (2015).

Analisando o Gráfico 1, pela frequência absoluta, é possível apreender que entre os três autores mais citados, está a autora da primeira tese defendida que deu origem ao Grupo de Pesquisa MHTX, Lima, G. A.B.O., confirmado, assim, a premissa que todos os outros trabalhos acadêmicos defendidos partiram, inicialmente, das temáticas pesquisadas nesta tese. Nota-se, portanto, que a primeira autoria posicionada no gráfico é de Campos, M. L. A. que se justifica por ter sido uma das referências metodológicas da pesquisa inicial, e também, por ter suas publicações com temáticas relacionadas à área de Organização da Informação e do Conhecimento, mais especificamente relacionada à Teoria da Análise Facetada, à Teoria do Conceito, de Dahlberg, e à Teoria da Terminologia de Wüster, teorias essas muito utilizadas nas pesquisas realizadas pelos orientandos do Grupo. Em segundo lugar aparece a Dahlberg, I. estudiosa da Teoria do Conceito, e da área da Organização do Conhecimento, temáticas relevantes em quase todos os trabalhos acadêmicos do Grupo MHTX. Em posições semelhantes aparecem respectivamente, Cunha, M. B. que trabalha com a temática sobre bibliotecas digitais, Lara, M. L. G., que estuda os aspectos linguísticos das linguagens e Terminologia, Hjørland, B. que, também, publica estudos sobre a área da Organização do Conhecimento, e Lancaster F.W., estudioso sobre indexação, principalmente sobre análise de assunto. Os 12 dos outros autores aparecem em posição semelhantes, indo para um decrescendo destacando que 50% dos autores citados aparecem com maior relevância, e os outros 50% dos autores em destaque menor, caracterizando a utilizando das referências em trabalhos mais pontuais. Salienta-se, portanto, que os autores Maculan, B.C.M.S e Pontes, F.V. tiveram a dissertação e tese defendida, respectivamente, sob a orientação da coordenadora do Grupo MHTX, e são citados como referências, demonstrando uma característica do Grupo que é a de uma pesquisa dar continuidade nas outras pesquisas subsequentes, visto que nem sempre se esgota um tema em uma única pesquisa.

O Gráfico 2 apresenta o índice de citação de cada autor por ano, pela frequência absoluta, a partir de 2004, data que foi defendida a primeira tese, a 2015, quando fez o corte temporal para esta análise. Porém, ressalta que existe um intervalo entre os anos 2005 e 2006, por não possuir nenhum documento acadêmico defendido, neste período.

Verifica-se, que os autores mais citados por ano, coincide com os dados do Gráfico 1 dos autores mais citados por peso, anteriormente descrito. A autora Campos, M.L.A. teve 42 citações de seus trabalhos, destas, o índice de maior ocorrência foi em 2015, com nove citações, seguido pelos anos 2008 e 2011, com seis citações, respectivamente, e o ano de 2009, com cinco citações. O ano de 2010 aparece com três citações, e, o ano de 2007, com duas ocorrências. Em seguida, aparece Dahlberg, I. com 29 citações, com maior ocorrência no ano de 2009, com nove citações. A autora volta a ser muito citada, novamente, nos anos de 2011 e 2015, com sete e seis citações, respectivamente. Nos anos de 2010 e 2008 a autora foi citada com bem menos frequência, ocorrendo quatro citações no ano de 2010, e em 2008, três citações. Lima, G.A.B.O foi citada por 21 vezes, sendo que o ano de 2009 foi o ano que ocorreu um maior número índice de citações, com sete ao todo.

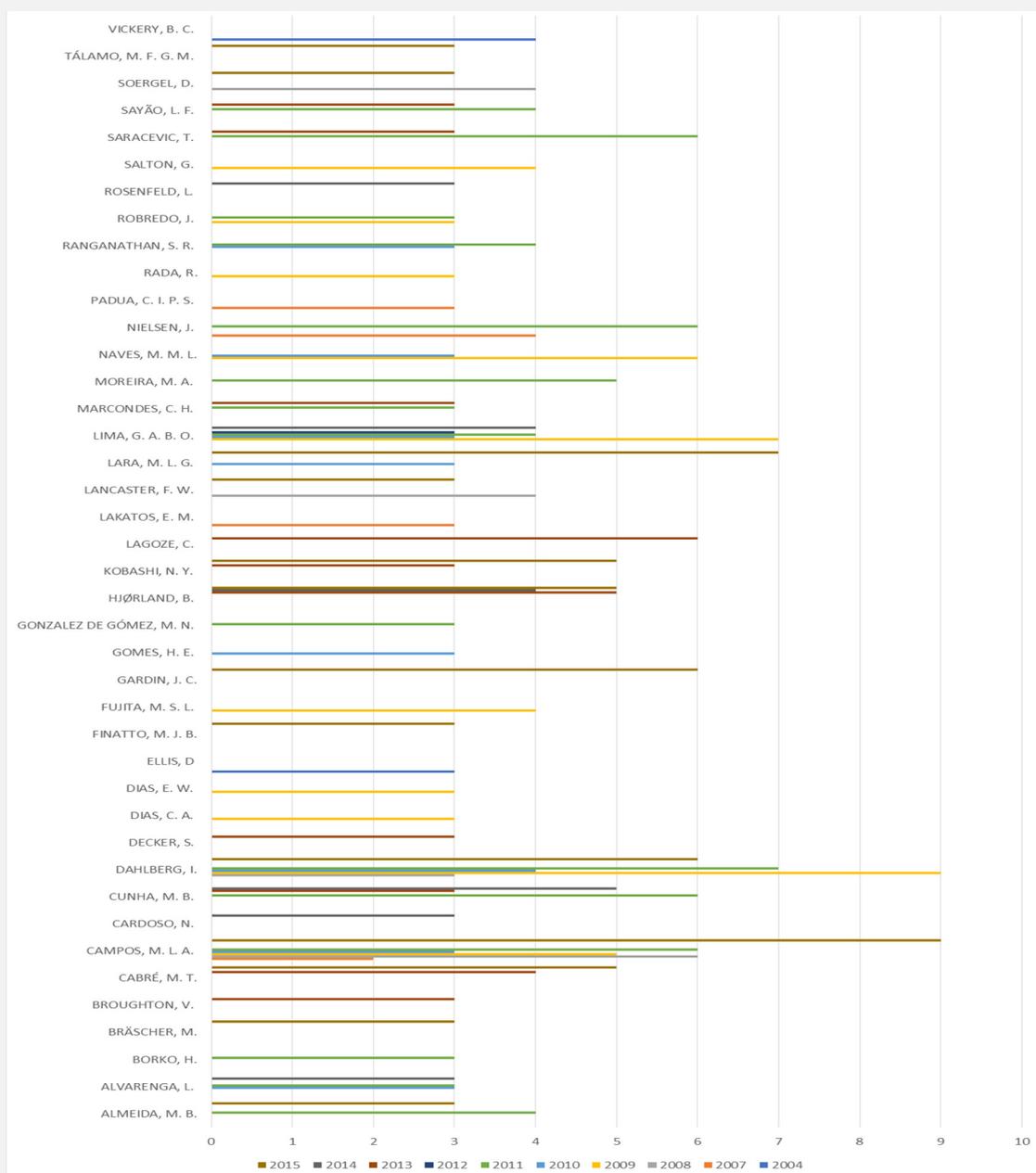


Gráfico 2. Índice de citação de autores por ano

Referência: Resultado da pesquisa, (2017).

Ademais, verifica-se que os autores que alcançaram acima de dez citações foram Hjørland, B. com 14 citações, sendo cinco em 2013, quatro em 2014 e cinco em 2015; e Cunha, M. B. com 14 citações, também, com seis em 2011, três em 2013, e cinco em 2014; todos os outros ficaram abaixo de dez citações, neste período de tempo.

Observa-se que no Gráfico 3, que entre os trabalhos mais citados encontra-se a tese “Mapa hipertextual (MHTX): um modelo para organização hipertextual de documentos”, de autoria de Lima, G.A.B.O com ocorrências 12 citações. Este fato pode ser justificado porque foi este o documento que deu origem a todas as outras pesquisas subsequentes. Logo após, vem o trabalho “A organização de unidades do conhecimento em hiperdocumentos: o modelo conceitual como um espaço comunicacional para a realização de autoria”, de autoria de Campos, M.L.A., o artigo de “As we may think” de Vannevar Bush, que ainda é um trabalho seminal para os estudos sobre hipertexto, e o artigo de Dias, E.W “Contexto digital e tratamento da informação”. A ocorrência destes quatro trabalhos caracteriza muito bem a interseção das duas áreas que subsidiaram os estudos iniciais do Grupo MHTX: Biblioteconomia e Ciência da Informação (BCI) e a Ciência da Computação.

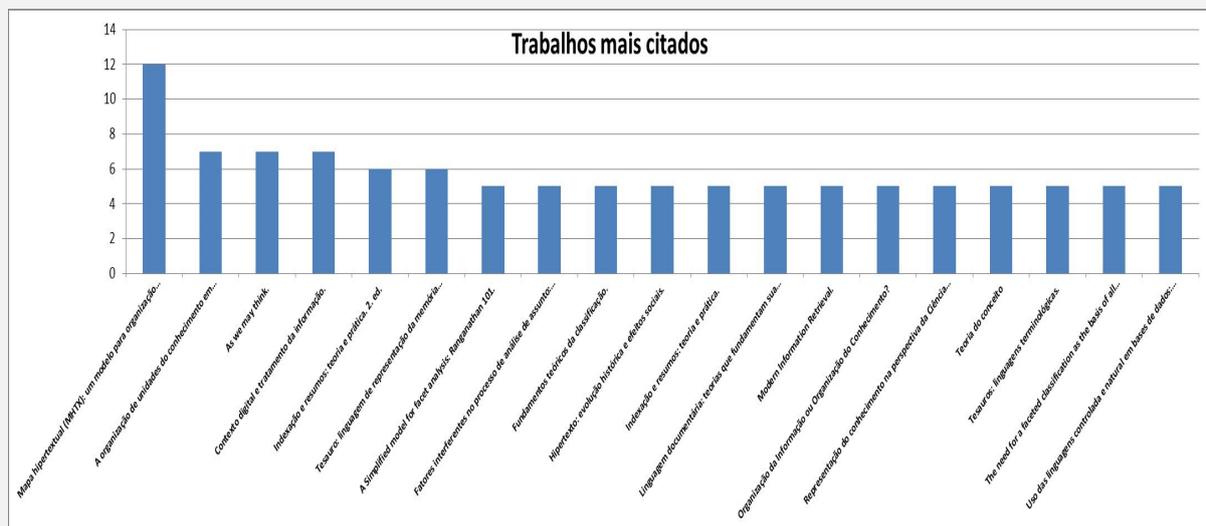


Gráfico 3. Os trabalhos mais citados

Fonte: Resultado da pesquisa, (2015)

Ao analisar as variáveis dos Gráficos 4, 5, 6, respectivamente, sobre os periódicos científicos mais citados, dos eventos científicos mais citados e, das editoras mais citadas, pode-se perceber que, as referências bibliográficas dos 16 trabalhos do Grupo de Pesquisa MHTX foram publicados em 17 periódicos, dez eventos científicos e 24 editoras.

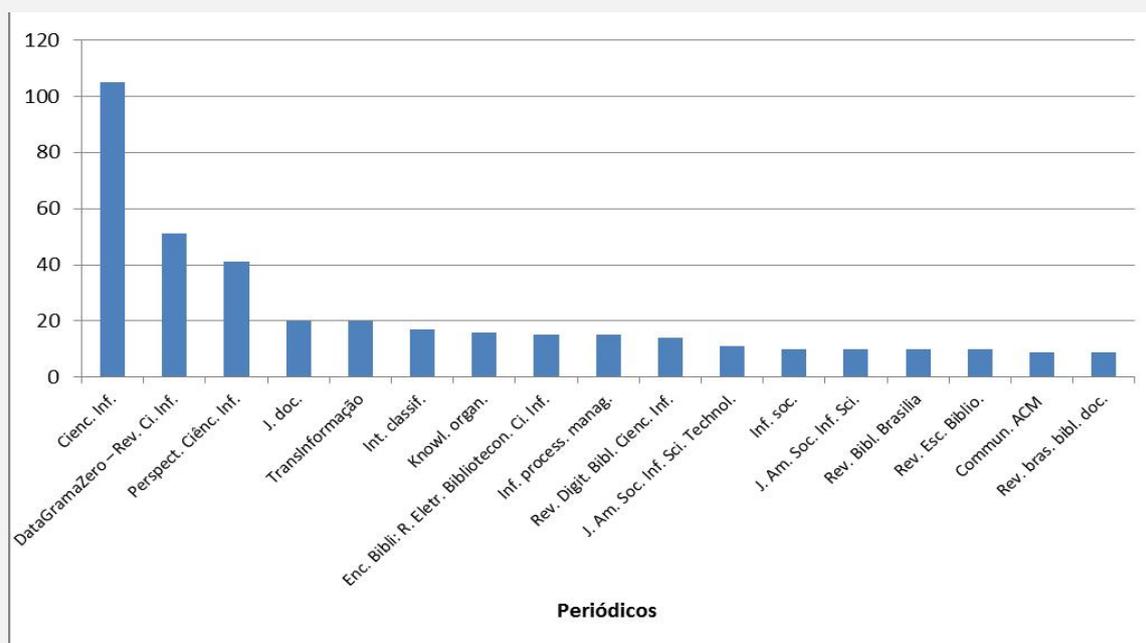


Gráfico 4. Os periódicos científicos mais citados

Fonte: Resultado da pesquisa, (2015).

Em relação aos periódicos mais citados destaca-se a Revista Ciência da Informação, publicada pelo IBICT/Rio de Janeiro/Brasil, com 105 citações, que publica artigos da área Organização da Informação e do Conhecimento e assuntos relacionados às novas tecnologias, visto que está sob a responsabilidade do Instituto Brasileiro de Informação, Ciência e Tecnologia.

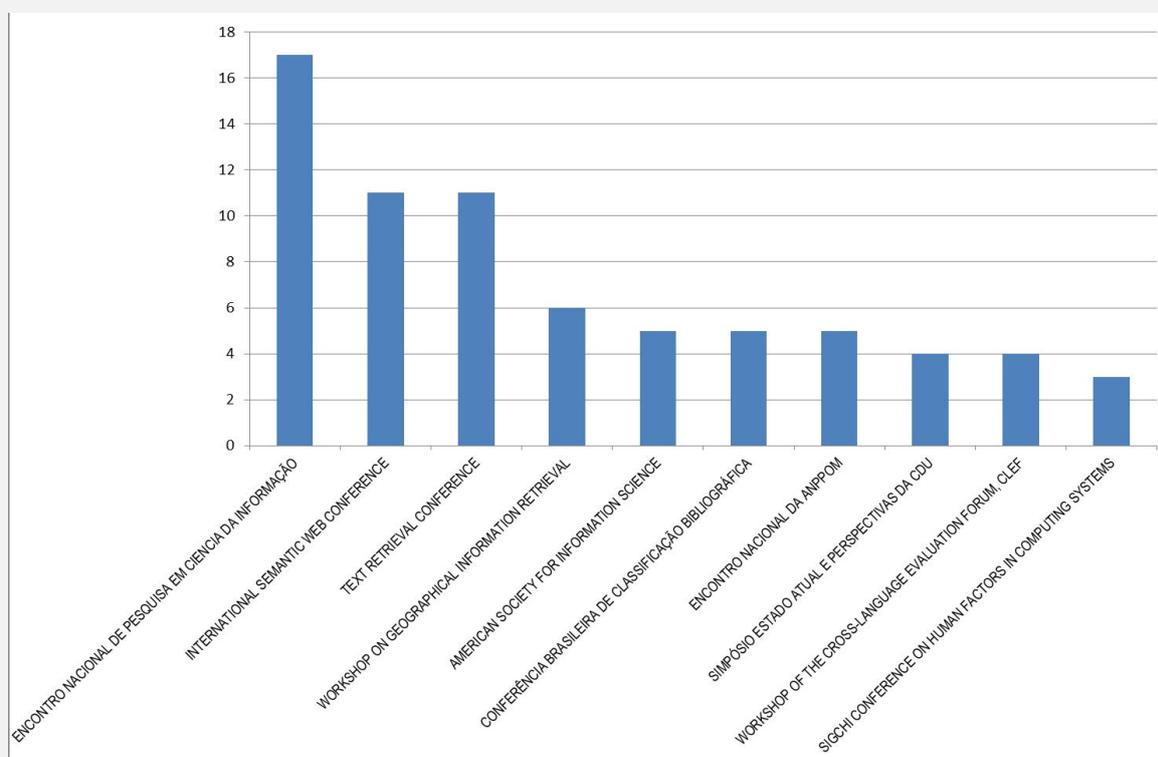


Gráfico 5. Eventos científicos mais citados

Fonte: Resultado da pesquisa, (2015).

Além dos periódicos científicos, houve uma grande representatividade de publicações em evento científico, como por exemplo, o Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), teve destaque com 17 citações foram publicadas, justificada por ser este o maior evento de pesquisa na área no Brasil, que ocorre anualmente.

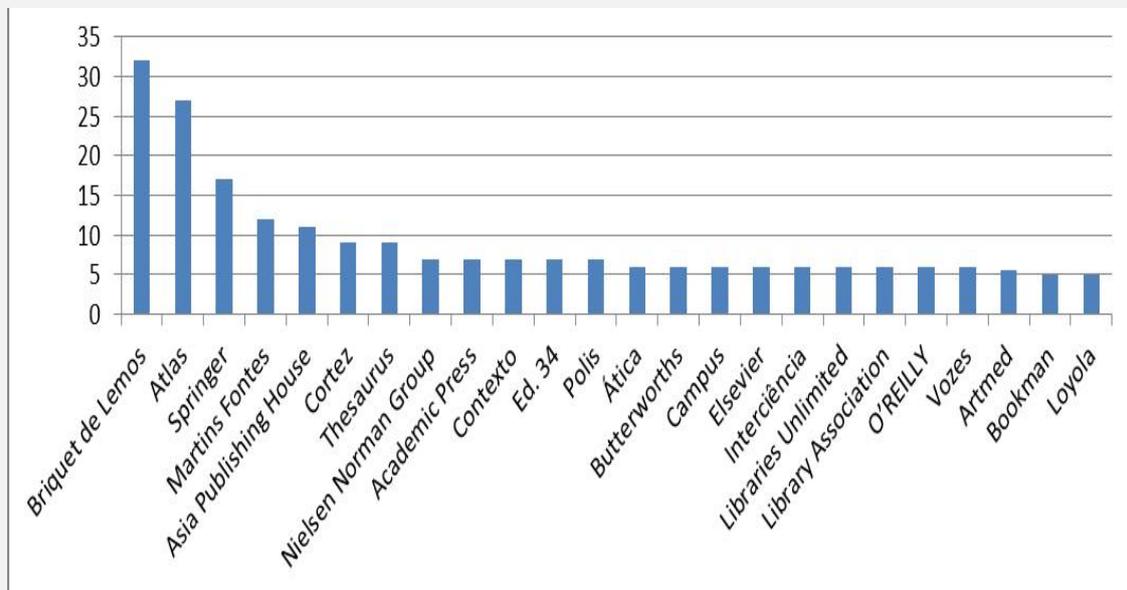


Gráfico 6. As editoras mais citadas

Fonte: Resultado da pesquisa, (2015).

Em relação a Editora, sobressai a Editora Briquet de Lemos, com 32 citações, que publica exclusivamente livros da área da BCI, fato que justifica ser ela a mais citada.

No Gráfico 7 apresenta as principais palavras-chave abordadas pelos trabalhos, com o intuito de mapear se as mesmas temáticas se mantiveram nos trabalhos subsequentes neste período, e quais outras surgiram pela necessidade de complementação de conhecimento advinda de outras áreas correlatas do conhecimento. Os trabalhos analisados utilizaram em 71 palavras, que posteriormente foram agrupadas semanticamente equivalentes (em diferentes graus), resultando em 36 palavras-chave, para serem representadas graficamente. As principais foram: Organização e Representação da Informação e Biblioteca Digital, utilizadas em seis trabalhos, seguidas de Recuperação da Informação, e Interface utilizadas em cinco, e Teoria da Classificação Facetada e Hipertexto com quatro ocorrências cada. Conforme pode ser observado no Gráfico 8, outras palavras-chave foram utilizadas pelos autores, porém com menos ocorrências, tais como: Teoria da Análise de Domínio, Taxonomias facetadas e dinâmicas, modelagem conceitual, indexação manual, Teoria do Conceito, Protótipo/Protótipo MHTX, Mapa Conceitual e Arquitetura da Informação. As outras palavras chave não citadas aqui foram utilizadas somente uma vez nos documentos.

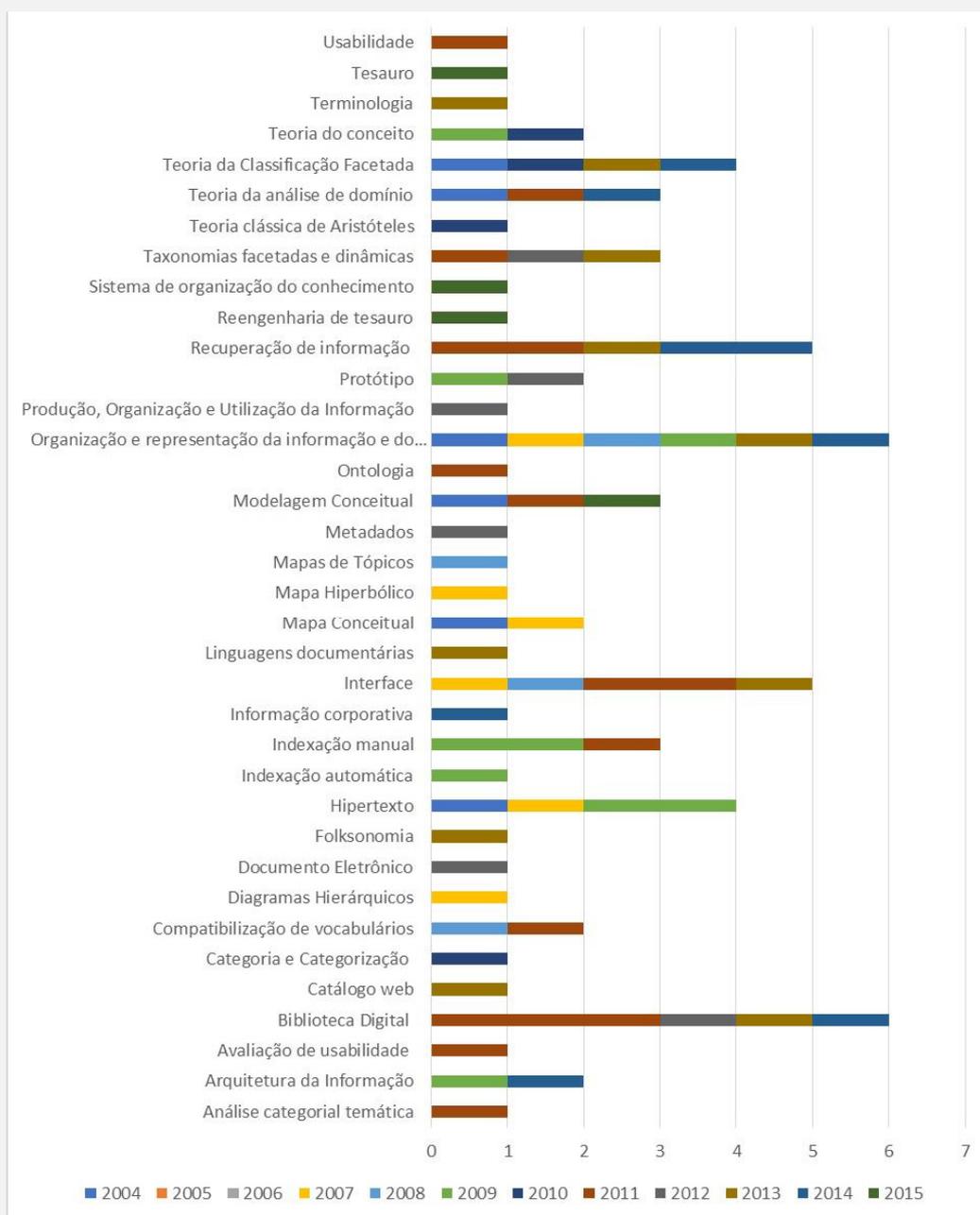


Gráfico 7. Índice de citação de palavras-chave por ano

Fonte: Resultado da pesquisa, (2017).

Considera-se que a metodologia utilizada foi adequada porque possibilitou chegar aos resultados previstos na investigação, conforme a questão a ser investigada e aos objetivos propostos. Este estudo trouxe elementos que permitiram a análise, tanto quantitativa quanto qualitativa, que retrata a coerência nos fundamentos teóricos metodológicos adotados pelos trabalhos realizados pelo Grupo de Pesquisa MHTX.

4 RESULTADOS

Os resultados alcançados demonstram:

- a) Na frequência absoluta o autor mais citado foi Lima (2004), isto se justifica porque todas as pesquisas realizadas subsequentes tiveram como problema de pesquisa questões que levavam a solucionar as lacunas não resolvidas por esta pesquisa inicial;
- b) O trabalho mais citado, por atribuição de peso, foi a tese de Campos (2001), que também foi utilizada como referencial teórico metodológico no trabalho de Lima (2004) que deu origem ao Grupo de Pesquisa MHTX;
- c) Permitiu fazer um mapeamento de quais os periódicos e editoras foram mais citados, bem como quais as instituições de ensino e pesquisa foram mais referenciadas e a data de publicação referenciadas nas teses e dissertações defendidas, utilizando o campo “data de publicação” da referencia bibliográfica, e analisada por meio de uma tabela de temporalidade por década;
- d) Pode-se mapear os assuntos mais recorrentes em cada trabalho do Grupo de Pesquisa, a partir das palavras chaves, relacioná-los com a referência bibliográfica utilizada em cada tese e dissertação defendida pelos membros do grupo de Pesquisa MHTX, e as novas temáticas que foram surgindo à medida que as pesquisas avançavam e buscavam aportes teóricos em outras áreas do conhecimento.
- e) Possibilitou evidenciar quais autores continuaram sendo utilizados como referência em todas as teses e dissertações defendidas, ou se outros autores foram utilizados para embasar a mesma temática, ou mesmo se novos autores foram sendo referenciados a partir da ampliação do escopo de pesquisa do Grupo MHTX, com as áreas interdisciplinares tais como a Ciência da Computação, Linguística e Ciências Cognitivas.

5 CONCLUSÕES

Os métodos estatísticos da bibliometria, cujas principais Leis são as Lei de *Lotka*, em 1926, Lei de *Bradford*, em 1934, e Lei de *Zipf*, em 1936, são importantes para a coleta de dados quantitativos para análises qualitativas da produção científica de um grupo de pesquisadores. Nesse contexto, as análises de citação podem dar subsídios para investigar a realidade dos estudos de um grupo de pesquisadores, dando subsídios para verificar possíveis lacunas para novos estudos. Neste trabalho, adotamos uma abordagem híbrida, uma vez que não havia a intenção de verificar o fator de impacto de um autor e nem mesmo o grau de colaboração de coautorias, e, sim, os autores mais citados e suas contribuições.

Os resultados dos dados quantitativos respaldaram as análises qualitativas, que teve com objetivo identificar as bases teórico-metodológicas que estão sendo empregadas pelos integrantes do Grupo de Pesquisa MHTX. Na análise aponta indícios de que, prioritariamente, o grupo utiliza os aportes de Ranganathan (Teoria da Análise Facetada) e de Dahlberg (Teoria do Conceito e a área da Organização do Conhecimento). Verificou-se, assim, que a base teórica inicial do grupo, que foi com a Teoria da Análise Facetada, continua como ponto norteador da maior parte das pesquisas do grupo, mas foram incorporadas, também, tendo por base o estudo do índice de citação dos autores, as teorias advindas da Terminologia.

Porém, novas temáticas, oriundas das áreas da Ciência da Computação, da Linguística e das Ciências Cognitivas, têm surgido, à medida que as pesquisas avançam em busca de solucionar a proposta inicial do Grupo MHTX.

AGRADECIMENTOS

Agradecimentos ao apoio das Agências de fomento: FAPEMIG, CAPES e CNPq/Brasil.

REFERÊNCIAS BIBLIOGRÁFICAS

Alves, M. A. M. (1987). A Biblioteca Nacional, banco de dados da produção científica e cultural brasileira. In *Anais do 5 Seminário Nacional de Bibliotecas Universitárias* (pp. 149-166). Porto Alegre.

Araújo, C. A. A. (2006). Bibliometria: evolução histórica e questões atuais. *Em Questão*, Porto Alegre, 12(1), 11-32. Recuperado de: <http://www.redalyc.org/html/4656/465645954002/>. Acessado em: 09/set./2017.

Campos, M. L. A. (2001). *A organização de unidades do conhecimento em hiperdocumentos: o modelo conceitual como um espaço comunicacional para realização da autoria*. (Tese de Doutorado, Instituto Brasileiro de Informação em Ciência e Tecnologia, Rio de Janeiro). Recuperado de: <http://www.conexaorio.com/bit/tesemlcampos/TeseMLCampos.pdf>. Acessado em: 09/set./2017.

Chen, Y., Chong, P. P. & Tong, M. Y. (1994). The simon-yule approach to bibliometric modeling. *Information Processing & Management*, 30(4), 535-56. Recuperado de doi: 10.1016/0306-4573(94)90038-8. Acessado em: 09/set./2017.

Lima, G. A. B. O. (2004). *Mapa hipertextual (MHTX): um modelo para a organização hipertextual de documentos* (Tese de Doutorado, Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte). Recuperado de: <http://www.bibliotecadigital.ufmg.br/dspace/handle/1843/LHLS-6BUPG9>. Acessado em: 09/set./2017.

Lima, G. A. B. M. O. (2007). Modelo hipertextual-MHTX: um modelo para organização hipertextual de documentos. *DataGramaZero*, 8(4). Recuperado de <http://www.brapci.ufpr.br/brapci/index.php/article/download/7550>. Acessado em: 09/set/2017.

Maltrás-Barba, B. (2003). *Los indicadores bibliométricos: fundamentos y aplicación al análisis de la ciencia*. Gijón: Trea.

Ranganathan, S. R. (1967). *Prolegomena to library classification*. London: Asia Publishing House.

Vickery, B. C. (1972). The administration of research in institution. *Anais do The Advisory Board on Research of the Library Association*, The Library Association, London, pp. 33-38.



CIÊNCIA DA INFORMAÇÃO E WEB SEMÂNTICA: UMA RELAÇÃO EFETIVA OU APÓCRIFA?

Luís Machado¹, Maria da Graça Simões², Renato Rocha Souza³, Maurício Almeida⁴

¹Universidade de Coimbra, 0000-0003-3403-5618, luismachado@fe.uc.pt

²Universidade de Coimbra, gsimoes@darq.uc.pt

³Fundação Getúlio Vargas, 0000-0002-1895-3905, renato.souza@fgv.br

⁴Universidade Federal de Minas Gerais, mba@eci.ufmg.br

RESUMO A Web é o local privilegiado para procurar informação. As potencialidades de uma *web semântica* para recuperar informação afiguram-se essenciais. Na Ciência da Informação (CI), esta ideia surge com Otlet ao desenvolver uma “web semântica analógica”, justificando a ligação entre: Web Semântica (WS) e CI. Procura-se verificar e refletir sobre a presença da WS no *output* dos cursos de mestrado/doutoramento em CI, Portugal e Brasil (2005-2015). Para tanto: identificam-se os cursos, dissertações de mestrado (DM) e teses de doutoramento (TD) em CI associados à WS; identificam-se pontos de confluência entre as duas áreas nestes trabalhos; caracterizam-se o perfil académico dos autores e orientadores dos mesmos. Partiu-se do paradigma qualitativo: revisão bibliográfica e estudo empírico, centrado na análise de conteúdo dos trabalhos recuperados (RCAAP e OASISbr). Verificou-se 4 registos únicos no RCAAP; 40 no OASISbr. Identificaram-se 4 cursos de mestrado e 3 de doutoramento (Portugal), e 16 de mestrado e 9 de doutoramento (Brasil). Recuperaram-se 0 DM/TD (Portugal) e 16 DM, 5 TD (Brasil). Identificaram-se 3 pontos de confluência: a) contribuição mútua (62%), b) WS enquanto tema interdisciplinar (29%) e c) complementaridade de conhecimentos (9%). Contabilizaram-se 11 trabalhos (52%) onde os investigadores responsáveis apresentam, no conjunto, formação nas áreas CI/Informática. O estudo aponta para a WS como um campo de investigação da CI, embora essencialmente teórico, sendo o trabalho prático interdisciplinar, CI/Informática, incipiente e pouco representativo.

PALAVRAS-CHAVE *Web Semântica; Ciência da Informação; cursos de Mestrado/Doutoramento; Portugal; Brasil.*

ABSTRACT The Web is the privileged place for seeking information. The potential of the so-called *semantic web* to enhance retrieval of relevant information seems essential. In Information Science (IS), this idea was developed from the Otlet's “analogic semantic web”, justifying the connection between the Semantic Web (SW) and the field of IS. We aimed to exam and analyze the presence of subjects related to SW in the output of IS master and PhD courses in Portugal and Brazil (2005-2015). Specifically, we aimed to: identify IS courses, master's dissertations (MD) and PhD thesis (PD) associated to SW; identify the points of confluence between the IS and the SW in these works; characterize the academic profile of these works' authors and advisors. Based on a qualitative approach, we performed a literature review and an empirical study focused on content analysis of the retrieved works (RCAAP and OASISbr). As results, we have: verified 4 unique records in RCAAP and 40 in OASISbr; identified 4 Master's and 3 PhD courses (Portugal) and 16 masters and 9 PhD courses (Brazil); retrieved no MD/PD in Portugal, 16 MD and 5 PD in Brazil; identified 3 confluence points: a) mutual contribution (62%), b) WS as an interdisciplinary theme (29%) and c) complementary knowledge (9%); identified 11 works (52%) where the respective researchers have training in IS and Computer Science (CS). The study points to the WS as a field of investigation in IS, although essentially theoretical, insofar as practical interdisciplinary works in CI/CS are incipient and not representative.

KEYWORDS *Semantic Web; Information Science; Master's/PhD courses; Portugal; Brazil.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Nas últimas décadas tem-se assistido a um aumento exponencial de informação produzida no espaço da *World Wide Web* (WWW ou simplesmente *Web*). Na transição da informação, do analógico para o digital, uma carência começou a evidenciar-se: a necessidade de uma recuperação mais eficaz da mesma, considerando-se a precisão e a revocação (Caro-castro, 2010; Patel-Schneider & Horrocks, 2007; Souza & Alvarenga, 2004). Neste contexto uma “nova” *Web*, denominada de *Web Semântica*, apresenta-se como solução para esta situação. Muitas esperanças foram colocadas nesta *Web Semântica* (WS), nomeadamente, como sendo capaz de criar um ambiente onde *agentes*, um tipo específico de *software* “inteligente”, processando e partilhando “conhecimento”, conseguem realizar tarefas complexas de acordo com as necessidades dos utilizadores (Berners-Lee, 2006; Berners-Lee & Fischetti, 2000; Berners-Lee, Hendler, & Lassila, 2001). A questão que se coloca neste cenário é a necessidade de tornar as máquinas capazes de raciocínios similares ao dos humanos decorrendo, daí, uma associação à chamada Inteligência Artificial, que os investigadores ligados à WS, logo nas publicações iniciais, procuraram afastar relacionando-a, antes, à modelação do conhecimento humano utilizando meios computacionais por forma a possibilitar inferências automáticas (Berners-Lee & Hendler, 2001; Shadbolt, Hall, & Berners-Lee, 2006) e, deste modo, garantir uma recuperação da informação eficaz.

A revisão da literatura relativa a este conceito (*Web Semântica*), revelou a necessidade de uma clarificação relativamente a outros conceitos associados, comumente utilizados como sinónimos, tais como: *Web de Dados* (*Web of Data*) ou *Web de Dados Ligados* (*Web of Linked Data* ou simplesmente *Linked Data*), conceitos esses que, apesar de intrinsecamente relacionados, são passíveis de distinção conceitual.

A WS apresenta-se, de acordo com a análise efetuada, como um constructo abrangente cujo objetivo final pode ser colocado como o de tornar a WWW uma base de dados global. Tornando possível, por um lado, uma pesquisa transversal (independente do tipo, formato e fonte dos dados), por outro, a obtenção de dados semanticamente interrelacionados, i.e., informação e não apenas uma listagem de documentos, muitas vezes sem ligação semântica entre si, como acontece na atual “web de documentos” (World Wide Web Consortium [W3C], 2013, 2015).

A esta *web*, resultante do desenvolvimento da WS, corresponde o conceito de *Web de Dados*, tal entendimento é reforçado pelo facto de se constatar que é esse o termo utilizado quando o objetivo da WS é mencionado (Berners-Lee, 1998; Bizer, Heath, & Berners-Lee, 2009; Hendler, Berners-Lee, & Miller, 2002; W3C, 2016).

Ao contrário da *Web de Dados*, ainda longe de se tornar efetiva, a *Web de Dados Ligados* está já em funcionamento (Hendler, 2014; Sletten, 2014), fazendo uso do conjunto de procedimentos e regras para a publicação de dados abertos (*Linked Open Data*), de especificações como RDF e SPARQL e das chamadas “ontologias leves”, tais como SKOS, FOAF ou Dublin Core (Bizer et al., 2009; Cambridge Semantics, 2016; W3C, 2016). A *Web de Dados Ligados* não é considerada como o objetivo da WS pois, ao contrário do que se pretende com a *Web de Dados*, a primeira mencionada apresenta grandes limitações na abrangência de ligações e, sobretudo, nas relações semânticas das mesmas (Hendler, 2011, 2014).

Pode-se, desta forma, entender a Web de Dados Ligados como uma parte da Web da Dados e esta como parte da WS que engloba, também, todas as especificações (padrões, linguagens, vocabulários, protocolos) desenvolvidas ou associadas à sua operacionalização, comumente designadas por “tecnologias da web semântica” (Bikakis, Tsinaraki, Gioldasis, Stavrakantonakis, & Christodoulakis, 2013; Parreiras, 2012).

A WS pretende, assim, potenciar a ideia original da WWW de disponibilizar o conhecimento humano, de forma abrangente, promovendo o trabalho colaborativo (Berners-Lee, 2009). Todavia, meio século antes desta *Web* digital, já o projeto da *Rede Universal de Documentação* de Paul Otlet, assenta nos mesmos ideais sendo considerada por vários autores como uma “web semântica” de madeira e cartão (Heuvel & Rayward, 2011; Laaff, 2011). Esta visão expandida da Documentação, de Otlet, é considerada, por muitos autores, como estando na gênese da área que, atualmente assume a nomenclatura de Ciência da Informação (CI) (Borges & Casado, 2009; A. M. da Silva & Ribeiro, 2008). Neste contexto, existe uma ligação entre o propósito da WS e as fundações teóricas da CI (Robredo, 2010; Souza & Alvarenga, 2004).

Concettualmente, a CI surge com frequência, na literatura, como uma área científica cujas demarcações, de conhecimento e de intervenção, se mostram complexas (Barreto, 2008; Saracevic, 1996; J. L. C. Silva & Freire, 2012; Zins, 2007a). Em termos epistemológicos a CI apresenta-se como uma área excessivamente flexível em termos de solidez disciplinar, a ponto de a comprometer (Galvão, 1998; Souza & Almeida, 2009). Parte desta problemática encontra-se relacionada com as intrincadas interações disciplinares presentes ao longo de todo o trajeto histórico da CI, em particular com a Biblioteconomia, a Arquivística e a Documentação (Borko, 1968; Le Coadic, 1996; Mikhailov & Giljarevskij, 1971; L. V. R. Pinheiro, 1997; A. M. da Silva & Ribeiro, 2008; Wersig & Neveling, 1975). Importa ressaltar que, relativamente ao conceito de disciplina, este é perspetivado segundo a visão epistemológica de Morin (2007), enquanto categoria autónoma de um corpo científico mais vasto.

Outra ligação disciplinar de relevo, que contribui para a complexidade conceitual da área da CI, é a que esta mantém com a Informática/Ciência da Computação (I/CC), a ponto da primeira ser “confundida” com a segunda (L. V. R. Pinheiro, 1997). Mesmo entre académicos relacionados com a área da CI é possível observar entendimentos similares, não só no que diz respeito à I/CC como à Biblioteconomia, à Comunicação ou à Matemática (Zins, 2007b).

Apesar destas díspares conceções da área, é possível, segundo Zins (2007a, 2007b, 2007c, 2007d), apresentar uma definição que reúne um razoável consenso entre especialistas contemporâneos, onde a CI é entendida como o estudo dos aspetos de mediação entre os utilizadores e os fenómenos dado, informação e conhecimento (enquanto artefactos humanos) e a sua implementação no domínio cultural/social (entendido como a forma da sociedade encarar a realidade).

A ligação entre a CI e a WS, embora concettualmente se afigura óbvia, em termos pragmáticos não se mostra tão evidente. O contexto teórico-prático que se consubstancia nas Dissertações de Mestrado (DM) e Teses de Doutoramento (TD) originárias de cursos de CI, cujo tema é a WS, é um campo de estudo profícuo, tanto mais que este tipo de estudos são indicadores importantes na aferição do desenvolvimento de uma área científica. Neste sentido, a reflexão que se propõe fornece indícios relevantes para a clarificação da relação CI-WS, em termos práticos e teóricos. A incidência do estudo em trabalhos produzidos em cursos de Portugal e do Brasil afigura-se pertinente, na medida em que estes países partilham afinidades históricas e culturais.

Assim, pretende-se: verificar e proceder a uma reflexão crítica no que respeita à presença da WS no *output* dos cursos de mestrado e doutoramento em Portugal e no Brasil em CI, considerados no período de 2005 a 2015. Para tal, propõe-se, caracterizar os repositórios agregadores de Portugal e do Brasil, respetivamente, RCAAP e OASISbr¹ dado serem a fonte de coleta de dados; identificar os cursos de mestrado e doutoramento em Portugal e no Brasil em CI certificados, respetivamente, pela A3ES e pela CAPES²; verificar o número de DMs e TDs em CI associados à WS, produzidas em Portugal e no Brasil (2005 a 2015), recuperadas de acordo com a metodologia desenhada para este estudo; identificar os pontos de confluência da CI com a WS, presentes nos trabalhos recuperados; caracterizar o perfil académico dos autores, orientadores e coorientadores dos trabalhos que constituem o corpus de estudo.

O desenho metodológico considerado para a efetivação destes propósitos assentou num paradigma qualitativo/pragmático, cuja descrição dos métodos e técnicas se apresentam no ponto que se segue.

METODOLOGIA

Desenhou-se uma metodologia de natureza qualitativa exploratória assente em duas fases. Na primeira recorreu-se a uma revisão bibliográfica para contextualizar os objetos de estudo: Web Semântica e Ciência da Informação, assim como para fundamentar teoricamente a parte empírica desenvolvida na segunda fase do estudo. A bibliografia recolhida foi analisada e cruzada. Relativamente à sua tipologia recorreu-se, essencialmente, a artigos científicos compreendidos, de um modo geral, entre 1990 e 2016.

No que se refere à segunda fase desenvolveu-se um estudo empírico. Iniciou-se por limitar a recolha dos dados aos repositórios RCAAP e OASISbr, dado que ambos se configuram como ponto de acesso à produção científica e académica nacional, multidisciplinar e de acesso aberto. Limitou-se, o estudo às DM e às TD, dado a, já referida, importância deste tipo de produção e sua ligação direta aos cursos superiores de uma área científica específica. Por fim, foi estabelecido como limite o período entre 2005 e 2015 pelo facto de, neste período, se ter assistido à evolução conceptual da Web Semântica, assim como da sua aplicabilidade em vários domínios do conhecimento, entre os quais na CI.

Seguiu-se a identificação dos cursos de mestrado e doutoramento através das estruturas oficiais associadas aos sistemas educativos dos respetivos países em estudo. Portugal: Direção Geral do Ensino Superior, Direção-Geral de Estatísticas da Educação e Ciência, Agência de Avaliação e Acreditação do Ensino Superior; Brasil: Conselho Nacional de Desenvolvimento Científico e Tecnológico, portal do Governo Federal do Brasil, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

Para a pesquisa foi selecionado o termo *web semântica*, uma vez que, após o cruzamento entre os termos de quatro glossários sobre a WS e os programas, conteúdos e ementas dos cursos de CI identificados, se verificou ser esse o termo que reunia a pertinência e relevância para os objetivos do estudo.

Procedeu-se a testes prévios nos dois repositórios, RCAAP e OASISbr, no sentido de aferir e ajustar os procedimentos a aplicar, nomeadamente, no que concerne à forma de recuperar possíveis variações do

¹ RCAAP - Repositório Científico de Acesso Aberto de Portugal, OASISbr - Portal Brasileiro de Acesso Aberto à Informação Científica.

² A3ES - Agência de Avaliação e Acreditação do Ensino Superior, CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

termo eleito, como aos campos e filtros a utilizar nas respetivas plataformas digitais. A escolha dos campos *Título* e *Assunto* garantiu uma uniformização de critérios entre repositórios e direcionou a pesquisa a trabalhos onde se assume que a WS tenha um papel central, dados os trabalhos recuperados apresentarem o respetivo termo em um, ou nos dois elementos relevantes de identificação das DM e TD. Os critérios usados na recolha de dados foram estabelecidos de acordo com os objetivos específicos dado a articulação intrínseca destes com o processo de recolha (Quivy & Campenhoudt, 1992).

Num total de 519 registos recuperados eliminaram-se 297 (duplicados ou falsos positivos) constituindo a amostra 222 trabalhos. Desta amostra, foi apurada uma subamostra de 21, número relativo aos trabalhos que reuniam as condições para o estudo — ser proveniente de um curso de CI e apresentar o termo *Web Semântica* no título e/ou nas palavras-chave. O corpus deste estudo é constituído por estes 21 trabalhos.

Na análise dos resultados foi aplicado um conjunto de técnicas e procedimentos metodológicos denominado de análise de conteúdo (Bardin, 2011), adotando-se níveis analíticos diferentes, porque o cumprimento dos objetivos assim o exigiram. Para determinar os pontos de confluência entre a CI e a WS e o grau de proximidade temática na abordagem efetuada nos 21 trabalhos que constituíram o corpus do estudo, analisaram-se (nível mais exaustivo) os seus resumos e respetivos elementos internos³, e os objetivos, expressos na introdução dos documentos. Estabeleceram-se critérios para determinar os pontos de confluência grau de confluência e o grau de proximidade temática entre a CI e a WS (ver figura 1), expressos em uma escala de intensidade negativa (Bardin, 2011).

c1: Critérios usados na análise dos pontos de confluência da Web Semântica (WS) com a Ciência da Informação (CI), expressos nos resumos dos trabalhos constituintes da amostra.										
Presença de pontos de confluência no resumo	Frequência mínima								Grau de confluência atribuído	
	de ligações (entre os conceitos CI e WS)					dos termos			quantitativo	qualitativo
	diretas	e/ou	indiretas	e/ou	c/intermediário	CI	WS	tipo de relação		
Sim	2	e	indiferente			n.a	n.a	n.a	0	maior
	1	ou	3	e	indiferente	n.a	n.a	n.a	-1	↓
	0	e	1	ou	3	n.a	n.a	n.a	-2	↓
	0	e	0	e	1	n.a	n.a	n.a	-3	↓
Não	n.a	n.a	n.a	n.a	n.a	1	1	direta	-4	↓
	n.a	n.a	n.a	n.a	n.a	1 deles		indireta	-5	↓
	n.a	n.a	n.a	n.a	n.a	0	0	n.a	-6	menor
<i>n.a: não aplicável</i>										
c2: Critérios usados na análise da centralidade temática da Web Semântica (WS), expressa nos objetivos dos trabalhos constituintes da amostra.										
Função do conceito WS desempenhada nos objetivos	Presença/frequência do termo WS no(s) objetivo(s):						Grau de proximidade atribuído			
	Geral			Específicos			quantitativo	qualitativo		
Configura-se como o próprio objeto da ação	Apresenta						Em 50% ou mais		0	maior
Enquanto objeto da ação apenas se encontra nos objetivos específicos							Em menos de 50%		-1	↓
Surge como auxiliar do objeto da ação ou como elemento contextual	Surge no geral ou em um dos específicos e não nos dois em simultâneo						-2	↓		
Relação conceitual não explicitada	Não apresenta			Igual a 0%			-3	↓		
							-4	menor		

Figura 1. Critérios de análise dos resumos (c1) e objetivos (c2) dos trabalhos e respetiva escala gradativa

³ Introdução, objetivos, metodologia, resultados e conclusões.

Relativamente ao perfil académico dos autores, orientadores e coorientadores dos 21 trabalhos (nível analítico de grau moderado), a análise incidiu na formação académica dos mesmos. Consideraram-se como fontes para a recolha desses dados as plataformas digitais: *DeGóis*, para docentes portugueses e a *Lattes*, para docentes brasileiros.

Confrontou-se, posteriormente, esta última análise com a soma das duas anteriores (aos resumos e aos objetivos), no sentido de verificar a existência de uma correlação entre o grau de confluência/proximidade da relação WS-CI e a diferença quantitativa das formações em I/CC relativamente às formações em CI dos autores e orientadores de cada trabalho. Utilizou-se uma escala bipolar onde o polo de intensidade positiva máxima corresponde a um maior peso da área I/CC na formação e o polo inverso ao peso da área CI.

Por fim, com um nível analítico de menor grau, verificou-se a presença da Web Semântica enquanto tema dos trabalhos recuperados de acordo com as áreas científicas.

RESULTADOS

Verificou-se, quanto aos repositórios RCAAP e OASISbr, diferenças tanto no enquadramento normativo como nas funcionalidades (o segundo apresenta, e.g., mais opções de “pesquisa avançada”). Observou-se a existência de trabalhos recuperáveis apenas em um dos repositórios (quatro no RCAAP e 40 no OASISbr), ao contrário do espectável dado ambos agregarem os documentos existentes no outro.

Relativamente aos cursos de CI, ativos em 2016, identificaram-se quatro de mestrado e três de doutoramento em Portugal e, no Brasil, 16 de mestrado e 9 de doutoramento (ver tabela 1).

Tabela 1. Cursos de mestrado e doutoramento em CI, ativos em 2016, em Portugal e no Brasil certificados, respetivamente, pela A3ES e pela CAPES

Instituição de Ensino Superior	Ref. ^a	Designação do Curso	Tipologia ^b
Universidade de Coimbra	P.01	Ciência da Informação	M & D
Fundação Ensino e Cultura Fernando Pessoa	P.02	Ciência da Informação	D
Universidade do Porto	P.03	Ciência da Informação	M
Universidade do Porto & Universidade de Aveiro ^c	P.04	Informação e Comunicação em Plataformas Digitais	D
Universidade de Lisboa	P.05	Ciências da Documentação e Informação	M
Universidade do Algarve	P.06	Ciências Documentais	M
Universidade de São Paulo	B.01	Ciência da Informação	M & D
Universidade Estadual Paulista Júlio de Mesquita Filho/Marília	B.02	Ciência da Informação	M & D
Universidade Estadual de Londrina	B.03	Ciência da Informação	M
Universidade Federal da Bahia	B.04	Ciência da Informação	M & D
Universidade Federal da Paraíba/João Pessoa	B.05	Ciência da Informação	M & D
Universidade Federal de Pernambuco	B.06	Ciência da Informação	M
Universidade Federal Santa Catarina	B.07	Ciência da Informação	M & D
Universidade Federal Fluminense	B.08	Ciência da Informação	M & D
Universidade Federal do Rio de Janeiro	B.09	Ciência da Informação	M & D
Universidade de Brasília	B.10	Ciência da Informação	M & D
Universidade Federal de Minas Gerais	B.11	Ciência da Informação	M & D

Universidade do Estado de Santa Catarina	B.12	Gestão da Informação	M(P)
Universidade Federal do Rio Grande do Norte	B.13	Gestão da Informação e do Conhecimento	M(P)
Fundação Casa de Rui Barbosa	B.14	Memória e Acervos	M(P)
Universidade Federal do Estado do Rio de Janeiro	B.15	Biblioteconomia	M(P)
	B.16	Gestão de Documentos e Arquivos	M(P)

^a P.- Portugal; B.- Brasil. ^b Tipologia: M - mestrado; D - doutoramento; M(P) - mestrado profissional. ^c Curso lecionado em parceria pelas duas instituições.

Tendo-se verificado, nos cursos de Portugal, a continuação da utilização de nomenclaturas como *Ciência da Documentação* ou *Ciências Documentais* e a inexistência, nestas designações, de termos como *Biblioteconomia*, *Arquivística* ou *Gestão da Informação*. No Brasil, estes últimos termos, ou similares, surgem associados a mestrados profissionais, enquanto o termo *Ciência da Informação* é adotado por todos os cursos (de ambos os graus) classificados como académicos. Interessa, ainda, constatar que os dois cursos da Universidade Federal de Rio de Janeiro, *Biblioteconomia* e *Gestão de Documentos e Arquivos*, deixaram de ser incluídos na área do conhecimento da CAPES *Ciência da Informação*, em dezembro de 2015, passando estes cursos a estar integrados, respetivamente, nas novas áreas *Biblioteconomia* e *Arquivologia* (designação usada no Brasil), indo ao encontro da proposta de alteração da Tabela de Áreas do Conhecimento (TAC), efetuada em 2005 (ver tabela 2).

Tabela 2. Relação entre a Ciência da Informação, a Biblioteconomia e a Arquivística nas versões da TAC da CAPES

Ano da versão	Ciência da Informação (CI)	Biblioteconomia e Arquivística
1976	Subárea da <i>Comunicação</i>	Especialidades da área CI
1982	Uma área conjunta: <i>Ciência da Informação, Biblioteconomia e Arquivologia</i>	
1984	Área	Subáreas da área CI
2005 ^a	Três áreas autónomas: (i) <i>Ciência da Informação</i> ; (ii) <i>Biblioteconomia</i> ; (iii) <i>Arquivologia</i>	

^a Proposta preliminar da Comissão Especial de Estudos das entidades: CNPq (Conselho Nacional de Desenvolvimento Científico), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) e FINEP (Financiadora de Estudos e Projetos).

Nas DM e TD de CI cujo tema é a WS recuperaram-se, respetivamente, 16 e cinco, produzidas no Brasil. Estes 21 trabalhos (ver tabela 3) representam 12% de um total de 173 desenvolvidos neste país (cf. figura 2).

Tabela 3. Identificação dos 21 trabalhos que compõem a amostra do presente estudo

Ref. ^a	Ano ^b	Título do trabalho	Curso ^c
dm.01	2005	Web semântica: Uma análise focada no uso de metadados	B.02
dm.02	2006	Web semântica: aspetos interdisciplinares da gestão de recursos informacionais no âmbito da Ciência da Informação	B.02
td.01	2006	Autoria de documentos para a Web semântica: Um ambiente de produção de conhecimento baseado em ontologias	B.10
dm.03	2008	A construção do conhecimento científico: a Web Semântica como objeto de estudo	B.02
dm.04	2008	Inserção da biblioteca digital de teses e dissertações no contexto da Web Semântica	B.12
dm.05	2008	Padrões de representação e descrição de recursos informacionais em bibliotecas digitais na perspetiva da Ciência da Informação: uma abordagem do <i>marcont initiative</i> na era da Web Semântica	B.02
dm.06	2009	Perspetivas na Web Semântica para a Ciência da Informação	d.01 ^d
dm.07	2010	Web semântica e o governo eletrônico brasileiro: Revisão histórica e teórico-conceitual da representação descritiva e estudo acerca da taxonomia e adoção de metadados na representação da informação	B.04
td.02	2010	Representação iterativa: Um modelo para repositórios digitais	B.02
dm.08	2011	Estudo exploratório sobre compatibilização semântica de vocabulário utilizando ontologia: Fundamentação teórica-metodológica visando ao aperfeiçoamento do MHTX	B.11

dm.09	2011	Registros de marcas da junta comercial do estado de Pernambuco: Um olhar semântico para a organização de um acervo histórico	B.06
dm.10	2012	Artigos científicos digitais na web: Novas experiências para apresentação, acesso e leitura	B.08
dm.11	2012	As dimensões semânticas e pragmática da web e dos mecanismos de busca no ciberespaço	d.02 ^d
dm.12	2012	Organização em sistema caótico: Uso das <i>tags</i> para classificação da informação pelos usuários da rede social <i>twitter</i>	B.09
dm.13	2013	Arquitetura de repositório semântico de dados para organização de pesquisa agropecuária	B.09
dm.14	2013	As tecnologias da Web Semântica no domínio bibliográfico	B.02
dm.15	2013	Representação da informação dinâmica em ambientes digitais	B.01
td.03	2013	Web Semântica e repositórios digitais educacionais na área de saúde: Uma modelagem com foco no objetivo de aprendizagem para refinar resultados de busca	B.10
dm.16	2014	Wikis semânticos: Da web para a Web Semântica	B.02
td.04	2015	Interoperabilidade e mapeamentos entre sistemas de organização do conhecimento na busca e recuperação de informações em saúde: Estudo de caso em ortopedia e traumatologia	B.01
td.05	2015	Representação e recuperação de informação em acervos digitais nos contextos da Web Semântica e web pragmática: Um estudo crítico	B.01

^a dm: dissertação de mestrado; td tese de doutoramento. ^b ano de defesa do trabalho. ^c referência que consta na Tabela 1. ^d cursos que em 2016 se encontravam desativados: d.01- Ciência da Informação (Pontifícia Universidade Católica de Campinas); d.02- Gestão da Informação (Universidade Estadual de Londrina).

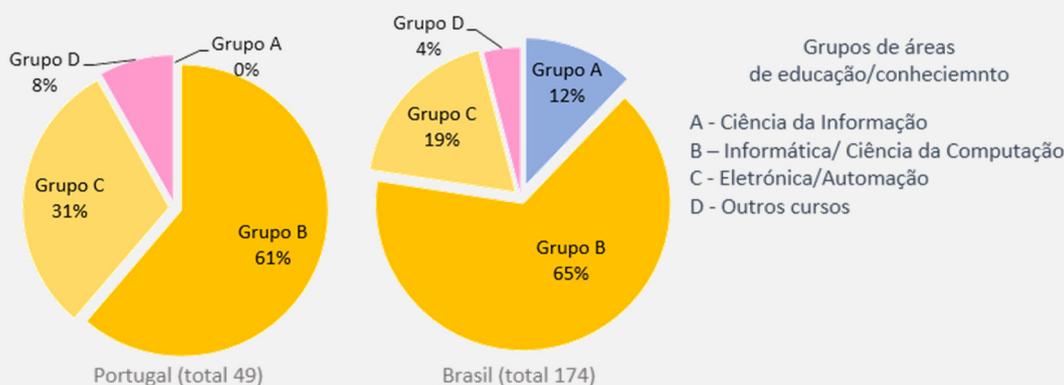


Figura 2. Distribuição percentual dos trabalhos recuperados (DM e TD), de acordo com o país de origem, pelos grupos de áreas de educação/conhecimento

Em Portugal, a inexistência de trabalhos cujo tema é a WS, produzidos em cursos considerados de CI, contrasta com os quatro (8% de 49) provenientes das áreas: Ciências da Educação; Ciências Dentárias; Gestão e Administração; e Tecnologia dos Processos Químicos, (incluídas no grupo D).

Identificaram-se três pontos de confluência entre a CI e a WS: (i) contribuição mútua entre as duas áreas (62%), (ii) a WS enquanto tema interdisciplinar entre a CI e a I/CC (29%) e (iii) a utilização conjunta de conhecimentos das duas áreas (9%). Na distribuição temporal destes pontos pelo elemento onde foram encontrados (ver figura 3) destaca-se o primeiro ponto de confluência pela presença ao longo de todo período analisado. Este ponto apresenta 54% das suas ocorrências no elemento *objetivos* (entre 2006 e 2013) e é o único a ser referido no elemento *introdução* (em 2014). Relativamente ao terceiro ponto, verifica-se que as duas únicas ocorrências surgem no último ano, 2015, ambas no elemento *metodologia* de teses de doutoramento.



Figura 3. Distribuição temporal dos pontos de confluência pelos elementos do resumo onde foram identificados

Relativamente ao perfil académico dos 17 orientadores e dois coorientadores dos trabalhos considerados neste estudo, apurou-se que: dois (10%) apresentam formação só em CI; 11 (58%) em CI e outras áreas; e seis (32%) só em outras áreas. A formação na área da I/CC surge em sete (35%) docentes, sendo que destes apenas um acumula essa formação com a da área de CI. Nos 21 autores, contabilizaram-se: 10 (48%) que apresentam formação só em CI e 11 (52%) em CI e outras áreas. A formação na área da I/CC surge em nove (43%) autores, todas obtidas em graus anteriores ao mestrado.

Em termos de correlação, verificou-se a existência de uma bastante forte (expressa no coeficiente 0,781) entre as duas variáveis analisadas nos trabalhos: o grau de confluência encontrado nos resumos e o grau de proximidade determinado nos objetivos, ao contrário da correlação muito residual (coeficiente 0,145) entre a diferença quantitativa das formações em I/CC relativamente às de CI e o grau de confluência/proximidade da relação WS-CI (ver figura 4).

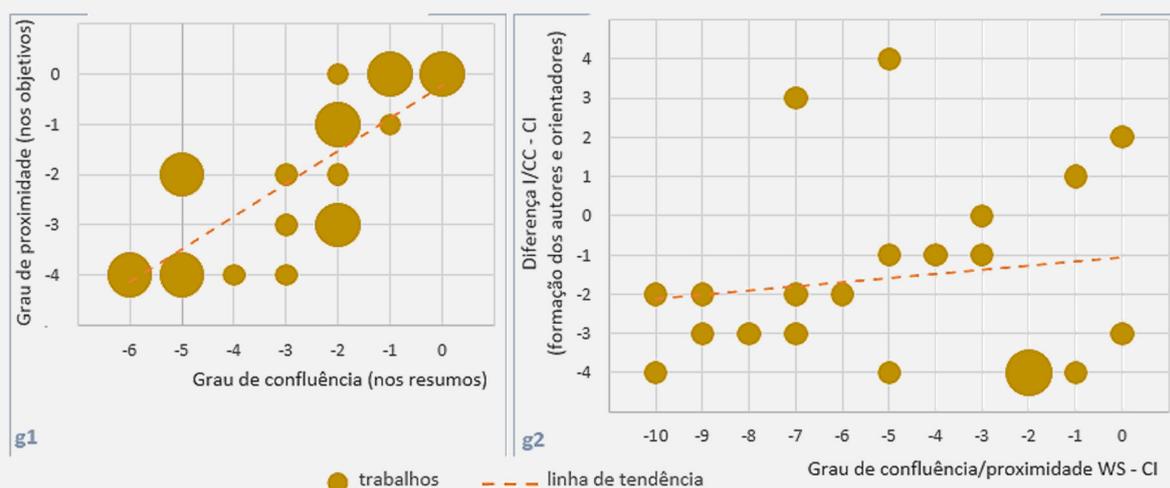


Figura 4. Gráficos de correlação entre pontos de confluência WS-CI e centralidade temática da abordagem à WS (g1) e entre relação WS-CI e diferença quantitativa das formações I/CC-CI (g2)

A confrontação efetuada entre a formação dos responsáveis pelos trabalhos e o grau de proximidade aferido, da relação WS-CI, permitiu também verificar uma correlação ainda mais afastada (coeficiente 0,092) se a ponderação for o equilíbrio entre os pesos de ambas as formações em vez da aproximação à I/CC.

CONCLUSÕES

Relativamente às fontes de obtenção de dados (RCAAP e OASISbr), conclui-se existir discrepâncias entre a informação que consta nos campos *assuntos* e *data* destes e os campos correspondentes nos repositórios das Instituições de Ensino Superior (RIES) de origem dos documentos. Discrepâncias com impacto nos trabalhos recuperados que, infere-se, resultam da existência, em vários RIES, de mais de um campo com informação passível de ser considerada “assunto” assim como com vários campos com diferentes datas registadas.

Conclui-se que em Portugal o panorama apresenta uma mudança de atitude no que concerne aos cursos da área de CI. A posição inicial pode ser descrita como uma associação entre a visão de Otlet (1934) e a posição de López Yepes (1995), i.e., uma visão da Documentação enquanto área que abarca as outras áreas relacionadas “elevando” o carácter científico desse conjunto às “Ciências Documentais”. A posição de chegada, pode ser apresentada como uma visão holística da área, materializando-se na CI enquanto entidade que incorpora aqueles saberes e outros, num todo maior que a soma das suas partes. No Brasil, conclui-se que existiu um percurso relativamente fluído da Documentação para a CI. Quanto ao relacionamento desta com a Biblioteconomia e a Arquivística, observa-se, o que se infere ser, uma procura de identidade da CI enquanto um todo e, simultaneamente, uma tentativa de afirmação por parte daquelas áreas disciplinares específicas.

No que respeita à distribuição da amostra pelos dois países, conclui-se existir um desfasamento maior que o esperado entre ambos, dado que todos os trabalhos recuperados considerados de CI, cujo tema é a WS, foram produzidos no Brasil. Assim, infere-se a existência de uma relação entre os contextos de origem da CI, europeu, para Portugal, e americano, no caso do Brasil, e a diferente ligação à I/CC, área dominante no que diz respeito ao tratamento do assunto em questão.

Relativamente aos três pontos de confluência entre as duas áreas (CI e WS) identificados (i: contribuição mútua; ii: a WS enquanto tema interdisciplinar; iii: utilização conjunta de conhecimentos), conclui-se, atendendo à relevância percentual dos mesmos, sua distribuição ao longo do período analisado e o elemento do resumo onde é referido, que a relação entre a CI e a WS apresenta-se como um campo, essencialmente, de investigação teórica. Dentro da amostra analisada, apenas em 2014 a contribuição recíproca entre as duas áreas é tomada como premissa deixando de ser uma hipótese a investigar e, em 2015, surgem trabalhos que se propõem fazer uso efetivo dos conhecimentos das duas áreas.

Em termos do perfil académico dos autores e orientadores dos trabalhos analisados, conclui-se que a amostra (21) é pouco representativa do carácter interdisciplinar da WS enquanto tema trabalhado pelas áreas da CI-I/CC, atendendo a que em apenas 11 trabalhos (52%) os investigadores responsáveis apresentam, no conjunto, formação nestas duas áreas (10% se apenas for atendido a formação dos orientadores) e a correlação quase inexistente desta com uma maior confluência/proximidade entre a WS e a CI.

Assim, da análise da amostra, conclui-se que, embora se possa entender a WS como um campo de estudo de direito próprio da CI, este configura-se, essencialmente, de investigação teórica sendo o trabalho prático interdisciplinar, CI-I/CC, incipiente e pouco representativo.

Considera-se ter contribuído com: uma clarificação de conceitos relativos à WS, indicadores sobre o seu estudo, na área CI, e relativos à relação CI-I/CC; subsídios para uma reflexão sobre as relações

disciplinares da CI no contexto dos cursos de ensino superior; alguns indicadores relacionados com os repositórios usados, nomeadamente quanto à ausência ou duplicação de registos. Questões que se julga merecedoras de futuros aprofundamentos pelas suas implicações tanto teóricas como práticas.

REFERÊNCIAS BIBLIOGRÁFICAS

- Bardin, L. (2011). *Análise de conteúdo*. (L. A. R. A. Pinheiro, Trans.). São Paulo: Almedina.
- Barreto, A. de A. (2008). Uma quase história da ciência da informação. *DataGramaZero*, 9(2), 1–18. Retrieved from <http://eprints.rclis.org/17637/1/DataGramaZero> - Revista de Ciência da Informação - Artigo 01_Aldo.pdf
- Berners-Lee, T. (1998, October 14). A roadmap to the Semantic Web. Retrieved March 21, 2016, from <https://www.w3.org/DesignIssues/Semantic>
- Berners-Lee, T. (2006). The World Wide Web: Past, present and future (FAQ). Retrieved March 16, 2016, from <http://www.w3.org/People/Berners-Lee/FAQ.html>
- Berners-Lee, T. (2009). Web 2.0 Summit 09 discussion: A Conversation with Tim Berners-Lee interview by Tim O'Reilly [video]. San Francisco: O'Reilly Media. Retrieved from <https://youtu.be/KY5skobffk0>
- Berners-Lee, T., & Fischetti, M. (2000). *Weaving the Web*. HarperCollins.
- Berners-Lee, T., & Hendler, J. A. (2001). Publishing on the Semantic Web. *Nature*, 410(April), 1023–1024. Retrieved from <https://github.com/AKSW/semann/blob/master/pdf/T.B. Lee - Publishing on the semantic web.pdf>
- Berners-Lee, T., Hendler, J. A., & Lassila, O. (2001). The Semantic Web. *Scientific American*, 284(5), 35–43. Retrieved from http://ldc.usb.vt/~yudith/docencia/UCV/ScientificAmerican_FeatureArticle_TheSemanticWeb_May2001.pdf
- Bikakis, N., Tsinaraki, C., Gioldasis, N., Stavrakantonakis, I., & Christodoulakis, S. (2013). The XML and Semantic Web Worlds: Technologies, Interoperability and Integration. A survey of the state of the art. In I. E. Anagnostopoulos, M. Bieliková, P. Mylonas, & N. Tsapatsoulis (Eds.), *Semantic Hyper/Multimedia Adaptation* (pp. 319–360). Berlin: Springer. https://doi.org/10.1007/978-3-642-28977-4_12
- Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked data - The story so far. University of Southampton, ePrints Soton. Retrieved from <http://eprints.soton.ac.uk/271285/>
- Borges, M. M., & Casado, E. S. (Eds.). (2009). *A Ciência da Informação criadora de conhecimento* (Vol. I). Coimbra: Imprensa da Universidade de Coimbra. <https://doi.org/http://dx.doi.org/10.14195/978-989-26-0319-3>
- Borko, H. (1968). Information Science: What Is It? *American Documentation*, 19(1), 3–5. Retrieved from <http://cdigital.uv.mx/bitstream/123456789/6699/2/Borko.pdf>

Cambridge Semantics. (2016). Many Names for the Semantic Web. Retrieved May 15, 2017, from <http://www.cambridgesemantics.com/semantic-university/many-names-semantic-web>

Caro-castro, C. (2010). Las clasificaciones bibliográficas: De los estantes a la web. *Tabula*, 13, 11–23. Retrieved from <http://hdl.handle.net/10366/121954>

Galvão, M. C. B. (1998). Construção de conceitos no campo da ciência da informação. *Ciência Da Informação*, 27(1), 46–52. <https://doi.org/10.1590/S0100-19651998000100006>

Hendler, J. A. (2011). “Why the Semantic Web will never work” (note the quotes) [Video]. Retrieved from <https://youtu.be/oKiXpO2rbJM>

Hendler, J. A. (2014). The Semantic Web: The Inside Story [video]. In *Summer School in cognitive Science: Web Science and the Mind*. Montréal: Institut des sciences cognitives, UQAM. Retrieved from <https://youtu.be/3Ap5FsxvjTQ>

Hendler, J. A., Berners-Lee, T., & Miller, E. (2002). Integrating Applications on the Semantic Web. *Journal of the Institute of Electrical Engineers of Japan*, 122(10), 676–680. Retrieved from <http://www.w3.org/2002/07/swint>

Heuvel, C. Van Den, & Rayward, W. B. (2011). Facing interfaces: Paul Otlet’s visualizations of data integration. *International Review of Research in Open and Distance Learning*, 62(12), 2313–2326. <https://doi.org/http://dx.doi.org/10.1002/asi.21607>

Laaff, M. (2011, July 22). Internet visionary Paul Otlet: Networked knowledge, decades before Google. (E. Ornstein, Trans.), *SPIEGEL ONLINE - International*. Retrieved from <http://www.spiegel.de/international/world/internet-visionary-paul-otlet-networked-knowledge-decades-before-google-a-775951.html>

Le Coadic, Y.-F. (1996). *A ciência da informação*. (M. Y. F. Gomes, Trans.). Brasília: Briquet: de Lemos/Livros.

López Yepes, J. (1995). *La Documentación como disciplina: Teoría e historia* (2nd ed.). Pamplona: Ediciones Universidade de Navarra.

Mikhailov, A. I., & Giljarevskij, R. S. (1971). An Introductory Course on Informatics/Documentation. *International Federation for Documentation*. Retrieved from <http://eric.ed.gov/?id=ED060875>

Morin, E. (2007). *Educação e complexidade: Os Sete Saberes e outros ensaios*. (M. da C. de Almeida & E. de A. Carvalho, Eds., E. de A. Carvalho, Trans.) (4th ed.). São Paulo, Brasil: Cortez Editora.

Otlet, P. (1934). *Traité de documentation: Le livre sur le livre: Théorie et pratique*. Bruxelles: Ediciones Mundaneum Palais Mondial. Retrieved from http://lib.ugent.be/fulltxt/handle/1854/5612/Traite_de_documentation_ocr.pdf

Parreiras, F. silva. (2012). *Semantic Web and Model-driven Engineering*. Piscataway: John Wiley & Sons, Inc.

Patel-Schneider, P. F., & Horrocks, I. (2007). A comparison of two modelling paradigms in the Semantic Web. *Web Semantics*, 5(4). <https://doi.org/http://dx.doi.org/10.1016/j.websem.2007.09.004>

Pinheiro, L. V. R. (1997). *A Ciência da Informação entre sombra e luz: Domínio epistemológico e campo interdisciplinar*. Universidade Federal do Rio de Janeiro. Retrieved from <http://ridi.ibict.br/handle/123456789/35>

Quivy, R., & Campenhoudt, L. Van. (1992). *Manual de investigação em ciências sociais*. (J. M. Marques, M. A. Mendes, & M. Carvalho, Trans.) (6th ed.). Lisboa: Gradiva.

Robredo, J. (2010). Ciência da informação e Web semântica: Linhas convergentes ou linhas paralelas? In *Passeios pelo bosque da informação: estudos sobre representação e organização da informação e do conhecimento* (pp. 12–47). Brasília: IBICT. Retrieved from <http://repositorio.unb.br/handle/10482/7952>

Saracevic, T. (1996). Ciência da informação: origem, evolução e relações. *Perspectivas Em Ciência Da Informação*, 1(1), 41–62. Retrieved from <http://www.brapci.inf.br/article.php?dd0=0000003224&dd90=bc40e71f02>

Shadbolt, N., Hall, W., & Berners-Lee, T. (2006). The semantic web revisited. *IEEE Intelligent Systems*, 21(3), 96–101. <https://doi.org/10.1109/MIS.2006.62>

Silva, J. L. C., & Freire, G. H. de A. (2012). Um olhar sobre a origem da ciência da informação: indícios embrionários para sua caracterização identitária. *Encontros Bibli*, 17(33), 1–29. <https://doi.org/10.5007/1518-2924.2012v17n33p1>

Silva, A. M. da, & Ribeiro, F. (2008). *Das “ciências” documentais à ciência da informação: Ensaio epistemológico para um novo modelo curricular* (2nd ed.). Porto: Edições Afrontamento.

Sletten, B. (2014, January 13). Keep On Keeping On. Dataversity, Smart Data. Retrieved from <http://www.dataversity.net/keep-on-keepin-on/>

Souza, R. R., & Almeida, M. B. (2009). Representação do conhecimento: identidade ou esvaziamento da Ciência da Informação? In *A Ciência da Informação criadora de conhecimento, vol.1, IV Encontro Ibérico EDIBCIC 2009* (pp. 157–165). Coimbra: Imprensa da Universidade de Coimbra.

Souza, R. R., & Alvarenga, L. (2004). A Web Semântica e suas contribuições para a ciência da informação. *Ciência Da Informação*, 33(1), 132–141. Retrieved from <http://ref.scielo.org/mp2fh8>

Wersig, G., & Neveling, U. (1975). The phenomena of interest to Information Science. *Information Scientist*, 9(4). Retrieved from <http://sigir.org/files/museum/pub-13/18.pdf>

World Wide Web Consortium. (2013, December 11). W3C Semantic Web Activity Homepage. Retrieved January 27, 2016, from <https://www.w3.org/2001/sw/>

World Wide Web Consortium. (2015). W3C Standards: Semantic Web. Retrieved March 21, 2016, from <https://www.w3.org/standards/semanticweb/>

World Wide Web Consortium. (2016). W3C Data Activity: Building the Web of Data. Retrieved March 21, 2016, from <http://www.w3.org/2013/data/>

Zins, C. (2007a). Classification Schemes of Information Science: Twenty-Eight Scholars Map the Field. *Journal of the American Society for Information Science and Technology*, 58(4), 645–672. <https://doi.org/10.1002/asi.20508>

Zins, C. (2007b). Conceptions of Information Science. *Journal of the American Society for Information Science and Technology*, 58(4), 335–350. <https://doi.org/10.1002/asi.20508>

Zins, C. (2007c). Conceptual Approaches for Defining Data, Information, and Knowledge. *Journal of the American Society for Information Science and Technology*, 58(4), 479–493. <https://doi.org/10.1002/asi.20508>

Zins, C. (2007d). Knowledge Map of Information Science. *Journal of the American Society for Information Science and Technology*, 58(4), 526–535. <https://doi.org/10.1002/asi.20505>



EL ÁRBOL DE DOMINIO COMO HERRAMIENTA EN ORGANIZACIÓN DEL CONOCIMIENTO: APLICACIÓN AL ÁREA DE EDUCACIÓN SUPERIOR

Mario Barité¹, Mirtha Rauch²

¹Universidad de la República de Uruguay, mario.barite@fic.edu.uy

²Universidad de la República de Uruguay, mrauch@gmail.com

RESUMO Dos son las modalidades tradicionales de ordenación de términos de clasificación/indización en el cuerpo de los sistemas de organización del conocimiento: la alfabética y la sistemática. En cualquiera de las dos modalidades, los clasificacionistas necesitan establecer orientaciones, principios y métodos para seleccionar, ordenar y relacionar las voces especializadas, es decir, los términos. Desde hace veinticinco años se han dado variadas aproximaciones teóricas y metodológicas entre la Organización del Conocimiento y la Terminología. Una contribución desde la Terminología es la utilización de árboles de dominio como herramientas metodológicas necesarias en el proceso de elaboración de productos terminológicos. Este trabajo se enfoca en el proceso de construcción del árbol de dominio que acompaña y controla la creación y desarrollo de un banco de datos terminológicos especializado en Educación Superior. Se trata de un proyecto en curso que toma como base la terminología habitual de los países del Cono Sur americano, y que puede ofrecer como productos en el futuro tesauros, listas de descriptores, diccionarios u otros sistemas de organización del conocimiento especializados en Educación Superior. De las 65 fuentes del corpus analizadas hasta el presente se extrajeron 510 términos, que fueron organizados en un árbol de dominio conforme a procedimientos *bottom-up*, en siete grandes subclases o subramas. Se presenta un cuadro con la distribución de términos por rama. Entre otras conclusiones, se afirma que la claridad del método de construcción del árbol y sus resultados, confirman su valor como herramienta de encuadre, relación y control terminológico. Al servir para fines de organización y sistematización de terminología, los árboles de dominio podrían integrarse a los métodos de construcción de tesauros, taxonomías y ontologías, y ser considerados, en consecuencia, por los responsables de normas y estándares.

PALAVRAS-CHAVE *Árvore de domínio, Controle de vocabulário, Educação Superior.*

ABSTRACT ATraditionally, there are two ways of organizing terms in Knowledge Organization Systems: alphabetically and systematically. In either of these two arrangements, classifiers need to establish guidelines, principles and methods to select, sort and relate specialized words, i.e. terms. In the last 25 years, there have been various theoretical and methodological approaches between Knowledge Organization and Terminology. A contribution from Terminology has been the use of domain trees as crucial methodological tools in the process of development of terminological products. This paper focuses on the building process of the domain tree which goes along with and controls the creation and development of a terminological data bank specialized in Higher Education. It is an ongoing project based upon the usual terminology of the Southern Cone countries which will have the possibility to provide in the future products such as thesauri, lists of descriptors, dictionaries or other systems of knowledge organization specialized in Higher Education. 510 terms have been extracted until now from the 65 sources making up the corpus and then organized in a domain tree according to the bottom-up procedures in seven big subclasses or sub branches. A table with the distribution of terms by branch is presented. Among other conclusions, it is stated

that the clarity of the construction method of the tree and its results confirm its value as a tool of terminological framework, relationship and control. As they are useful for terminology organization and systematization, domain trees could be part of the methods of construction of thesauri, taxonomies and ontologies and consequently be considered by those responsible of establishing rules and standards.

KEYWORDS *Domain tree, Vocabulary control, Higher Education.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUCCIÓN

Dos son las modalidades tradicionales de ordenación de términos de clasificación/indización en el cuerpo de los sistemas de organización del conocimiento (SOC): i) la alfabética, que continúa la tradición enciclopedista de presentación de tópicos conforme al orden arbitrario del alfabeto. En esta modalidad se incluyen las listas (de encabezamientos de materias, de descriptores y de autoridades), los índices (temáticos, onomásticos), y gran parte de los glosarios y diccionarios, tanto los generales de las lenguas como los especializados; ii) la sistemática, fundada sobre organizaciones taxonómicas o funcionales de tópicos, cuyo antecedente formal más antiguo es el llamado Árbol de Porfirio, el cual se basaba “*en una ley elemental de la lógica: la dicotomía*” (San Segundo, 1996, p. 38). La ordenación sistemática sostiene las estructuras conceptuales de los sistemas de clasificación, los tesauros, las taxonomías y las ontologías. En cualquiera de las dos modalidades, los clasificacionistas necesitan establecer orientaciones, principios y métodos para seleccionar, ordenar y relacionar las voces especializadas, es decir, los términos.

Desde hace veinticinco años se han dado variadas aproximaciones entre la Organización del Conocimiento y la Terminología (Dahlberg, 1992; Cabré, 1993; Campos, 2001; Alexiev & Marksby, 2010; Woźniak-Kasperek, 2014), probablemente fundadas en que ambos campos temáticos trabajan sobre el mismo material de origen: los términos como etiquetas de representación, y el conocimiento especializado, debidamente registrado en fuentes documentales legitimadoras de los términos propios de las disciplinas (documentos legales, patentes, normas técnicas, revistas científicas, obras canónicas, *proceedings* y *advances*).

Si bien el objeto de estudio de la Terminología son los términos (Cabré, 1993), y el de la Organización del Conocimiento es el conocimiento registrado en documentos, el punto de conexión está dado en que los términos son, en ambos campos temáticos, las claves para construir -a la manera de ladrillos- el edificio semántico que forman las estructuras conceptuales de las disciplinas. Es sabido que las áreas del saber se construyen no solo a partir de la identificación de sus términos de referencia, sino también de las relaciones -estables, aceptadas y verificadas, pero también las contingentes u ocasionales- que se establecen entre ellos. Como señala Fedor de Diego, “*en terminología, los conceptos no se consideran como elementos aislados; cada concepto individual es parte integrante de un sistema conceptual que a la vez refleja la ordenación, la estructuración de algún campo del conocimiento*” (Fedor de Diego, 1995, p. 39). Ese es el fundamento racional de la metodología de desarrollo de tesauros y de las operaciones de control de vocabulario, que exigen que todo descriptor esté relacionado con otros de alguna forma, jerárquica, sinonímica o asociativa. La norma ISO relativa a principios y métodos de la terminología reafirma esas ideas cuando señala que “*la terminología de un dominio (...) no debe constituir un conjunto arbitrario de términos, sino más bien un sistema terminológico coherente y*

correspondiente a su sistema nocional” (International Organization for Standardization, 1987, p. 11). De hecho, tanto en investigaciones terminológicas como en los procesos de creación o revisión de sistemas de organización del conocimiento se evita la presencia de términos huérfanos, aquellos que en una presentación final de una estructura conceptual, quedarían aislados por no tener establecida con ningún otro.

Otros puntos de conexión entre Terminología y Organización del Conocimiento son de naturaleza metodológica y no han sido suficientemente explorados en la literatura. La Organización del Conocimiento, por ejemplo, ofrece estándares de alta formalidad para la construcción de vocabularios monolingües o plurilingües (British Standard Institution, 2005; International Organization for Standardization, 2011). La Terminología, por su parte, aporta una metodología elaborada para la selección de un corpus de investigación, así como métodos y técnicas de extracción y selección de términos, todo lo cual puede aprovecharse en la creación de tesauros, taxonomías y listas. Pero además contribuye con la utilización de árboles de dominio como herramientas metodológicas necesarias en el proceso de elaboración de productos terminológicos.

La noción de árbol de dominio está instalada desde hace por lo menos cuarenta años en la Terminología (Auger et Rousseau, 1978), y su inclusión en la metodología de creación de glosarios y diccionarios constituyó un punto de inflexión respecto a la influencia de la Lexicografía en la manera de concebir el desarrollo de glosarios y diccionarios especializados. En efecto, la matriz lexicográfica de la terminología llevaba a considerar casi exclusivamente organizaciones alfabéticas de términos, mientras que los árboles de dominio obligaron a volver la mirada sobre la presentación taxonómica de las nociones propias de un campo de conocimiento. Claro está que no como figura sustituta de las presentaciones alfabéticas de términos, sino como un cuadro sistemático necesario e insoslayable en el proceso de construcción de glosarios y diccionarios, más allá de que estos se organizaran finalmente bajo la tradicional ordenación alfabética.

En un proceso paralelo, la noción de árbol de dominio fue capturada también por la Informática con un significado peculiar, aunque emparentado; destacan a este respecto los trabajos pioneros de Missikoff, (1982) y Mokapetris (1983; 1987). También se utiliza la figura del árbol de dominio en áreas como la Biología (Pace, 2009). En Informática incluso se ha avanzado hacia la noción de bosque como un conjunto de árboles de dominio conectados mediante relaciones (por ejemplo, Belk and Brown, 2005)

Si nos ceñimos al campo de la Terminología, un árbol de dominio (o árbol de campo, esquema gráfico o campo de conceptos como también se menciona bajo distintos autores o traducciones) es una representación gráfica que contiene una estructura vertical de conceptos, y constituye un esquema básico de la composición interna de un área del saber (De Bessé, Nkwenti-Azeh & Sager, 1997). Por ello afirman Krieger y Finatto que el árbol de dominio *“es un diagrama jerárquico compuesto por términos clave de una especialidad, semejante a un organigrama”* (Krieger y Finatto, 2004, p. 134). En su expresión formal, el árbol de dominio es una variedad de taxonomía, pues solo establece relaciones jerárquicas entre sus elementos, siendo cada elemento un término.

Para crear un árbol de dominio se requiere previamente la identificación del dominio sobre el cual se realizará la investigación terminológica. Puede entenderse un dominio como un *“ámbito del saber o de la experiencia humana que se caracteriza por su especialización, o por la posibilidad de ser delimitado en función de las prácticas o los intereses propios de una comunidad de usuarios”* (Barité, et al, 2015, p. 64). Dicho de otro modo, un dominio -tanto en Terminología como en Organización del Conocimiento- puede ser cualquier recorte de la realidad que pueda ser estudiado desde una o varias

perspectivas disciplinares. Dentro de ese molde, un dominio tanto puede ser una ciencia, como una técnica, una disciplina social o humanística, una macrodisciplina, un campo temático interdisciplinario o un conjunto de estudios alrededor de un tópico. La expresión 'dominio' no es en absoluto ajena a la Organización del Conocimiento, sobre todo desde que a principios de los años noventa se introdujo la noción de análisis de dominio. Seguramente se debería hablar más en la literatura del área sobre el emparentamiento que tiene el concepto de dominio con lo que conocemos como análisis de dominio (Albrechtsen, 1993; Hjørland and Albrechtsen 1995).

El término que denomine al dominio será a la vez término tope y punto de partida de esa taxonomía.

A continuación, corresponde decidir si el árbol de dominio va a seguir un proceso de construcción del tipo *top-down* o del tipo *bottom-up*. En el primer caso, los diseñadores seleccionan los términos que encabezarán cada rama del árbol de acuerdo a criterios preestablecidos. Esta distribución por subcategorías del dominio “*puede seguir criterios más o menos canónicos (división de la Medicina en sus cuatro ramas tradicionales, Anatomía, Fisiología, Patología y Terapéutica), o puede realizarse a partir de criterios alternativos, por ejemplo, conforme a los objetivos que se persiguen o a las características de los destinatarios (como si se dividiera a la Medicina por síntomas)*” (Barité, 2017). Una vez definidas las subcategorías, se van agrupando los términos que se extraigan de las fuentes seleccionadas para la investigación (es decir, su corpus), bajo cada una de ellas.

Si se sigue la modalidad *bottom-up*, las expresiones que encabezarán las ramas del árbol solo se obtendrán luego de un análisis inductivo de los términos recogidos previamente, hasta una instancia que se considere suficiente o representativa del conjunto de la terminología del dominio. En este caso, entonces, los términos primariamente seleccionados se subagrupan considerando características y afinidades, y cada agrupamiento constituirá una rama del árbol con sus términos adscriptos.

Los términos que se consideren integrantes del dominio pueden situarse en al menos dos anillos: i) un anillo central con los términos que claramente pertenecen al dominio, llamados términos nucleares (*core terms*); así, por ejemplo, 'Jurisprudencia' es un término nuclear del Derecho); ii) un anillo de términos próximos o incluso pertenecientes a otros dominios, también denominados *fringe terms*, que suelen utilizarse frecuentemente en el dominio que se está estudiando (por ejemplo, *cannabis sativa* es un término de la Botánica que puede considerarse un término afín en los estudios sobre Adicciones).

La situación de un término dentro de una estructura conceptual puede variar según la perspectiva. Puede ser considerado un *core term* en un dominio y un *fringe term* en otro, e inclusive puede ser un *core term* si el dominio se estructura de una manera, y un *fringe term* si se organiza de otra manera alternativa.

Aunque su presentación final suele ser esquemática y despojada (un término tope, varias ramas similares a clases de sistemas de clasificación, y los términos propios de cada rama adscriptos a ella), los árboles de dominio cumplen con importantes objetivos y utilidades:

- contribuir decisivamente a la delimitación del dominio estudiado y sus fronteras conceptuales, “*al dar cuenta de sus partes y de las relaciones de ellas entre sí y con el conjunto*” (Auger y Rousseau 1978, p. 17);
- al encuadrar cada término en alguna de sus ramas, garantizar tanto la existencia del término como su pertenencia al dominio. Para facilitar este proceso suelen establecerse criterios de inclusión/exclusión para cada rama.

- operar como un elemento de control terminológico, permitiendo la visualización de lagunas, inconsistencias y desigualdades de la estructura de conceptos (Barité *et al.*, 2015, p. 31).
- favorecer la reconstrucción conceptual del dominio, al volver visible “*la estructura latente de un área del conocimiento, tal como es en un determinado punto de su evolución, o tal como puede ser vista desde una determinada perspectiva*” (Barité, 2017).

Hoy día constituye una pauta terminológica necesaria incluir el árbol de dominio en los preliminares de la publicación que resulte de la investigación, sea que se trate de una tesis (Cervantes, 2006), un diccionario (TermCat, 1992) o un tesaurus (Roggau, 2001). Su inclusión en la publicación constituye un elemento de visualización general del campo por parte de los lectores o usuarios, y favorece la evaluación crítica del trabajo.

La metodología habitual para construir un árbol de dominio es cualitativa, sin perjuicio de que, a partir de los datos numéricos que arroje la cantidad de términos ubicados en el mismo, puedan establecerse diversos análisis cuantitativos e incluso tomarse decisiones respecto a la inclusión o exclusión de términos.

Este trabajo se enfoca en el proceso de construcción del árbol de dominio que acompaña y controla la creación y desarrollo de un banco de datos terminológicos (BDT) especializado en Educación Superior. Se establecen los criterios terminológicos de formación del banco, se presenta la estructura del árbol con sus características particulares, y se indican las ramas creadas por inducción (enfoque *bottom-up*). Se brinda un avance de los datos cuantitativos de los términos, como resultado parcial de la investigación, y se consignan conclusiones en este punto de avance de la investigación.

METODOLOGÍA

El proyecto en el que se inserta la construcción del árbol se centra en la creación de un banco de datos terminológicos en Educación Superior (en adelante BDT-ES), sostenido por el programa TEMATRES, cuyo desarrollo y actualización permita la publicación de tesauros, diccionarios, taxonomías o listas de términos del dominio.

El subgrupo de investigación responsable de este proyecto integra el Grupo de Investigación sobre Terminología y Organización del Conocimiento (GTERM), radicado en la Facultad de Información y Comunicación (Universidad de la República, Uruguay). Dicho subgrupo se integra con tres bibliotecólogos (uno de ellos con formación informática), una traductora y un consultor en el área estudiada. También participan dos estudiantes de grado a través de pasantías curriculares.

El BDT-ES recopila la terminología de la Educación Superior aceptada y utilizada en los cinco países del Cono Sur americano (Argentina, Brasil, Chile, Paraguay y Uruguay). Los datos que se indican en cada registro terminológico son los siguientes:

- término de entrada en español,
- término equivalente en inglés,
- término equivalente en portugués (Brasil),
- rama del árbol en la que encarta el término,

- términos sinónimos o cuasi sinónimos,
- términos relacionados,
- definiciones extraídas del corpus,
- definición original,
- otros datos.

El corpus del proyecto se integra con 150 fuentes de información, ponderadas de manera diferente conforme al valor ficto que se asigna a cada una:

- Fuentes nivel I: normativas y documentos oficiales nacionales y federales sobre Educación Superior de los países del Cono Sur americano, ya sean nacionales o del Mercosur (tratados, leyes, decretos, reglamentos, protocolos, ordenanzas, resoluciones, convocatorias, bases, convenios). Valor ficto de estas fuentes: 4.
- Fuentes nivel II: diccionarios y glosarios sobre Educación Superior y Educación en general, documentos técnicos, informes y programas de eventos. Valor ficto de estas fuentes: 2.
- Fuentes nivel III: cualquier otra obra o documentación que aporte información fiable de interés para la investigación, de los países de la región o extrarregionales (información periodística confiable, entrevistas a autoridades de la educación, etc.). Valor ficto de estas fuentes: 1.

Hasta la fecha han sido estudiadas 65 de las 150 fuentes del corpus, ninguna de las cuales corresponde a Brasil hasta el momento (salvo en lo relativo a la terminología de los posgrados), por lo que la información que se brinda en este avance es todavía parcial.

Se estableció como criterio general de selección que los términos sean propios del dominio de la Educación Superior (los llamados nucleares, como por ejemplo 'Doctorado'), o que sean utilizados en forma habitual en el ámbito de la Educación Superior aunque pertenezcan al área Educación en general (los llamados afines, como por ejemplo 'Orientador') o a otras áreas (por ejemplo, 'Plagio'). Si un concepto se expresa con diferencias, matices o variantes en los diferentes países del Cono Sur americano, se presenta esa variedad de forma fidedigna. Los casos dudosos se incluyen para una valoración posterior.

Los términos seleccionados deben aparecer por lo menos en tres fuentes distintas o en fuentes de dos países diferentes, pero una de ellas siempre deberá ser una fuente nivel I. Hasta el presente se han incorporado al BDT-ES 510 términos, en base a estos criterios.

Llegados a este punto, se construyó un árbol de dominio para situar los 510 términos seleccionados. La metodología utilizada fue básicamente inductiva, a partir del significado explícito o sobreentendido de los términos y de la identificación de sus asociaciones semánticas o funcionales. Se consideró apropiado el establecimiento de una rama con un mínimo de treinta términos adscriptos.

RESULTADOS

La compilación de términos en las fuentes calificadas del corpus demostró que con casi medio millar de términos es posible contener -por lo menos en los países del Cono Sur americano- los principales conceptos que pueden asociarse al dominio de la Educación Superior, los cuales se encuentran en sus

documentos oficiales o de referencia. Este universo de elementos terminológicos debería ser suficiente, además, para clasificar o indizar la mayor parte de sus documentos. Por ese motivo, su inclusión en un banco de datos terminológicos que pueda tomarse como base para la construcción de tesauros, diccionarios, glosarios o listas de descriptores sobre Educación Superior, parece totalmente pertinente.

Para el procedimiento de agrupación en ramas de los 510 términos recolectados no se consultó ninguna clasificación previa del dominio de Educación Superior, establecida en un tesoro o en un sistema de clasificación. Por el contrario, se prefirió considerar las asociaciones paradigmáticas -o de otra índole- que se encontraron entre los términos en los documentos del corpus, para enfatizar la modalidad *bottom-up* de creación del árbol.

Luego de sucesivas aproximaciones y ajustes, se identificaron ocho ramas principales del dominio: Actores, Comunicación científica, Currícula, Documentos, Estructura universitaria, Evaluación, Procesos educativos universitarios y Titulaciones. Las ocho ramas logran dar cobertura a los 510 términos en su totalidad, por lo que el BDT-ES no cuenta con términos huérfanos.

Dos de las ramas fueron subdivididas a su vez, en forma provisional: Comunicación científica en Arbitraje y Eventos; Procesos educativos en Evaluación universitaria y Otros procesos. Se está evaluando la posibilidad de subdividir la rama Currícula en Grado y Posgrado. A cada una de las ramas se le asignaron criterios de inclusión/exclusión de términos. A modo de ejemplo se mencionan los criterios de inclusión /exclusión de la rama Documentos, y el listado de términos adscriptos:

DOCUMENTOS

Incluye términos que representan diferentes tipos documentales propios del ámbito de la educación superior, o que se concretan como consecuencia del quehacer universitario habitual.

Lista primaria de 55 términos adscriptos, que incluye sinónimos no preferidos (entre paréntesis), y términos en estudio (entre signos de interrogación): Acta, Anales, Artículo científico (Paper), Artículo de divulgación, Bibliografía, Boletín, Certificado de estudios, Comunicación, Convenio, Convenio específico, Convenio marco, Currículum, Diccionario, Diploma, Documento de trabajo, Documento técnico, Ficha de lectura, Glosario, Guía de estudio, ¿Informe de lectura?, Informe técnico, Manual, Material didáctico, Memorándum, Memoria, Monografía (Trabajo monográfico), Obra colectiva, Parecer, ¿Perfil de investigación?, ¿Perfil de proyecto?, Pericia, Plan de Estudios, Ponencia, Póster, Pre-print, Preproyecto, Protocolo, Proyecto, Proyecto de investigación, Publicación periódica, Reglamento, Reseña, Resumen (Abstract), Revista, Tesina, Tesis, Tesis de doctorado, Tesis de licenciatura, Tesis de maestría, Título, Trabajo de fin de curso, Tutorial.

En la tabla 1 se exhibe la distribución de los términos en las ramas del árbol.

Tabla 1. Árbol de dominio sobre Educación Superior

Rama del árbol	Cantidad de términos	Observaciones
Actores	101	
Comunicación científica	48	Se integra con dos subramas: Arbitraje y Eventos
Currícula	153	A estudio subdivisión en grado y posgrado
Documentos	55	
Estructura universitaria	53	
Procesos	70	Se integra con dos ramas: Evaluación universitaria y Otros procesos
Titulaciones	30	
Total términos	510	

CONCLUSIONES

Existe una intersección teórico-metodológica entre Terminología y Organización del Conocimiento que, en virtud de tratarse de campos temáticos relativamente jóvenes, no ha sido explorada ni en su extensión ni en su potencial. Uno de los aportes que ofrece la Terminología a la Organización del Conocimiento es la creación de árboles de dominio como herramientas estructuradoras de terminología, como una fase metodológica necesaria en la creación de sistemas de organización del conocimiento. Los árboles de dominio, a partir de su finalidad sistematizadora pueden ordenar los procedimientos previos al control de vocabulario, y facilitar en especial el establecimiento de relaciones jerárquicas entre términos.

En este trabajo se explica el sentido y la utilidad de la creación de un árbol de dominio, en el marco de un proyecto de desarrollo de un banco de datos terminológicos sobre Educación Superior, basado en la terminología habitual de los cinco países del Cono Sur americano. De las 150 fuentes del corpus, se analizaron 65 hasta hoy. De estas 65 fuentes se extrajeron hasta el presente 510 términos. Este caudal fue organizado conforme a procedimientos *bottom-up*, en ocho grandes subclases o subramas.

El árbol creado en el marco de la investigación gobernará la inclusión/exclusión de los términos que surjan del corpus todavía no analizado, y adjudicará las ramas de referencia para ellos.

La claridad del método y sus resultados, aún cuando constituyen un avance de la investigación, confirman el valor de los árboles de dominio como herramientas de encuadre, relación y control terminológico. Aplicados en forma rigurosa los árboles de dominio permiten controlar las fronteras del dominio que se está estudiando, tomar decisiones respecto a la orientación, la dimensión y el alcance del dominio, evitar lagunas y fugas de términos importantes, y anticipar -en el caso de los tesauros y las taxonomías- los vínculos jerárquicos entre los candidatos a descriptores.

El proyecto de investigación en curso permite avanzar la idea de que no existen mayores obstáculos para elaborar un árbol de dominio desde una perspectiva *bottom-up*, y que quizás sea ese el enfoque más adecuado, en virtud de que parte de un respaldo en la garantía literaria, es decir, en la justificación de descriptores en base a documentos reales, que son los que en definitiva serán clasificados e indizados.

Al servir, por tanto, para fines de organización y sistematización de terminología, los árboles de dominio podrían integrarse a los métodos de construcción de tesauros, taxonomías y ontologías, y ser considerados, en consecuencia, por los responsables de normas y estándares.

REFERENCIAS BIBLIOGRÁFICAS

Albrechtsen, H. (1993). Subject analysis and indexing: from automated indexing to domain analysis. *The Indexer*, 18(4): 221-224.

Alexiev, B. & Marksby, N. (2010). Terminology as organized knowledge. *Advances in Knowledge Organization*, 12, 363-370.

Auger, P. y Rousseau, L.-J. (1978). *Méthodologie de la recherche terminologique*. Québec: Office de la langue française.

Barité, M., Colombo, S., Duarte Blanco, A., Simón, L., Cabrera Castromán, G., Odella, M. L., Vergara, M. (2015). *Diccionario de Organización del Conocimiento: Clasificación, Indización, Terminología*. (6ª edición ampliada y revisada). Montevideo: CSIC.

Barité, M. (2017). Los árboles de dominio. In Álvarez Catalá, S. y Barité, M. (orgs.), *Teoría y praxis en Terminología*. Montevideo: CSIC. En imprenta.

Belk, J. M. And K. S. Brown (2005). Forest diagrams for elements of Thompson's Group F. *International Journal of Algebra and Computation* 15(5/6): 815-850.

British Standards Institution. (2005). *Structured vocabularies for information retrieval: BS 8723-1*. S.I.: Author.

Cabré, M. T. (1993). *La Terminología: Teoría, metodología, aplicaciones*. Barcelona: Antártida.

Campos, M.L. de A. (2001). *Linguagem documentária: Teorias que fundamentam sua elaboração*. Niteroi: EdUFF.

Cervantes, B. M. N. (2006). *Terminologia do processo de Inteligência Competitiva: Estudo teórico e metodológico*. Londrina: EDUEL.

Dahlberg, I. (1992). Knowledge organization and terminology: Philosophical and linguistic bases. *International Classification*, 19(2), 65-71.

De Bessé, B., Nkwenti-Azeh, B. & Sager, J.C. (1997). Glossary of terms used in Terminology. *Terminology*, 4(1), 119-156.

Fedor de Diego, A. (1995). *Terminología: teoría y práctica*. Caracas: Universidad Simón Bolívar.

Hjørland, B., Albrechtsen, H. (1995). Toward a new horizon in information science: domain-analysis. *Journal of the American Society for Information Science*, 46(6), 400-425.

International Organization for Standardization = ISO. (1987). *Norme internationale: Principes et méthodes de la terminologie: ISO 704*. Geneva: Author.

International Organization for Standardization = ISO. (2011). *Thesauri and interoperability with other vocabularies: Part 1: Thesauri for information retrieval*. Geneva: Author.

Krieger, M.G. e Finatto, M.J.B. (2004). *Introdução à Terminologia: Teoria e prática*. São Paulo: Contexto.

Missikoff, M. (1982). A domain based internal schema for relational database machines. En ACM SIGMOD '82 International Conference on Management of Data. *Proceedings*, 215-224.

Mokapetris, P. (1983). Domain names: implementation and specification. ISI. Recuperado de <https://buildbot.tools.ietf.org/html/rfc883>

Mokapetris, P. (1987). *Domain names: concepts and facilities*. ISI. Recuperado de <https://tools.ietf.org/html/rfc1034>

Pace, N. R. (2009). Mapping the Tree of Life: Progress and Prospects. *Microbiology and Molecular Biology Reviews*, 73(4): 565-576.

Roggau, Z. (2001). *TELL : Tesouro de Lengua y Literatura*. Rosario: Nuevo Parhadigma.

San Segundo, R. (1996). *Sistemas de organización del conocimiento: La organización del conocimiento en las bibliotecas españolas*. Madrid: Universidad Carlos III ; Boletín Oficial del Estado.

TermCat (1992). *Diccionari de Lingüística*. Barcelona: Fundació Barcelona.

Woźniak-Kasperek, J. (2014). Terminology as a picture of Knowledge Organization in a scientific discipline. *Advances in Knowledge Organization*, 14, 305-311.



ANÁLISE DO CONCEITO DE “IDENTIFICAÇÃO ARQUIVÍSTICA” À LUZ DOS PRINCÍPIOS DA TEORIA DO CONCEITO DE DAHLBERG

Danilo André Bueno¹, Ana Célia Rodrigues², Maria Luiza de Almeida Campos³

¹Programa de Pós-graduação em Ciência da Informação, Universidade Federal Fluminense
danilobuenoarq@gmail.com

²Programa de Pós-graduação em Ciência da Informação, Universidade Federal Fluminense
anyrodrigues@yahoo.com.br

³Programa de Pós-graduação em Ciência da Informação, Universidade Federal Fluminense
marialuizalmeida@gmail.com

RESUMO Estudo dos aspectos teóricos que envolvem a definição do conceito de identificação arquivística na literatura espanhola e brasileira, refletindo sobre a necessidade de debate teórico mais aprofundado sobre esse conceito, visto que se encontram dispersos e pouco sistematizados. Apresenta e analisa o conceito de “identificação arquivística” à luz da Teoria do Conceito de Ingetraut Dahlberg (1978), identificando suas características constitutivas essenciais (que lhe dão identidade) e suas características constitutivas acidentais (funções, aplicações, contexto), a partir da aplicação do método triangular conceitual. Como resultado deste estudo, considera-se a identificação arquivística como uma metodologia de pesquisa, desenvolvida com rigor científico, a partir de procedimentos e instrumentos próprios, que subsidiam o arquivista no desenvolvimento das funções de classificação, avaliação, descrição e também o planejamento da produção dos documentos de arquivos. Nesta perspectiva a identificação arquivística tem por finalidade o reconhecimento do sujeito produtor (elemento orgânico e funcional) e do objeto produzido (tipos documentais), conhecimento essencial para o estudo da gênese do documento de arquivo, para fins de organização de documentos acumulados em arquivos ou para o planejamento de programas de gestão de documentos.

PALAVRAS-CHAVE *Identificação arquivística, Teoria do Conceito, Tipologia Documental.*

ABSTRACT Study of the theoretical aspects that involve the definition of the concept of archival identification in the Spanish and Brazilian literature, reflecting on the need for a more detailed theoretical debate on this concept, since they are scattered and little systematized. It presents and analyzes the concept of "archival identification" in the light of Ingetraut Dahlberg's Theory of Concept (1978), identifying its essential constitutive characteristics (which give it identity) and its accidental constitutive characteristics (functions, applications, context) from the Application of the conceptual triangular method. As a result of this study, archival identification is considered as a research methodology, developed with scientific rigor, from its own procedures and instruments, which subsidize the archivist in the development of classification, evaluation, description and also production planning functions Of the file documents. In this perspective, archival identification has the purpose of recognizing the producer subject (organic and functional element) and the object produced (documentary types), essential knowledge for the study of the genesis of the archival document, for the purpose of organizing documents accumulated in archives or For the planning of document management programs.

KEYWORDS *Archival identification, Concept Theory, Documentary Typology.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Na literatura da área verifica-se que os métodos e teorias arquivísticas, até pouco tempo estavam voltados ao tratamento das massas documentais que se acumulavam sem critérios técnicos ao longo dos séculos. Apenas recentemente, na segunda metade do século XX, passou-se a perceber as mudanças que envolvem a gestão, o acesso e o uso dos documentos/informação arquivística.

Nesse cenário, a Arquivística tem apresentado elementos alinhados à contemporaneidade a partir de uma visão global e integrada dos arquivos, especialmente a partir do desenvolvimento da teoria do ciclo de vida dos documentos. Essa evolução vai ao encontro da constante valorização da informação para os processos de gestão organizacional e para a tomada de decisão nesse âmbito.

O termo e o conceito de identificação surgiram na Espanha para designar as pesquisas desenvolvidas por grupos de arquivistas preocupados com a formulação de metodologias para solucionar o problema da acumulação irregular de documentos, sendo utilizada inicialmente como ferramenta para a organização de documentos acumulados em arquivos históricos, e posteriormente, para o planejamento de técnicas e procedimentos de Gestão de Documentos a partir da década de 1990. (Rodrigues, 2008, p. 181).

Nesse contexto, a construção teórica desta metodologia ganha destaque nos estudos arquivísticos nas três últimas décadas, na qual a *Identificação* foi reconhecida como uma fase independente da metodologia arquivística nas *Primeras Jornadas sobre Metodología para la Identificación de Fondos Documentales de las Administraciones Públicas*, realizadas em 1991, em Madrid na Espanha. Nesta jornada o conceito de identificação foi definido como um conjunto de procedimentos de pesquisa para a “[...] identificação do órgão produtor, identificação do elemento funcional e identificação do tipo documental.” (Martín-Palomino Benito; Torre Merino, 2000, p. 14, tradução nossa).

A partir deste momento o uso do termo e o conceito se consolidaram no meio arquivístico espanhol, sendo incorporados pelo Dicionário de Terminologia Arquivística daquele país que a define como “[...] fase do tratamento arquivístico que consiste na investigação e sistematização das categorias administrativas e arquivísticas em que se sustenta a estrutura de um fundo” (DTAE, 1993, p. 37, tradução nossa).

Ainda nos anos 1980, o Brasil integrou o movimento internacional que se formava em busca de referenciais metodológicos para resolver a superlotação dos arquivos, problema comum do qual compartilhavam os países ibero-americanos, desenvolvendo processos de identificação no Arquivo Nacional. Destas experiências profissionais de identificação de fundos, resultou a definição posta no Dicionário Brasileiro de Terminologia Arquivística (DIBRATE, 2005, p. 104), que também considera a *identificação* como uma fase do processamento técnico dos arquivos, definindo-a como o “processo de reconhecimento, sistematização e registro de informações sobre arquivos, com vistas ao seu controle físico e/ou intelectual”.

Rodrigues (2008, p. 13), afirma que:

Os aspectos teóricos que envolvem a metodologia da identificação e, especificamente, da identificação de documentos nos parâmetros da tipologia documental, estão pulverizados na literatura arquivística, refletindo a necessidade de um esforço de sistematização de ideias para abordar a construção teórica desta metodologia.

Diante dessa afirmativa, o objetivo deste artigo é apresentar e analisar o conceito de Identificação à luz da Teoria do Conceito de Ingetraut Dahlberg, que propõe um modelo analítico idealizado para elucidar a natureza e a estrutura dos conceitos, modelo que implica em analisar o conceito voltado para o seu referente, que pode ser “[...] um simples objeto, um conjunto de objetos considerados como uma unidade ou uma propriedade, uma ação, uma dimensão, etc. ou qualquer destas combinações” (Dahlberg, 1978, p.5).

O modelo analítico proposto por Dahlberg (1978) orienta a construção e estruturação de conceitos (Dahlberg, 1981, 1983) a partir do método “triangular conceitual”, com vistas a verificar as características constitutivas essenciais e acidentais dos conceitos a partir de predicções ou afirmações verdadeiras feitas sobre um determinado item do referente, o que implica em analisar e determinar essas características e suas relações, que formam os conteúdos dos conceitos.

O MÉTODO ANALÍTICO DA TEORIA DO CONCEITO

A Teoria do Conceito é proposta como um método analítico idealizado para tornar clara a natureza e a estrutura dos conceitos, sendo aplicado para sustentar os procedimentos apropriados para seu estudo, desenvolvido por Ingetraut Dahlberg na década de 1970, no âmbito do desenvolvimento de terminologias no campo das Ciências Sociais.

Conforme explica Campos (2001, p. 87), “a Teoria do Conceito possibilitou uma base mais sólida para a determinação e o entendimento do que consideramos conceito, para fins de representação/recuperação da informação”.

Conceito é definido como “uma unidade do conhecimento, compreendendo afirmações verdadeiras sobre um dado item de referência, representado numa forma verbal”. (Dahlberg, 1978, p. 7). Conceitos são, portanto, entendidos aqui nesta reflexão como ‘unidade do conhecimento’, e deste modo Dahlberg (1978), destaca que se o homem tem a capacidade de desenvolver afirmações corretas sobre as coisas reais (itens empíricos) e sobre as ideias da mente (cognição), então é possível construir um modelo de análise para a construção de conceitos.

O modelo de análise conceitual parte da seleção de um item de referência – referente - (que pode ser um simples objeto, ou uma propriedade, uma ação, uma dimensão, etc.). As afirmações sobre um referente são aceitas como verdadeiras numa forma verbal e pode ser convenientemente representada como um termo. Com a forma verbal os seres humanos são capazes de se comunicar oralmente ou por escrito sobre os conteúdos (afirmações sobre o referente) de um conceito, e o uso dessa forma verbal é aplicada para pensar no universo do discurso, na construção e representação do conhecimento.

Dahlberg (1978, p. 7) define três tipos de componentes necessários para a formação de uma unidade do conhecimento:

[...] afirmação verdadeira é a componente de um conceito que expressa um atributo do seu item de referência;

[...] item de referência é o componente do conceito para o qual sua afirmação verdadeira e sua forma verbal estão diretamente relacionadas, seu “referente”;

[...] forma verbal (termo/nome) de um conceito é o componente que resume convenientemente ou sintetiza e representa um conceito com o propósito de designar um termo, um conceito em comunicação.

Sendo assim, Dahlberg (1978) propõe três passos fundamentais para a construção de conceitos representados por:

- Passo referencial (A) – item de referência – referente.
- Passo predicacional (B) – afirmações sobre o referente – características.
- Passo representacional (C) – forma verbal – termo que representa o conceito.

Esses passos também são representados graficamente dentro do triângulo conceitual proposto por Dahlberg segundo o modelo reproduzido abaixo:

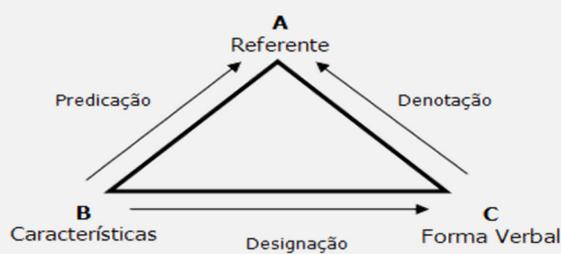


Figura 1. Triângulo Conceitual

Fonte: Dahlberg, 1978.

Este modelo evidencia os tipos de atividades/relações existentes entre os três componentes do modelo: Predição nominal (B-A); Designação (B-C); e Denotação (C-A). O esquema proposto por Dahlberg (1978) parece ser mais válido dentro das Ciências Sociais baseado nas linhas de categorias formais de Aristóteles, através do qual todos os itens de uma referência devem ser fundamentalmente ordenados em quatro categorias formais e doze subcategorias formais, começando-se pela categorização de um referente pertencente a algumas combinações delas.

Categorias formais	Subcategorias formais
Entidades	> princípios
	> objetos imateriais
	> objetos materiais
Propriedades	> quantidades
	> qualidades
	> relações
Atividades	> operações
	> estados
	> processos
Dimensões	> tempo
	> posição
	> espaço

Figura 2. Categorias formais e Subcategorias formais de Aristóteles

Fonte: Dahlberg, 1978.

Verifica-se, portanto, que a determinação final dos referentes, que segue a regra de predicação, pode tomar a forma de uma “cadeia de características” ou “cadeia de conceitos”, pois segundo Dahlberg (1978, p. 12) “Começando de qualquer base e levando os predicados até uma última subcategoria de forma, pode-se estabelecer a classe que forma uma das características de um conceito”.

De maneira analítica, podemos estabelecer que todos os nossos conceitos são abstrações da realidade no sentido de que são produtos e instrumentos da habilidade do homem de pensar e falar sobre a realidade na medida permitida por seu conhecimento da realidade. Eles diferem, no entanto, na escala de abstração, passando dos mais específicos e individuais para os específicos e mais gerais. (Dahlberg, 1978, p. 12).

Corroborando com essa análise, Campos (2001, p. 103) ressalta que “Para Dahlberg, o processo de determinação do conceito se dá no momento em que é selecionado um item de referência – um referente – e analisado dentro de um universo”. Assim, é nesse momento em que se atribui predicado ao referente, selecionando suas características ou atributos mais relevantes para a definição do conceito, que é a soma de todos os elementos que o compõem, ou seja, a soma das características que refletem as propriedades de um item de referência é que formarão o conteúdo e a designação apropriada para o mesmo.

Segundo Dahlberg (1978, p. 13) “[...] podemos reconhecer certos tipos de características de acordo com as afirmações – e derivável delas – sobre as propriedades de um referente, além daquelas que constituem a categoria final de um conceito, a saber, a sua especialização”.

Existem três tipos gerais de predicacões que produzem três tipos de características que podem ser afirmadas segundo Dahlberg (1978):

- 1) Características que se aplicam a todos os referentes de um dado tipo: “características essenciais”;
- 2) Características que se aplicam a somente alguns referentes de um dado tipo: “características acidentais”;
- 3) Características que se aplicam a um único referente: “características individualizantes”.

Dahlberg (1978, p.14) também faz distinção entre as características que constituem um conceito, características constitutivas dos conceitos (características necessárias = essenciais); opondo-se as características possíveis, que são qualquer uma daquelas que um dado tipo de referente pode adquirir, mas que não são necessárias para a constituição de um conceito; características consecutivas ou implícitas resultam daquelas constitutivas dos conceitos por implicação.

Por último, Dahlberg (1978, p. 15) faz distinção entre quatro grandes tipos de características, a saber:

- a) Características categoriais formais: aquelas que se referem a categorias formais e subcategorias formais;
- b) Características ontológicas: aquelas que se referem a categorias do ser, como, por exemplo, “ser um objeto material”, “ser um ser vivente”, “ser um ser humano”, “ser um suporte de informação”.
- c) Características constitutivas dos conceitos (características essenciais, acidentais e individualizantes); aquelas que se referem a quase todos, ou a alguns casos ou a um simples caso de um dado item de referência;
- d) Características consecutivas e implícitas: aquelas que resultam das características essenciais e acidentais por implicação ou possível diversidade.

A partir dessa reflexão, podemos verificar que para o processo de construção e análise de qualquer conceito é necessário que selecionemos um item de referência – referente – para analisá-lo dentro de um universo – contexto – atribuindo-lhe predicados – afirmações – a dito referente selecionando as suas características mais relevantes (essenciais, acidentais e individualizantes) que constituirão o conceito, bem como formarão seu conteúdo e designação – termo.

O CONCEITO DE IDENTIFICAÇÃO ARQUIVÍSTICA

O termo e o conceito de “identificação” surgem após a introdução do conceito de Gestão de Documentos no domínio de conhecimentos da Arquivística, trazendo consigo a necessidade de formulação de novas abordagens teórico-metodológicas para resolver os problemas decorrentes da crescente produção e acumulação de documentos que crescia exponencialmente nas administrações públicas e privadas verificadas a partir da década de 1940. (Rodrigues, 2008, p. 48).

Com vistas à resolução desses problemas, foram criados na década de 1980 grupos de trabalho em países ibero-americanos. Entre esses países que participaram da formulação do termo e do conceito de identificação estão Espanha, Brasil, Portugal e México, nos quais, Mendo Carmona (2004, p.41, tradução nossa), aponta que:

Tal definição foi sendo desenvolvida em reuniões profissionais para discutirem a realidade constante nos arquivos desses países com larga tradição administrativa e arquivística, cujos procedimentos administrativos e práticas diárias não tinham se preocupado com a transferência sistemática de seus fundos documentais, acumulando fundos indiscriminadamente nos depósitos de arquivos.

A partir desse momento, muitas experiências empíricas foram aplicadas nos arquivos para o desenvolvimento de uma metodologia que possibilitasse a resolução do problema do acúmulo indiscriminado de documentos. A Espanha, nesse sentido, se destacou no panorama arquivístico internacional com experiências desenvolvidas pela Direção de Arquivos Estatais do Ministério da Cultura, no âmbito de seus sistemas de arquivos. Rodrigues (2008, p.48) destaca que as discussões na Espanha sobre a normalização dos processos de identificação de documentos de arquivo para o controle da acumulação ou planejamento da gestão de documentos influenciaram toda a Arquivística ibero-americana, e nesse contexto, especialmente o Brasil.

O problema comum compartilhado pela Arquivística ibero-americana era a acumulação de fundos. A metodologia proposta visava recuperar a fluidez do sistema de arquivos, focando a questão sob dois pontos de vista: controlar as séries documentais desde o momento de sua produção e àquelas acumuladas nos arquivos, definindo um limite cronológico como ponto de partida, denominada identificação e avaliação de fundos documentais. (Rodrigues, 2008, p.50).

A identificação é um conceito recente na literatura Arquivística internacional que vem sendo introduzido gradativamente na Arquivística brasileira a partir da década de 1990. Mendo Carmona (2004, p. 41, tradução nossa), referindo-se ao conceito de identificação obtido na Espanha nas *Primeras Jornadas sobre Metodología para la Identificación de Fondos Documentales de las Administraciones Públicas*, realizadas em 1991, em Madrid na Espanha, define identificação como: “[...] o processo de investigação e sistematização de categorias administrativas e arquivísticas nas quais se sustenta a estrutura de um fundo, sendo um dos seus objetivos principais assegurar através de seus resultados a valorização das séries documentais”.

Martín-Palomino Benito e Torre Merino (2000), definem a Identificação Arquivística como “[...] fase do tratamento arquivístico que consiste na investigação e sistematização das categorias administrativas e arquivísticas em que se sustenta a estrutura de um fundo.” (Martín-Palomino Benito; Torre Merino, 2000, p. 14, tradução nossa).

Corroborando com a afirmativa, Mendo Carmona (2004, p. 42, tradução nossa), ressalta que a identificação é a melhor ferramenta para aplicar o princípio básico da Arquivística: o de respeito à proveniência e da estrutura interna do fundo, que:

Consiste na investigação das características dos elementos implicados na gênese do fundo: o sujeito produtor e o objeto produzido. Entende-se por sujeito produtor a pessoa física, família ou organismo que produziu e/ou acumulou o fundo. Entende-se por objeto produzido a totalidade do fundo e cada uma das agrupações documentais que o compõem [série documental].

A identificação arquivística, portanto, é imprescindível para todo o processo de organização de arquivos, e principalmente para o planejamento da gestão de documentos, conforme elucida Vieira (2007, p.16-17):

[...] a identificação do órgão produtor, de seu elemento funcional e do tipo documental é realizada mediante pesquisa nos documentos legais de sua constituição. Tais informações coletadas nesta etapa de identificação são registradas em diferentes instrumentos tais como: índice do órgão produtor, índices de tipos documentais, organograma, formulários de levantamentos de dados e outros, que posteriormente serão usados na avaliação, organização, descrição dos documentos de arquivo.

A fase de identificação para a Arquivística, conforme elucida Vieira (2007), consiste na investigação e na análise crítica dos elementos da gênese do fundo e do reconhecimento do contexto administrativo que gerou a documentação em qualquer organização. É a pesquisa sobre os elementos que caracterizam o contexto de produção (estrutura organizacional formal – elemento orgânico - e suas competências, funções, atividades e procedimentos administrativos – elemento funcional) e a tipologia documental (subprodutos do desenvolvimento dos procedimentos administrativos e jurídicos).

Evidenciando sua importância, Rodrigues (2005, p. 05), sustenta que:

A base metodológica que sustenta a identificação é a aplicação de maneira direta do princípio da proveniência, que consiste em não misturar os documentos gerados por diferentes órgãos produtores e, conseqüentemente, o da ordem original, que significa não romper a ordem natural de produção de tipos documentais oriundos dos procedimentos administrativos que o determinam.

Conhecer a estrutura, o funcionamento e as tipologias documentais é fundamental para o desenvolvimento de qualquer processo ou tarefa arquivística no âmbito da criação/produção, classificação, avaliação e descrição de documentos, em suma, esta metodologia subsidia o desenvolvimento de parâmetros para a gestão de documentos.

Sendo assim, a partir dessas definições tentar-se-á analisar através do método triangular conceitual a “Identificação Arquivística” à luz da Teoria do Conceito voltada para o referente e suas características.

ANÁLISE DO CONCEITO DE IDENTIFICAÇÃO ARQUIVÍSTICA

Dentre as definições sobre “Identificação” apontadas na literatura, nosso referente pode ser categorizado como um “Método”. Este método possui operações, estados e processos, e observa-se que nas definições analisadas ele é tomado ora como um processo, ora como uma atividade ou como uma metodologia, e o seu contexto de análise (universo) é o único que permanece estável, ou seja, no contexto arquivístico. Portanto, parece-nos mais conveniente optar em denominar nosso referente como um “método”, visto que um método é um modo de proceder a partir do desenvolvimento de uma atividade específica que possui operações, estados, processos específicos de uma determinada natureza, e no contexto arquivístico ele é qualitativo e investigativo.

Nesse sentido, como fora definido que o referente (que é um método) e a forma verbal (Identificação Arquivística) nos resta apontar e verificar as predicções sobre esse referente, ou seja, procedeu-se ao estabelecimento das características que constituem o conceito (características essenciais, acidentais e individuais) a partir da verificação de afirmações evidenciadas e analisadas na literatura especializada.

No entanto, para alcançar o objetivo proposto por este artigo, optou-se por verificar apenas as características constitutivas essenciais e acidentais do conceito de “Identificação Arquivística”, com vistas à tentativa de desenvolver uma análise reflexiva que contribua para a área de Arquivologia

Dessa forma, podemos apontar que as características essenciais propostas por Dahlberg (1978) são aquelas oriundas de afirmações sobre a identidade de um determinado referente, sobre o que ele é e como se constitui, ou seja, as características essenciais dos conceitos são aquelas que lhes dão identidade, que os constituem essencialmente. Por exemplo, nesta análise o que lhe dá identidade à Identificação arquivística é a aplicação direta do princípio da Proveniência na investigação da gênese do fundo de arquivo, os elementos que caracterizam o seu contexto de produção (estrutura organizacional e funcional) e o objeto produzido (tipos documentais) gerado ou acumulado.

Da análise realizada na literatura, temos que as características essenciais do método de Identificação arquivística, e que a sustenta, é a aplicação direta do princípio da Proveniência no reconhecimento da gênese do fundo, identificando as estruturas, competências, funções e atividades do contexto de produção dos tipos documentais, para a formação, classificação e avaliação das séries documentais.

Por outro lado, as características acidentais, ou seja, variáveis deste método podem ser verificadas a partir das possibilidades de sua aplicação na formação e/ou organização de um fundo de arquivo: tanto na fase de acumulação, quanto na fase de produção documental em organizações públicas ou privadas. Além disso, verifica-se que o método possui a função de subsidiar de maneira segura o desenvolvimento de programas de Gestão de Documentos, criando parâmetros para desenvolvimento das funções arquivísticas de planejamento da produção, classificação, avaliação e descrição de documentos.

Verifica-se a partir da aplicação da teoria do conceito de Dahlberg (1978), que a soma das características essenciais e acidentais, constituintes da Identificação Arquivística, lhe confere identidade como um método analítico de investigação da gênese do fundo de arquivo dentro das metodologias arquivísticas. Dessa forma, tem-se que o método de análise proposto pela Teoria do Conceito nos possibilita um raciocínio mais seguro para desenvolver ou analisar a natureza e estrutura dos conceitos, a partir da literatura, pois as características constitutivas dos conceitos formam o seu conteúdo (Campos, 2010).

A temática deste artigo, portanto, objetivou a exploração da literatura para obter a compreensão das características constituintes do conceito de Identificação Arquivística, por meio do método triangular conceitual proposto por Dahlberg (1978), explorando os principais conceitos relacionados ao tema, bem como o processo e a importância de sua aplicação nas atividades inerentes ao fazer arquivístico.

A Identificação arquivística, como método analítico de investigação da gênese do documento de arquivo, cria bases sólidas e seguras para analisar a acumulação desordenada de documentos com vistas à sua correta organização, bem como cria conhecimento para o planejamento e controle da produção de documentos, classificação de documentos e avaliação de documentos, no âmbito da Gestão de Documentos.

CONCLUSÕES

A partir da análise realizada com base na aplicação do método analítico da Teoria do Conceito, podemos verificar a possibilidade de formular um raciocínio mais seguro no âmbito do desenvolvimento de análises e compreensão da estrutura dos conceitos, a partir do triângulo conceitual proposto por Dahlberg (1978), revelando-se um método seguro para analisar conceitos no campo científico.

Sob esse prisma é possível conceituar a Identificação Arquivística a partir do método triangular conceitual. A par das discussões sobre suas características essenciais e acidentais constitutivas compiladas da literatura especializada, parece-nos possível considerar que a Identificação arquivística é um método analítico que consiste na investigação das características dos elementos implicados na gênese do fundo, a partir da aplicação direta do princípio da proveniência, cuja finalidade é identificar o sujeito produtor (elemento orgânico: estrutura formal – elemento funcional: competências, funções, atividades e procedimentos administrativos) de qualquer instituição, seja pública ou privada; e do objeto produzido (tipos documentais). Esse método, se bem desenvolvido, possibilita a intervenção do arquivista de forma segura em todas as fases do ciclo vital dos documentos, subsidiando o planejamento de programas de Gestão de Documentos.

A Identificação Arquivística, portanto, consiste na investigação e na análise crítica dos elementos da gênese do fundo e do reconhecimento do contexto administrativo que gerou a documentação em qualquer organização, seja pública ou privada. É a pesquisa que evidencia os elementos que caracterizam o contexto de produção (estrutura organizacional formal, competências, funções, atividades e procedimentos administrativos) e a tipologia documental (subprodutos do desenvolvimento de atividades e dos procedimentos administrativos).

A Arquivologia, historicamente, tem buscado afirmar seu estatuto científico e epistemológico por meio da constante reflexão sobre seu objeto de estudos, o documento arquivístico e de seu arcabouço teórico-metodológico. No entanto, verifica-se que foi principalmente na Diplomática que a Arquivística encontrou as bases e fundamentos teóricos para a formulação mais segura de seus métodos de investigação e, principalmente para a formulação do método de Identificação Arquivística, conferindo cientificidade a esse método para a partir dos estudos de Tipologia Documental, próprio do campo científico da Arquivologia.

REFERÊNCIAS BIBLIOGRÁFICAS

- Campos, M. L. de A. (2001). *Linguagem Documentária: Teorias que fundamentam sua elaboração*. EDUFF, Niterói, RJ. 133p.
- Campos, M. L. de A. (2010). O papel das definições na pesquisa em ontologia. *Perspectivas em Ciência da Informação (Impresso)*, v.15, p. 220-238.
- Dahlberg, I. (1978). A Referent-oriented analytical concept theory of interconcept. *International Classification*. Frankfurt, v.5. nº3, p.142-150. (traduzido do inglês por Vânia Teixeira Gonçalves, Bolsista do CNPQ – 31p.).
- Dahlberg, I. (1983). Terminological definitions: characteristics and demands. In: *Problèmes de la définition et de la synonymie en terminologie*, 13–51. Québec: GIRSTERM.
- Dahlberg, I. (1981). Conceptual definitions for Interconcept. *International Classification* 8 (1).
- Diccionario de Terminología Archivística (DTAE). (1993). Madrid: Dirección de Archivos Estatales, Ministerio de Cultura (Espanha).
- Dicionário Brasileiro de Terminologia Arquivística (DIBRATE). (2005). Rio de Janeiro: Arquivo Nacional (Brasil).
- Martín-Palomino Benito, M., Torre Merino, J. L. L. (2000). Metodología para la identificación y valoración de fondos documentales. Madrid: Ministerio de Educación, Cultura y Deporte. 109 p. (Escuela Iberoamericana de Archivos: Experiencias y materiales).
- Mendo Carmona, C. (2004). Consideraciones sobre el método arquivístico. *Documenta & Instrumenta*, v.1, p.24-35.
- Rodrigues, A. C. (2005). A identificação como função arquivística. In: *Congresso de Arquivologia do Mercosul, 6., 2005, Campos do Jordão. Anais...* São Paulo: Associação dos Arquivistas de São Paulo, 2005.
- Rodrigues, A. C. (2008) *Diplomática contemporânea como fundamento metodológico da identificação de tipologia documental em arquivos*. Tese de Doutorado em História Social – Faculdade de Filosofia, Letras e Ciência Humanas da Universidade de São Paulo (USP), São Paulo, Brasil.
- Rodrigues, A. C. (2012). Identificação: uma metodologia de pesquisa para a Arquivística. In: *Estudos avançados em Arquivologia*. Marília, São Paulo: Oficina Universitária, Cultura Acadêmica, 2012, p. 197-215. Disponível em: <http://www.marilia.unesp.br/Home/Publicacoes/estudos_avancados_arquivologia.pdf>. Acesso em: 20 mar. 2014.
- Vieira, A. da S. L. (2007). *A identificação no processo de implementação e organização do arquivo permanente*. Marília: UNESP/FFC, 84f. TCC (Graduação em Arquivologia). Universidade Estadual Paulista, Marília.



REUSO DE ONTOLOGIAS: UMA REVISÃO DE LITERATURA NO CONTEXTO DA WEB SEMÂNTICA

Helder Firmino¹, Gercina Lima²

¹Doutorando em Gestão e Organização do Conhecimento/UFMG/Brasil, 0000-0001-8311-8618, hfirmino@eci.ufmg.br

²Professora da Escola de Ciência da Informação/UFMG/Brasil, 0000-0003-0735-3856, glima@ufmg.br

RESUMO Este artigo tem como objetivo, apresentar os resultados de uma revisão de literatura acerca do reuso de ontologias. O conceito de ontologia tem a sua gênese na Filosofia. Contudo, muitas áreas do conhecimento humano, como são os casos da Biblioteconomia e Ciência da Informação e da Ciência da Computação, adotaram o conceito de ontologia. As ontologias têm ganho enorme protagonismo nesses campos, trazendo um grande contributo na organização do conhecimento e modelagem de realidades. Como resultado dessa aposta nas ontologias, surgiram diversas aplicações práticas nomeadamente em áreas como a Biomedicina, o Direito, a Agricultura, com vários projetos de ontologias, tais como: o BFO (*Basic Formal Ontology*), o GFO (*General Formal Ontology*), o DOLCE (*Descriptive Ontology for Linguistic and Cognitive Engineering*), o *Blood Project*, GO (*The Gene Ontology*), além de vocabulários controlados como o AGROVOC, o projeto GALEN (*Generalized Architecture for Languages, Encyclopedias and Nomenclatures in Medicine*) dentre outros. Investigando o estado da arte, centrada no reuso de ontologias e inserida no contexto da Web semântica, mapeou-se o que de mais relevante tem sido produzido nos últimos anos pela comunidade científica das áreas acima mencionadas. A metodologia de investigação se enquadra na abordagem qualitativa e como procedimento metodológico seguiu-se a técnica de pesquisa bibliográfica de materiais publicados. A revisão de literatura baseou-se na revisão temática. Destaque para as seguintes bases de dados: Portal de Periódicos da CAPES, RCAAP (Repositório Científico de Acesso Aberto de Portugal), *Science Direct*, *Recherche Isidore*, *Scopus* e *Web of Science*. Analisou-se trabalhos publicados em periódicos revisados por pares e dentro do espaço temporal recomendado para a área das Ciências Sociais e em vários idiomas (português, inglês e francês). Como resultado da pesquisa bibliográfica obteve-se um panorama geral sobre o reuso de ontologias, bem como as lacunas existentes e os contributos teórico-metodológicos de cada uma das áreas analisadas.

PALAVRAS-CHAVE *organização do conhecimento, reuso de ontologias, web semântica, pesquisa bibliográfica, revisão temática.*

ABSTRACT This article aims to present the results of a literature review on ontologies reuse. The concept of ontology has its genesis in Philosophy. However, many areas of human knowledge, such as library and information science and computer science, have adopted the concept of ontology. Ontologies have gained enormous prominence in these fields, making a great contribution in knowledge organization and modeling of realities. As a result of this focus on ontologies, several practical applications have arisen in areas such as Biomedicine, Law and Agriculture, with several ontology projects such as BFO (*Basic Formal Ontology*), GFO (*General Formal Ontology*), DOLCE (*Descriptive Ontology for Linguistic and Cognitive Engineering*), *Blood Project*, GO (*The Gene Ontology*), as well as controlled vocabularies such as AGROVOC, the GALEN (*Generalized Architecture for Languages, Encyclopedias and Nomenclatures in Medicine*) among others. The state of the art was conducted with the focus on the reuse of ontologies and inserted in the context of the semantic Web. The most relevant works produced in the recent years by the scientific community of the

mentioned areas, was mapped. The research methodology is part of the qualitative approach and as methodological procedure we followed the literature search technique of published materials. The literature review was based on the thematic review. The following databases were highlighted: CAPES's Periodicals Portal, RCAAP (Open Access Scientific Repository of Portugal), Science Direct, Recherche Isidore, Scopus and Web of Science. Analyzes were carried out in peer-reviewed journals and within the recommended information retrospective for the Social Sciences area and in several languages (portuguese, english and french). As a result of the bibliographic research, a general overview was obtained on the reuse of ontologies, as well as the existing gaps and the theoretical and methodological contributions of each analyzed areas.

KEYWORDS *knowledge organization, ontologies reuse, semantic web, bibliographical seeking, thematic review.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Com a presente revisão de literatura, pretende-se investigar o que a comunidade científica das áreas da Biblioteconomia e Ciência da Informação e da Ciência da Computação têm produzido em matéria de organização do conhecimento, mais concretamente sobre o reuso de ontologias

O processo de produção de conhecimento científico deve sempre se basear em premissas sólidas. Toda e qualquer afirmação deve estar devidamente baseada em conhecimento produzido por especialistas da área de investigação, conhecimento que foi avaliado e validado pelos seus pares.

Berndtsson, Hansson, Olsson & Lundell (2008), afirmam que no processo de busca de informação para compor o seu trabalho, o investigador enfrenta muitas dificuldades e que não se resumem apenas às atividades rotineiras e racionais, sobressaindo também os aspetos emocionais. O pesquisador poderá enfrentar sentimentos como a confusão, a frustração e a dúvida, outrossim envolve o alívio e a satisfação pelo sucesso alcançado. Esses mesmos autores trazem como exemplo o *Information Seeking Process* (ISP), método desenvolvido pela Carol Kuhlthau, sendo uma das principais autoras em estudos sobre o comportamento de busca da informação. Na tabela 1, modelo de ISP da Kuhlthau adaptado por Berndtsson *et al.* (2008), podemos observar de forma resumida os sentimentos e pensamentos que o pesquisador enfrenta no processo de busca de informação.

Tabela 1. Modelo do processo de busca de informação de Kuhlthau

Estágios	Iniciação da Tarefa	Seleção do Tópico	Exploração pré-foco	Formulação do Foco	Coleta de Informação	Fecho da Busca	Início da escrita
Sentimentos	incerteza	otimismo	confusão/ frustração	clareza	senso de direção/ confiança	alívio	satisfação ou insatisfação
Pensamentos	ambiguidade				especificidade		
				aumento	interesse		
Ações	buscando informações relevantes					buscando informações pertinentes	

Referência: Adaptado de Berndtsson *et al.* (2008)

Da leitura de vários autores constatou-se que o reuso das ontologias é das primeiras situações a se ter em conta na atividade de construção de ontologias. Ora o reuso de ontologias pode poupar ao investigador muito esforço de codificação e tentativas de resolução de problemas que outras pessoas, adequadamente já resolveram. O artigo de Kless, Jansen & Milton (2016), publicado na revista *Semantic Web* vem corroborar esse fato, afirmando que “a criação de ontologias de conhecimento mais densas pode acarretar um tremendo consumo de tempo. Por esta razão, é desejável o reuso de modelos, já existentes de ontologias”.

Mas antes de adentrar na temática do reuso de ontologias, faz-se necessária a definição do conceito de ontologia. Almeida (2014), advoga que o assunto “ontologia” não está limitado a um campo de estudo específico e argumenta que “ontologia”, com letra inicial minúscula, refere-se a um artefacto e que “Ontologia”, com letra inicial maiúscula, diz respeito ao seu campo de pesquisa original, a Filosofia.

Segundo os autores Kless *et al.* (2016), “na Ciência da Informação, ontologias são declarações necessariamente verdadeiras, sobre características comuns de entidades de uma realidade em uma linguagem formal computável”.

Existem várias definições para o conceito de ontologia, no entanto a definição que é mais citada na literatura é a proposta por Gruber (1993), em que afirma que “ontologia é uma especificação explícita de uma conceptualização”. A figura 1 esclarece a noção de conceitualização, que pode ser entendida como abstrações da realidade que são baseadas em conjuntos de conceitos representados em uma linguagem.

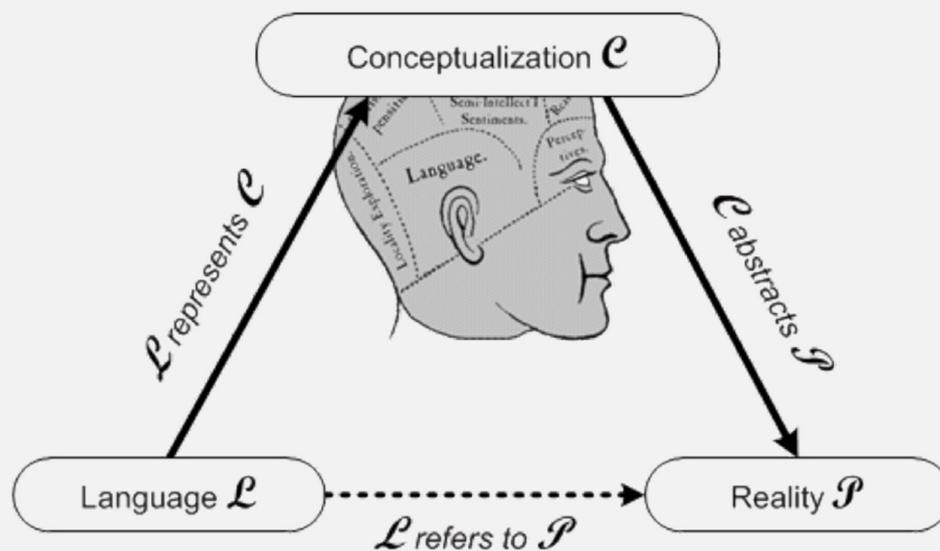


Figura 1. Triângulo de Ullman (Daniele & Brandt, 2013)

Deparou-se que o termo ontologia formal tem sido bastante focado na literatura. De acordo com Herre (2013), ontologias formais são, a grosso modo, aquelas que integram aspectos da filosofia, lógica formal, inteligência artificial e ciências cognitivas.

São exemplos de ontologias formais, o DOLCE (*Descriptive Ontology for Linguistic and Cognitive Engineering*), o *General Formal Ontology* (GFO) e o *Basic Formal Ontology* (BFO). Centrando-se nas diferenças filosóficas, temos o caso do BFO que assenta em pressupostos do realismo aristotélico. BFO é uma ontologia de alto nível inicialmente desenvolvida para uso em domínios científicos, como a biomedicina.

Por seu turno, o GFO preconiza, segundo Herre (2013), a ideia de ontologia de categorias. Diferentemente do BFO, o *General Formal Ontology* introduz uma nova forma de realismo, o chamado realismo integrativo. Ainda, segundo o mesmo autor, o realismo integrativo postula uma relação particular entre a mente e a realidade material independente.

Para Herre (2013), as ontologias exibem diferentes níveis de abstração; as ontologias de alto nível por exemplo, podem ser aplicadas a qualquer domínio de interesse, enquanto que as ontologias de domínio superior e de domínio estão relacionadas a domínios mais restritos. Ainda esse mesmo autor afirma que não há regras estabelecidas para separar esses níveis de abstração, embora exista a tendência para entender os axiomas de uma ontologia de alto nível como verdades analíticas.

Além das ontologias de alto nível, tem-se as ontologias de domínio. Um exemplo é o projeto *Gene Ontology* (GO), cujo objetivo é o de fornecer um caminho uniforme para descrever as funções dos produtos genéticos de organismos em todos os reinos da vida e, assim, permitir a análise de dados genômicos¹.

Refocando nossa atenção na questão do reuso de ontologias, tem-se o trabalho de Bastos e Café (2012) em que foi feito um levantamento sobre a utilização do assunto ontologias nas teses e dissertações dos programas de pós-graduação das universidades públicas brasileiras. Segundo os autores os resultados demonstram o crescente uso das ontologias nos campos da Ciência da Informação, Ciência da Computação e Engenharias.

Campos *et al.* (2013), asseveram que “as ontologias podem ser reutilizadas de diversas formas: que ora resultam na criação de uma ontologia independente a partir dos conceitos de outras (podendo ser estendidos e adaptados), ora preservam as ontologias originais”.

Fernández-López, Gómez-Pérez & Suárez-Figueroa (2013), desenvolveram uma metodologia que possibilita: (a) A identificação do tipo de ontologia genérica a ser reutilizado; (b) Encontrar quais axiomas e definições que poderão ser reutilizadas; (c) Tomar a decisão, usando uma análise conceitual formal, sobre o qual ontologia genérica vai ser reutilizada.

No que concerne às ontologias de domínio os autores Jiménez *et al.* (2013) asseveram que o objetivo do reuso de ontologias de domínio é de encontrar e selecionar uma ou mais ontologias de domínio que podem ser reutilizadas para desenvolver uma nova ontologia.

De acordo com as diretrizes metodológicas propostas na Metodologia NeOn, este processo de reutilização é composto por quatro atividades: A primeira consiste na pesquisa de ontologias de domínio, que poderão ser candidatas a suprir as necessidades da ontologia em construção. A segunda atividade, avaliação de ontologia de domínio, consiste em avaliar quais conjuntos de ontologias de domínio deverão ser incorporadas ou descartadas. Para uma análise mais criteriosa de cada ontologia é sugerido a título ilustrativo a seguinte heurística: Há lugar a uma rejeição de uma ontologia se a ontologia candidata não possuir escopos e/ou propósitos semelhantes. A atividade que se segue, seleção da ontologia de domínio, tem como objetivo a descoberta de quais ontologias de domínio serão as mais adequadas para o desenvolvimento da rede de ontologia. Por último temos a quarta atividade, integração

¹ Fonte: <<http://www.obofoundry.org/ontology/go.html>>

de ontologia de domínio, cujo objetivo é integrar as ontologias de domínio selecionadas na atividade 3, na ontologia em desenvolvimento (Jiménez *et al.*, 2013).

No campo da BCI, tem-se o estudo de Campos *et al.* (2013), em que não chegam a propor uma metodologia de reuso de ontologias, mas é demonstrado que os critérios de compatibilização desenvolvidos na ciência da informação são válidos para o reuso de ontologias. Etapas de reuso da ontologia podem ser assim resumidas: (a) Encontrar e selecionar as ontologias candidatas; (b) Avaliação das ontologias candidatas, por especialistas em domínio e engenheiros de ontologia; (c) Seleção final das ontologias a serem integradas; (d) Aplicação de operações de integração ontológica, que consideramos como um procedimento semi-automático.

A metodologia OntoForInfoScience da autoria de Mendonça (2015) é uma metodologia de construção de ontologias baseada em três metodologias, o NeOn, o Methontology e o Método 101. O reuso é uma tarefa que é enfatizada e altamente recomendada na proposta. Essa metodologia já foi usada e aplicada na criação de duas ontologias de domínio, a Hemonto no campo da Biomedicina e a Ontolegis na área jurídica.

No campo de Web Semântica identificou-se a tese de doutorado, da autora Paslaru-Bontas (2007), que propôs uma metodologia para reuso de ontologias.

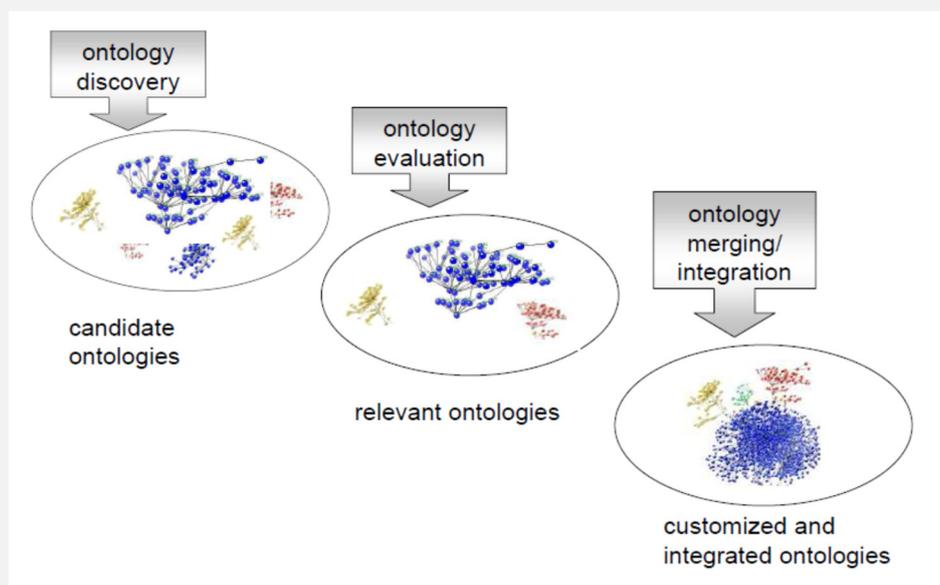


Figura 2. Processo de Reuso de Ontologias (Paslaru-Bontas, 2007)

A figura 2, mostra de forma sintética as três fases do processo de reuso de ontologias de Paslaru-Bontas. A primeira fase, chamada de Descoberta de Ontologia: é dedicada a encontrar uma lista de potenciais candidatos. A fase que se segue, Avaliação da ontologia, tem como objetivo avaliar a usabilidade das fontes ontológicas relevantes. E por último vem o *Merging* e integração de Ontologias; nesta etapa as ontologias apropriadas são customizadas e integradas na aplicação final. O resultado é na forma de um único ou grupo de ontologias representadas em uma linguagem de representação particular (Paslaru-Bontas, 2007).

Allemang & Hendler (2011) enumeram boas e más práticas de modelação. E precisamente uma dessas boas práticas aborda essa particularidade de se modelar para o reuso. Uma das principais diretivas da criação de um modelo semântico é que ele venha a ser utilizado por uma outra pessoa, que o poderá

aplicar num novo contexto. O “ontologista” deve ter o cuidado de nomear as entidades do seu modelo com nomes sugestivos e auto-explicativos.

O reuso é de acordo com Bonacin *et al.* (2016) altamente recomendável em qualquer processo de engenharia de software, mormente quando se fala do reuso de ontologias. Reuso significa que se pode tirar proveito de trabalhos anteriores e originárias de fontes credíveis, mas no entanto não é uma tarefa trivial, pois é preciso despende algum esforço para entender o que a comunidade de um domínio específico já desenvolveu para avaliar as possibilidades existentes e para adaptar os artefactos reutilizados.

O W3C (*Ford Vide Web Consortium*) acabou de publicar a sua mais recente recomendação, *Data no te Web Besta Pratiques*, datada de 31 de janeiro de 2017 e que vai precisamente suprir essa necessidade de padronizar os dados na Web e um dos benefícios esperados pela adoção desta recomendação é precisamente o reuso.

“O reuso de dados é uma outra forma de publicar dados; trata-se simplesmente de uma republicação. Ele pode assumir a forma de combinação de dados existentes com outros conjuntos de dados, criação de aplicativos Web ou visualizações, ou reagrupando os dados num novo formulário, como por exemplo uma tradução” (Lóscio, Burle, & Calcaria, 2017).

METODOLOGIA

O presente trabalho tem um carácter exploratório e descritivo, alicerçado numa abordagem qualitativa. Utilizou-se igualmente a técnica de pesquisa bibliográfica, baseada em pressupostos que são usualmente aceites na comunidade científica e concebida a partir de materiais publicados.

A primeira atividade que deverá ser levado a cabo, em um projeto de investigação, é o *Information Seeking Process*, método mencionado na secção propedêutica. Tomou-se igualmente como guia o *Research Skills Tutorial*², desenvolvido pela *Suny Empire State College*, uma universidade pública do estado de Nova York, EUA. Este tutorial cobre diversas temáticas ligadas ao processo de pesquisa e pensamento crítico, auxiliando na compreensão da complexidade do ambiente informacional moderno. (Morehouse, 2014).

A revisão de literatura, segundo Webster & Watson (2002), é essencial a um projeto académico, e para que ela seja completa deverá: (a) Ser bem fundamentada em artigos relevantes, cobrindo a literatura no tópico; (b) Ser sistêmica, com uma estruturação coerente; (c) Ser extensa; (d) Ser profunda; (e) Sublinhar as discrepâncias entre o que sabemos e o que precisamos saber; (f) Não estar confinada a uma metodologia de investigação, a um conjunto de revistas científicas, nem a uma região geográfica; (g) Justificar todas as opções tomadas.

Creswell (2014), um dos autores mais referenciados da área de metodologias de investigação, propõe cinco etapas para a condução de uma revisão de literatura. Independentemente da escolha de uma abordagem qualitativa ou quantitativa, a revisão de literatura não segue um caminho prescrito, ela dependerá do planeamento e permeará as cinco etapas de forma intercalada. As etapas atrás

² Fonte: <<http://subjectguides.esc.edu/researchskillstutorial>>.

mencionadas são as seguintes: 1) Identificação das palavras-chave a serem utilizadas na busca da literatura; 2) Localização da literatura sobre o tópico, consultando diversos tipos de materiais e bases de dados, incluindo os documentos disponíveis nas bibliotecas universitárias e na Web; 3) Avaliação criteriosa da literatura a ser selecionada para compor a revisão; 4) Organização da Literatura selecionada, construindo abstrações ou tomando notas sobre a literatura e desenvolvendo um diagrama visual (por ex.: mapa mental) sobre essa literatura; 5) Escrita da revisão de literatura reportando de forma sumarizada a literatura que deverá ser incluída na redação final da pesquisa.

Conforme defendido por Creswell (2011), existem dois modelos de revisão de literatura. Uma denominada *thematic review of the literature* e outra de *study-by-study review of the literature*. Na primeira, o pesquisador identifica um tema e cita a literatura de forma breve para documentar esse mesmo tema. Neste modelo, o autor apenas discute as ideias principais e os resultados dos estudos ao invés de detalhar cada caso particular. No caso do *study-by-study review of the literature*, é esperado que o autor forneça um resumo mais detalhado de cada estudo, agrupados em um tema mais amplo. Esta forma de revisão de literatura aparece geralmente em artigos de revistas científicas e em teses e dissertações.

O uso do mapa conceitual, que é uma técnica de organização do conhecimento, é aconselhado por vários especialistas para uma melhor definição e clarificação dos caminhos a percorrer durante a pesquisa; autores como Berndtsson *et al.* (2008) e Lima (2015) também defendem o uso dos mapas conceituais. O mapa conceitual pode ser visto “como uma representação que descreve a relação das ideias do pensamento, relação esta, pré-adequada ao longo de um processo de aprendizagem na construção do conhecimento e que vai se arquivando na memória” (Lima, 2015).

A presente revisão de literatura segue o tipo de revisão de literatura temática explorando as principais ideias e resultados de trabalhos científicos publicados no intervalo temporal de 2007 a 2017. Esta decisão de se avaliar os trabalhos publicados na última década vai de encontro à orientação sugerida por Creswell.

A pesquisa de informação, realizada entre dezembro de 2016 e janeiro de 2017, centrou-se fundamentalmente no Portal de Periódicos da CAPES, que está indexada às principais bases de dados, como o Google Acadêmico, *Science Direct*, *Scopus*, *Web of Science*, dentre outros. Utilizou-se as palavras-chave, “reuso de ontologias” para a recuperação de informação. Para a obtenção de informação mais abrangente, foram consultadas obras de referência, como dicionários e tesouros, para a captura de sinônimos.

Como estratégia de busca, utilizou-se as palavras-chave, “reuso de ontologias” para a recuperação de informação. Vale lembrar que se fez a busca nos idiomas português, inglês e francês, e sempre que possível combinando as palavras de busca com operadores lógicos para delimitar por exemplo, as áreas de conhecimento, como Ciência da Informação, Web Semântica e Agricultura. Para a obtenção de informação mais abrangente, foram consultadas obras de referência, como dicionários e tesouros, para a captura de sinônimos e termos relacionados como por exemplo, Integração de ontologias, *Merging* de ontologias.

Outra decisão tomada foi a de se consultar preferencialmente artigos publicados em periódicos revisados por pares. O *Peer Review*, é a forma mais rigorosa de controle de qualidade que existe no âmbito das fontes de informação acadêmicas. Ainda na pesquisa de artigos, vale ressaltar que uma das formas de garantir a qualidade é ter em atenção os fatores de impacto, classificação dos periódicos

(QUALIS, *Journal Citation Report*, *CiteSeer*^x). Índices de citação como o H-index (que avalia o autor) é uma excelente opção, uma vez que transmite ao pesquisador a confiança e informa da credibilidade do autor e é também uma forma de valorizar o esforço e trabalho de outros.

As fontes de informação podem ser organizadas, com recurso a ferramentas de gestão de citações bibliográficas, como são os casos de *EndNote*, *BibMe*, *Citation Machine*, *Zotero* e *Mendeley*, estes dois últimos são ferramentas de utilização livre e com suplementos para os editores de texto *Microsoft Word* e *OpenOffice Writer*.

RESULTADOS

Esta pesquisa visava essencialmente a pesquisa bibliográfica do tema, reuso de ontologias no contexto da Web semântica. Para a materialização desse desiderato foram identificados trabalhos nos campos da Biblioteconomia/Ciência da Informação e da Ciência da Informação, que tinham como objeto de estudo o reuso de ontologias ou mencionavam o reuso no processo de construção de ontologias. A tabela 2 apresenta o ponto de vista de cada um dos campos científicos, mostrando quais são os contributos teóricos que elas trazem para a organização do conhecimento e mais especificamente que metodologias, com enfoque no reuso, têm sido propostos para a construção de ontologias.

Tabela 2. Comparativo sobre Reuso de Ontologias, segundo a visão da Ciência da Informação e da Ciência da Computação

	Ciência da Informação	Ciência da Computação
Reuso de Ontologias (contributos/noções)	<p>Fornecendo uma base semântica para integração. São exemplos de sustentação teórica: Princípios de organização do conhecimento de Hjørland, a Teoria da Análise Facetada (TAF) de Ranganathan e a Teoria do Conceito de Dahlberg, que permitem a representação de domínios de conhecimento.</p> <p>Fundamentos filosóficos e lógicos das ontologias formais na ciência da informação, em particular o BFO (<i>Basic Formal Ontology</i>) de Barry Smith.</p> <p>Methodological Aspects Applied to Ontology Reuse Tools: A Study Based on Genomic Annotations in the Domain of Trypanosomatides (Campos <i>et al.</i>, 2013)</p>	<p>Processo em que conhecimento ontológico existente está baseado em entradas para gerar novas ontologias (Paslaru-Bontas, 2007)</p> <p>O objetivo do reuso de ontologias de domínio é encontrar e selecionar uma ou mais ontologias de domínio que possam ser reutilizadas para desenvolver uma nova ontologia (Jiménez, Suárez-Figueroa, Mateos, Gómez-Pérez, & Fernández-López, 2013)</p> <p>O reuso de ontologias envolve a construção de uma nova ontologia através da maximização da adoção de ontologias pré-usadas ou componentes de ontologia. A reutilização tem várias vantagens. Primeiro, reduz o trabalho humano envolvido na formalização de ontologias a partir do zero (Lonsdale, Embley, Ding, Xu, & Hepp, 2010)</p>
Metodologias de Reuso de Ontologias	<p>Tomou-se conhecimento da OntoForInfoScience, uma metodologia que suporta o processo de construção de ontologias da área da Ciência de Informação, em que são descritos e relacionados explicitamente as ontologias reutilizadas. Esta metodologia tem como base três metodologias: o NeOn Methodology, Methontology e o Método 101 (Mendonça, 2015).</p>	<p>Metodologia NeOn (Jiménez <i>et al.</i>, 2013) (Fernández-López <i>et al.</i>, 2013)</p> <p>Automated ontology reuse approach (Lonsdale <i>et al.</i>, 2010).</p> <p>Processo de Reuso de Ontologias (Paslaru-Bontas, 2007)</p> <p>A Methodology for Ontology Integration (Pinto & Martins, 2001)</p> <p>Conceptualizing the Impacts of Agriculture on Water Resources: experiences and ontology engineering challenges (Bonacin <i>et al.</i>, 2013)</p> <p>Ontology Evaluation for Reuse in the Domain of Process Systems Engineering (Trokanas, N. & Cecelja, F., 2016)</p>

Referência: Elaboração própria

Os procedimentos utilizados para a realização desta revisão de literatura foram os seguintes:

- 1) Desenho de um mapa conceitual para melhor compreensão da temática a explorar;
- 2) Ter em mente o processo de pesquisa informacional sustentado no ISP de Kuhlthau;
- 3) Definir as palavras-chave, nos idiomas português e inglês, conjugados com operadores booleanos e uso de chavetas para refinamento da busca;
- 4) Selecionar e delimitar as bases de dados de referência que suportariam o processo de pesquisa;
- 5) Gerir toda a bibliografia com recurso à ferramenta de organização de referências bibliográficas Mendeley;
- 6) Criar uma matriz de conceitos, cruzando a informação do Mendeley para ter uma base de conhecimento de conceitos, autores e citações.

CONCLUSÕES

O reuso de ontologias é uma área que está mais sedimentada no campo da Ciência da Computação, que é sustentada pela existência de diversas ferramentas conceituais (artefatos) e de metodologias de reuso de ontologias. Por seu turno, a área da Biblioteconomia e da Ciência da Informação empresta seus contributos na organização e representação do conhecimento, trazendo subsídios no domínio semântico para o reuso e integração, sustentada em aportes da área como os princípios de organização do conhecimento de Hjørland, a Teoria da Análise Facetada de Ranganathan e a Teoria do Conceito de Dahlberg.

A revisão de literatura reveste-se de suma importância, pois se ela for bem conduzida, facilita o trabalho de investigação. Com esta revisão constatou-se que apesar dos esforços, existem ainda poucos estudos no campo da Biblioteconomia e da Ciência da Informação, pelo que é um campo com grande potencial e margem de crescimento. Tomou-se conhecimento da metodologia de construção de ontologias OntoForInfoScience, uma metodologia proposta por Mendonça (2015) e que já foi usada na criação de duas ontologias de domínio, a Hemonto e a Ontolegis.

Por outro lado, na área da Ciência da Computação existe uma grande atividade de investigação e que passou a contar com os padrões e tecnologias da Web semântica. Como exemplos de tecnologias, consideradas pilares, da Web semântica tem-se o RDF (*Resource Description Framework*), o OWL (*Web Ontology Language*) e o SPARQL (*SPARQL Protocol And RDF Query Language*). Com essas tecnologias ganhou-se um grande impulso na representação do conhecimento.

O conhecimento científico não é estanque, surgindo sempre novas possibilidades de pesquisa. Identificou-se alguns cenários que poderão ser alvo de melhoramentos futuros, como a busca de informação em outras bases de dados e incluindo, por exemplo, trabalhos na língua castelhana.

Em jeito de síntese pode-se afirmar que a questão do reuso de ontologias tem sido cada vez mais estudado e aplicado no processo de criação de instrumentos de organização do conhecimento. Prova disso são as muitas metodologias surgidas com o intuito de auxiliar o processo de desenho de ontologias e em que a tônica do reuso está posta de forma explícita.

REFERÊNCIAS BIBLIOGRÁFICAS

- Allemang, D., & Hendler, J. (2011). *Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL* (2nd ed). Waltham, MA: Morgan Kaufmann/Elsevier.
- Almeida, M. B. (2014). Uma abordagem integrada sobre ontologias: Ciência da Informação, Ciência da Computação e Filosofia. *Perspectivas em Ciência da Informação*, 19(3), 242–258.
- Bastos, M.; Café, L. (2012). Ontologias: Abordagens nas teses e dissertações das Universidades Públicas Brasileiras. p. 81–98. Disponível em: <<http://www.ies.ufpb.br/ojs2/index.php/ies/article/view/10054/7759>>.
- Berndtsson, M., Hansson, J., Olsson, B. & Lundell, B. (2008). *Thesis Projects. A Guide for Students in Computer Science And Information Systems* (2nd ed). London: Springer-Verlag London.
- Campos, M. L. de A., Campos, M. L. M., Dávila, A. M. R., Gomes, H. E., Campos, L. M., & Lira e Oliveira, L. (2013). Information Sciences Methodological Aspects Applied to Ontology Reuse Tools: A Study Based on Genomic Annotations in the Domain of Trypanosomatides. *Knowledge Organization*, 40(1), 50–61.
- Creswell, J. W. (2011). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. (4^a edição). Boston, MA: Pearson Education, Inc.
- Creswell, J. W. (2014). *Research design: qualitative, quantitative, and mixed methods approaches*. (4^a edição). Los Angeles: SAGE Publications Inc.
- Daniele, L., & Brandt, P. (2013). Pragmatism versus formalism : The relation between Linked Open Data, semantics and ontologies. In E. Folmer, M. Reuvers, & W. Quak (Orgs.), *Pilot linked open data Nederland: Deel 1 Het Managementoverzicht* (p. 275–282). Amersfoort: Bizzprint.
- Fernández-López, M., Gómez-Pérez, A., & Suárez-Figueroa, M. C. (2013). Methodological guidelines for reusing general ontologies. *Data & Knowledge Engineering*, 86, 242–275. <https://doi.org/10.1016/j.datak.2013.03.006>
- Gruber, T. R. (1993). A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition* (Vol. 5). Palo Alto, CA.
- Jiménez, A., Suárez-Figueroa, M. C., Mateos, A., Gómez-Pérez, A., & Fernández-López, M. (2013). A MAUT APPROACH FOR REUSING DOMAIN ONTOLOGIES ON THE BASIS OF THE NEON METHODOLOGY. *International Journal of Information Technology & Decision Making*, 12(05), 945–968. <https://doi.org/10.1142/S021962201340004X>
- Kless, D., Jansen, L., & Milton, S. (2016). A content-focused method for re-engineering thesauri into semantically adequate ontologies using OWL. *Semantic Web*, 7(5), 543–576. <https://doi.org/10.3233/SW-150194>
- Lima, G. Â. de. (2015). *MHTX - MODELAGEM HIPERTEXTUAL PARA ORGANIZAÇÃO DE DOCUMENTOS Princípios e Aplicação*. Rio de Janeiro: Editora Interciência Ltda.

Lonsdale, D., Embley, D. W., Ding, Y., Xu, L., & Hepp, M. (2010). Reusing ontologies and language components for ontology generation. *Data & Knowledge Engineering*, 69(4), 318–330.

Lóscio, B. F., Burle, C., & Calegari, N. Data on the Web Best Practices (2017). W3C.

Morehouse, S. (2014). Research Skills Tutorial. Disponível em: <<http://subjectguides.esc.edu/researchskillstutorial>>. Acesso em: 9 jun 2017.

Mendonça, F. (2015). *OntoForInfoScience: Metodologia para construção de ontologias pelos cientistas da informação: Uma aplicação prática no desenvolvimento da ontologia sobre componentes do sangue humano (HEMONTO)*. (Tese de Doutorado). Universidade Federal de Minas Gerais, Belo Horizonte. Disponível em: <http://www.bibliotecadigital.ufmg.br/dspace/handle/1843/BUBD-A35H3K>

Paslaru-Bontas, E. (2007). *A Contextual Approach to Ontology Reuse: Methodology, Methods and Tools for the Semantic Web*. Freien Universitat Berlin.

Pinto, H. S., & Martins, P. (2001). A Methodology for Ontology Integration. In *K-CAP '01 Proceedings of the 1st international conference on Knowledge capture* (p. 131–138).

Trokanas, N. & Cecelja, F. (2016). Ontology Evaluation for Reuse in the Domain of Process Systems Engineering. *Computers & Chemical Eng.*, 85, p. 177–187.

Webster, J., & Watson, R. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, 26(2), 13–23.



RELACIONES CIENTÍFICAS Y ORGANIZACIÓN DEL CONOCIMIENTO

Carlos Olmeda-Gómez¹, Daniel Martínez-Ávila², María Antonia Ovalle-Perandones³

¹Universidad Carlos III de Madrid, Departamento de Biblioteconomía y Documentación, olmeda@bib.uc3m.es

²UNESP-Universidade Estadual Paulista, Department of Information Science, dmartinezavila@marilia.unesp.br

³Universidad Carlos III de Madrid, Departamento de Biblioteconomía y Documentación, movalle@bib.uc3m.es

RESUMEN El análisis de cocitación de autores (ACA) es un método importante para revelar la base intelectual de una especialidad. En el trabajo se ha empleado un análisis tradicional a partir de los artículos contenidos en la base Scopus, publicados en la revista *Knowledge Organization* entre 1993 y 2016. En este estudio se emplea una idea un método de representación visual de similitudes basado en distancias. Los resultados iniciales del análisis presentan cuatro clusters de autores. Los mapas permiten identificar los grupos de autores que tienen mayor interconexión entre sí y en el conjunto de la red.

PALABRAS CLAVE Análisis de cocitación de autores, redes bibliométricas, visualización de dominios, organización del conocimiento.

ABSTRACT Authors' cocitation analysis (ACA) is an important method for revealing the intellectual base of a specialty. In the work a traditional analysis has been used from the articles contained in the Scopus database, published in the *Knowledge Organization* journal between 1993 and 2016. In our approach a distance based visualization is presented. The initial results of the analysis present four clusters of authors. It suggest four scientific relatedness areas The maps allows to identify the groups of authors that have greater interconnection with each other and in the whole of the network.

KEYWORDS Author cocitation analysis, bibliometric networks, domain visualization, knowledge organization

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUCCIÓN

La Organización del Conocimiento es un campo científico que estudia la descripción, representación, archivo y organización de documentos, representaciones de documentos, materias y conceptos, tanto por humanos como por programas informáticos (Hjørland 2008). Uno de los principales canales de discusión y difusión de la investigación en la disciplina es la revista *Knowledge Organization* (KO), fundada en 1974 por Ingetraut Dahlberg bajo el nombre *International Classification* (ISSN 0340-0050). En 1989, con la fundación de la *International Society for Knowledge Organization* (ISKO), la revista se convirtió en su órgano oficial para la difusión de la investigación en dicha área, y, en 1993, cambió su nombre por el actual *Knowledge Organization*.

El objetivo de la revista, tal como queda plasmado en su primera editorial (Dahlberg 1974), fue el de superar los silos de las escasas sociedades nacionales sobre clasificación (como el británico *Classification Research Group* o el comité FID/CR de la *Federation Internationale de Documentation*) y proporcionar un foro internacional para el intercambio de conocimiento entre éstas y otras escuelas nacionales de investigación en cuestiones interdisciplinarias relacionadas con la organización del conocimiento. Tal como indica Hjørland (2008), no solo existen varias tradiciones y perspectivas dentro de la Organización del Conocimiento, sino que además la Organización del Conocimiento en sí misma es una entre tantos campos científicos que compiten en relación a la comunicación e intercambio de conocimiento. Autores, teorías y diversas escuelas de pensamiento compiten dentro de un campo científico para establecer los paradigmas dominantes y la configuración del campo y su relación respecto a otros

Dentro de la Organización del Conocimiento, los análisis bibliométricos y de cocitación están comúnmente asociados con el Análisis de Dominio. Birger Hjørland, creador del concepto de análisis de dominio tal como se entiende en Organización del Conocimiento junto a Anne Albrechtsen (1995), también reconoció que los estudios bibliométricos son una de las posibles y formas principales de realizar análisis de dominio de manera empírica (2002). En este sentido, se señalan los estudios de Garfield (ej. 1976) como pioneros del análisis de dominio realizados en el campo de la Bibliometría, y a White y McCain (1998) como pioneros en reconocer la influencia de Hjørland en sus estudios bibliométricos. Los análisis de dominio de conocimiento conforman una sub especialidad en el seno de los estudios de *Library and Information Science*. Se entiende por este término, un área interdisciplinar de conocimientos dirigida a cartografiar, realizar minería de contenidos, clasificar, analizar y posibilitar la navegación y la presentación de conocimientos (Shiffrin and Börner 2004).

En la literatura de Organización de Conocimiento existen numerosos ejemplos de análisis de dominio utilizando estudios bibliométricos (ver el ejemplo de Smiraglia 2015, para una amplia revisión del concepto). En la revista *Knowledge Organization* también hay ejemplos publicados sobre análisis de citas relacionados con esta subespecialidad. Por ejemplo, la investigación de España sobre Organización del Conocimiento, ha sido analizada de forma exhaustiva en dos trabajos que cubren dos periodos de estudio, entre 1992 y 2001 (López Huertas y Jiménez Contreras, 2004), y 2002-2010 (Moneda Corrochano et al., 2013). La conclusión principal de este segundo trabajo, es reconocer la consolidación de la investigación en KO en España y un incremento en la citación que obtienen los trabajos españoles, así como la renovación de los autores prolíficos, respecto de los datos obtenidos en el estudio anterior. Oliveira et al. (2010) realizan un análisis de cocitación de autores brasileños, a partir de 134 trabajos presentados entre 2003 y 2008 en *Encontros Nacionais de Pesquisa da Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (ENANCIBs)*. Otros estudios han abordado aspectos muy concretos, por ejemplo, la ética y el KO y su representación en las bases internacionales del *Web of Science* y en *Library, Information Science & Technology Abstracts* (Martínez-Ávila et al., 2015) o análisis sobre el conjunto de la obra producida por investigadores destacados en la especialidad, como Hope Olson (Wolfram, 2016).

Fuera de la revista *Knowledge Organization*, aunque también en el entorno ISKO por ser presentado en uno de sus congresos internacionales, Guimãraes et al. (2012) estudian los referentes teóricos en Organización del Conocimiento utilizando como corpus de investigación los artículos de la revista indexados por *Web of Science* en el periodo 1993-2011 y trabajando con redes de cocitación. También en el contexto ISKO, en un congreso de capítulo británico de ISKO, Guimãraes et al. (2015) presentaron

un análisis de las comunidades epistémicas en Organización del Conocimiento, utilizando como corpus los artículos de la revista indexados por Scopus en el periodo 1993-2013.

El problema de investigación que se indica y al que se da respuesta en este trabajo es la necesidad de identificar de manera empírica los autores más relevantes, relaciones y escuelas de investigación que configuran el dominio de la Organización del Conocimiento mediante la construcción de mapas de ciencia.

METODOLOGÍA

Como método para abordar el presente estudio, adoptamos un procedimiento genérico para los análisis de dominio y su visualización, que se realiza mediante la construcción de mapas de ciencia. El alcance o finalidad de los estudios mediante mapas de ciencia son las disciplinas científicas, un campo de investigación o un área de temas o materias en concreto. Con otras palabras, la unidad de análisis es un dominio de conocimiento tal y como queda reflejado por el conjunto de contribuciones intelectuales aportadas por los miembros de una determinada comunidad científica o subespecialidad, como en nuestro caso. La finalidad del análisis debe ser dar a conocer los componentes principales que son relevantes al programa de investigación en torno al cual se organiza el dominio o la subespecialidad. En nuestro trabajo, el objetivo inicial de nuestro interés se centra en conocer los autores que trabajan en el dominio y sus relaciones, tal y como quedan reflejadas en sus publicaciones mediante la técnica de análisis de cocitación de autores.

Datos. Los datos de análisis de entrada proceden de los registros de la revista *Knowledge Organization*. Se han recuperado a partir de los contenidos en la base bibliográfica *Scopus*. Se ha empleado el ISSN de la revista 0943-7444, en el campo correspondiente como argumento de búsqueda, a través de la suscripción a Scopus que mantiene en España la Fundación Española para la Ciencia y la Tecnología (FECYT). Los años de análisis son los comprendidos entre 1993 y 2016. La base de datos Scopus es una de la herramientas consideradas por Smiraglia (2015) para realizar análisis de dominio utilizando técnicas bibliométricas en Organización del Conocimiento. Los registros recuperados en esta primera operación en marzo de 2017, se han procesado para eliminar del análisis, aquellos que no fueran tipificados en Scopus como artículos, eliminándose así del análisis comunicaciones a congresos, revisiones, editoriales y cartas. Con estos criterios se han obtenido 359 artículos. Los registros se han descargado con toda la información disponibles en las secciones de información de citas, información bibliográfica, resumen y palabras clave, detalles de financiación y otra información, donde se incluyen las referencias.

Análisis de cocitación de autores. Uno de los enfoques para estudiar un dominio y mapearlo es la utilización de las técnicas de cocitación de autores. En este caso, la unidad de análisis son los autores. El análisis se basa la detección de aquellos autores que son citados conjuntamente en las listas de referencia por un trabajo citante escrito con posterioridad. La fuerza del vínculos entre los autores, queda determinado por el número de veces que los artículos citantes incluyen conjuntamente los trabajos citados conjuntamente en sus listados de referencias.

Se estima que los análisis de cocitación de autores representan mejor aquellos más antiguos en la especialidad, ya que para poder ser representados, deben tener acumulados un número importante de citas. Los valores de asociación entre los autores denotan un grado de vinculación cognitiva entre los representados, más afín entre los que sean más próximos.

Visualización. Existen diferentes aproximaciones técnicas que se han propuesto para visualizar dominios de conocimiento. Genéricamente pueden distinguirse entre las que basan el enfoque en las distancias, aquellas otras que se basan en la construcción de grafos y, por último, las basadas en representaciones temporales. De la lectura de trabajos previos, el enfoque basado en redes es el que, hasta la fecha, se ha empleado por autores que han analizado el dominio representado por la revista *Knowledge Organization*.

En nuestro caso, hemos seleccionado una herramienta *VOSviewer* (van Eck y Waltman 2010). *VOSviewer* contruye mapas con un enfoque basado en distancias. Las distancias entre dos nodos indican de forma aproximada el grado de relación entre los mismos. En general, cuanto menor sea la distancia entre dos nodos, mayor es su relación. En las redes bibliométricas existen con frecuencia, grandes diferencias entre los nodos que se representan debido a las diferencias entre los vínculos que hay entre los nodos; en nuestro caso, entre los autores medidas sus relaciones por los valores de cocitación que les unen. Por este motivo es necesario realizar un procesos de normalización de estas diferencias entre los nodos. En nuestro caso hemos empleado la técnica de normalización de los valores de coocurrencia basándonos en la fuerza de asociación.

Después de haber construido la red normalizada, se requiere que los nodos de la red se posicionen en un espacio bi-dimensional, de modo tal que los nodos más fuertemente enlazados se ubiquen más próximos unos con otros, mientras aquellos relacionados debilmente lo hagan muy lejos uno del otro. En este caso se ha empleado la técnica de visualización de similitudes de la que dispone el programa.

Por defecto, los nodos representados mediante *Vosviewer* se asignan a un cluster. Un cluster es un conjunto de nodos relacionados fuertemente entre sí. Cada nodo se asigna a un solo cluster, no permitiéndose, la vinculación de un mismo nodo a clusters diferentes. La técnica que emplea el programa para realizar el clustering es una variante de la función de modularidad introducida por Newman y Girvan (2004).

La presentación gráfica de la red de cocitación de autores realizada en este trabajo se muestra de dos modos. La primera, se representa mediante las técnicas de visualización de mapeo de similitudes, basadas en distancias en espacios bi-dimensionales, descritas más arriba de forma sumaria. En este caso son representados con gráficos de redes cuyos nodos son agrupados a partir de procedimientos de *clustering*. Los clusters de nodos se distinguen por la asociación de cada uno de ellos con colores diferentes. Un según modo de presentar los resultados, se ha realizado mediante la construcción de una visualización de densidades. En este tipo de visualización, los colores indican el modo en que los nodos se han distribuido en el espacio bidimensional que subyace en la representación. La visualización de densidades permite identificar de forma inmediata en qué área los autores se localizan más cercanos los unos de los otros y su vinculación cognitiva.

RESULTADOS

La Figura 1 muestra el número de artículos indexados por Scopus durante el periodo estudiado. El aumento de artículos indexados en la última década explicado por el aumento de números publicados por año por la revista, pasando de 4 números por año hasta 2010, a 6 números por año en el periodo 2011-2014, y 8 números por año a partir de 2015.

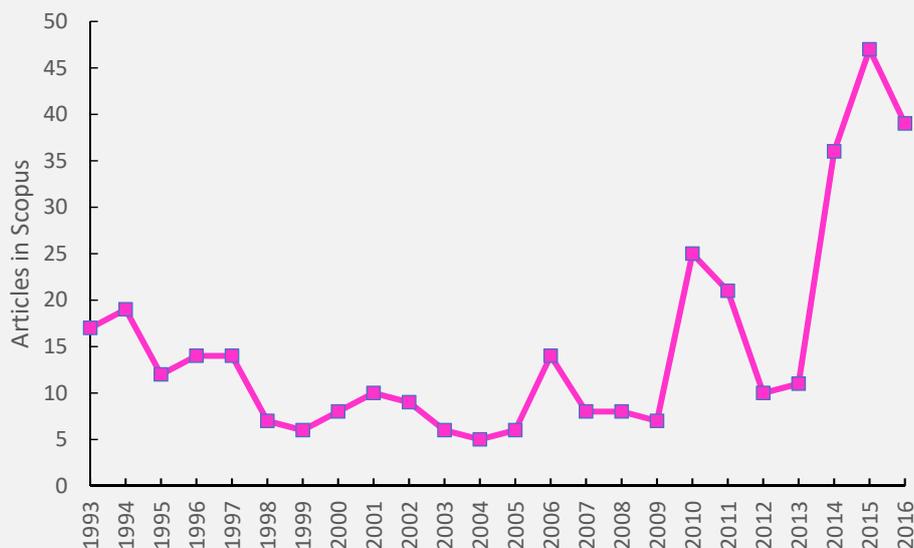


Figura 1. Distribución de registros bibliográficos de artículos en la revista *Knowledge Organization* indexados por Scopus. (1993-2016)

En la Figura 2 se muestran los principales clusters de cocitación de autores en la revista *Knowledge Organization*. El mayor de los clusters (representado en rojo), muestra un grupo de autores en Organización del Conocimiento, incluyendo bastantes autores británicos y canadienses, que trabajan en una dimensión aplicada de la Organización del Conocimiento. En este cluster se identifican autores "clásicos" como FW Lancaster, y otros que han trabajado en la elaboración de vocabularios controlados tradicionales como clasificaciones y tesauros, como Lois Mai Chan, Allan Gilchrist y Michèle Hudon, autores que han sido editores de la Clasificación Decimal Universal, como Ia McIlwaine, editores de la Clasificación Decimal de Dewey como Rebecca Green, o conocidos por su trabajo con OCLC, como Diane Vizine-Goetz, y también autores que trabajan en herramientas y estándares de organización del conocimiento más actuales como *Linked Data* y los Requisitos funcionales para registros bibliográficos (FRBR), como Marcia Lei Zeng, y etiquetado social, como Margaret Kipp y Alistair Shiri. El propio Melvil Dewey también aparece en este cluster.

El segundo cluster en tamaño, representado en la Figura 1 en color verde, representa la vertiente más epistemológica de la Organización del Conocimiento. En dicho cluster resalta de forma prominente la figura de Birger Hjørland, precursor junto a la también presente Anne Albrechtsen de la corriente de análisis de dominio y reconocido pragmatista. La presencia y cercanía en el cluster de un autor supuestamente ajeno a la Organización del Conocimiento como es Thomas Kuhn es un reflejo de la investigación en aspectos relacionados con la Historia de la Ciencia y configuración de dominios a través de análisis de dominio. Otros figuras prominentes en el cluster incluyen autores que han trabajado con teorías críticas y aspectos éticos en Organización del Conocimiento, como Hope Olson, famosa por su trabajo feminista, Clare Beghtol, famosa por su trabajo sobre la "garantía ética", Sanford Berman, pionero en el análisis crítico de los encabezamientos de materia, Oh Dong-Geun, traductor de la Clasificación Decimal de Dewey y autor de un seminal trabajo crítico sobre la representación de la clase de religión en Corea, y el español Antonio García Gutiérrez, autor de numerosos trabajos críticos y postepistemológicos que incluyen, entre otras, teorías postcoloniales (un aspecto que también comparte con Hope Olson). Otra figura que también sobresale en el cluster es Richard Smiraglia, editor de la revista *Knowledge Organization* desde 2004, y autor conocido principalmente su aportación al concepto de "obra" (*work*), y, más recientemente, por cuestiones relacionadas con el análisis de dominio. Desde

un punto epistemológico, la agrupación de autores "críticos" con los exponentes del análisis de dominio se justifica teóricamente por la asimilación del análisis de dominio en la escuela pragmatista, y la inclusión de las teorías críticas en esta escuela (ej. Hjørland 2013). De hecho, el análisis de dominio de Hjørland puede considerarse la continuación teórica de la corriente epistemológica creada a mitades del siglo pasado por Jesse Shera y Margaret Egan, con su proyecto de epistemología social fuertemente influenciado por los autores pragmatistas del siglo pasado como John Dewey y Charles Peirce, e incorporando las posteriores propuestas de teóricas de Kuhn. Por otro lado, un aspecto que sin embargo añade complejidad y merecería mayor reflexión teórica es la relación y características de Richard Smiraglia en el cluster, ya que se trata de un reconocido empirista (un aspecto que aunque probablemente no presenta problemas desde el punto de vista de Smiraglia sí que podría ser conflictivo desde el punto de vista de Hjørland) que a su vez también reconoce los estudios epistemológicos, como su propuesta de sinergia cultural (Smiraglia, 2014) también fuertemente influenciada por la epistemología social de Shera.

Otra figura que también merece especial mención en la presencia de Grant Campbell en el cluster de color rojo, ya que se trata de un autor famoso por su trabajo crítico en Teoría Queer y análisis de discurso. Sin embargo, esta posición se justifica por su fuerte relación con Margaret Kipp, situada en el cluster rojo, a quien dirigió la tesis doctoral, y su trabajo sobre etiquetado social.

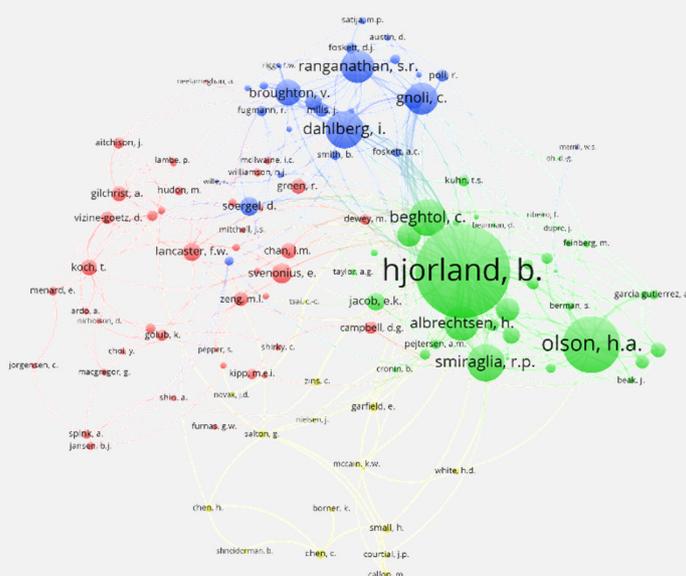


Figura 2. Un mapa basado en una red de cocitación de 131 autores y 2890 enlaces con cuatro clusters. *Knowledge Organization* (1993–2016)

El tercer cluster en tamaño, de color en azul en la Figura 1, representa una visión más tradicional y universalista de la Organización del Conocimiento, incluyendo autores conocidos también por su trabajo en clasificación y ontologías bajo una lógica racionalista. En este cluster se encuentra el renombrado clasificacionista indio Ranganathan, cuyo sistema Colon Classification ha sido discutido como racionalista en Hjørland (2014), discípulos de la escuela de Ranganathan como MP Satija, especialistas en clasificación también desde un punto de vista universalista como Vanda Broughton, conocida por su trabajo con la Clasificación de Bliss, autores que trabajan en ontologías como Roberto Poli y Barry Smith (este último, además de ser uno de los grandes nombres de la Filosofía

contemporánea, también desde una postura alejada de la óptica pragmatista, tal como es indicado en Martínez-Ávila y Fox, 2015), autores también famosos por su visión universalista, como Claudio Gnoli y, sobre todo, Ingetraut Dahlberg, fundadora de la revista *Knowledge Organization* y también de la ISKO (la cual ha sido identificada como representante de una escuela epistemológica antagónica a la de Hjørland en, por ejemplo, Guimãraes et al., 2015). Digno de reseñar también es la cercanía, aunque en el cluster de color rojo, de autoras contemporáneas a Dahlberg como Nancy Williamson y McIlwaine, un aspecto que puede sugerir un perfil en la edad de los trabajos cocitados.

Por último, el cuarto cluster marcado con color amarillo y situado en la parte inferior del grafo muestra un grupo heterogéneo de autores internacionales próximos a los estudios bibliométricos y relacionados genéricamente con la creación de mapas de ciencia, E. Garfield, H. Small, Howard D. White, y Katherine McCain; retrievalistas como George Salton; vinculados a la visualización de información, como Chaomei Chen y Katy Borner; con la interacción hombre-ordenador como Ben Shneiderman y Jacob Nielsen; o fundadores de los análisis de co palabras, Courtial y Callon.

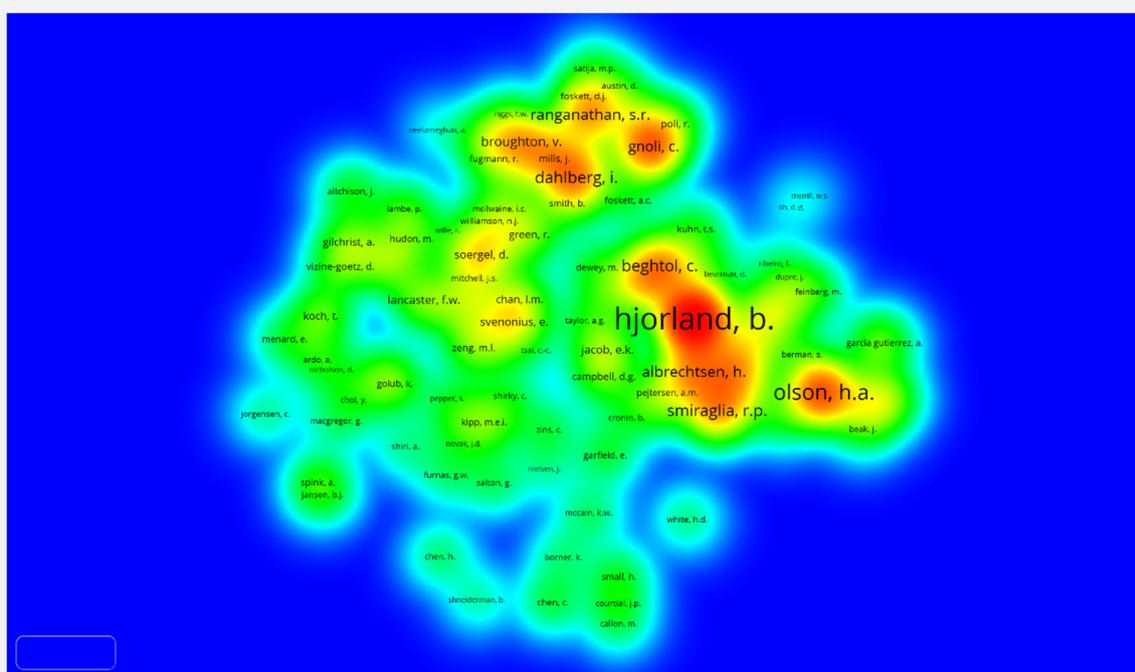


Figura. 3 Mapa de cocitación de autores basado en densidades. *Knowledge Organization* (1993-2016)

CONCLUSIONES

El objetivo principal de un análisis de cocitación de autores es personificar un campo de estudio por medio de la organización o disposición de los nombres de los autores, de modo tal que el dominio sea comprensible. Los autores que se han representado y cuyos nodos, por el tamaño de su diámetro son los mayores y agrupados en diferentes clusters, son aquellos que denotan orientaciones científicas o académicas próximas. Inicialmente el subdominio de la comunidad de autores que publican en *Knowledge Organization*, es una comunidad que, en el pasado, se ha construido sobre una base intelectual de autores que aparecen identificados, representados, y destacados en los mapas que se ha generado. En todo caso, la técnica solo muestra aquellos visibles a partir de los datos extraídos de la base de datos empleada. Una limitación del presente trabajo es la que se deriva de la delimitación del

dominio. Se ha entendido a partir las publicaciones contenidas en un único título de revista. En este sentido, los resultados, se deben entender que son provisionales y circunscritos únicamente a esta decisión. El resultado obtenido con esta aproximación resulta en una primera instancia, intelectualmente aceptable a nuestro entender.

REFERENCIAS BIBLIOGRÁFICAS

- Dahlberg, I. (1974). Editorial. *International Classification*, 1(1), 1.
- Garfield, E. (1976). Social-Sciences citation index clusters. *Current Contents*, 27, 5-11.
- Guimarães, J.A.C., de Oliveira, E.T., Gracio, M.C.C. (2012). Theoretical referents in knowledge organization: a domain analysis of Knowledge Organization journal. In A. Neelameghan & K.S. Raghavan (Eds.), *Categories, Contexts, and Relations in Knowledge Organization. Proceedings of the Twelfth International ISKO Conference, 6-9 August 2012, Mysore, India* (pp. 31-38). Würzburg: Ergon Verlag.
- Guimarães, J.A.C., Martínez-Ávila, D., Alves, B.H. (2015). Epistemic communities in knowledge organization: an analysis of research trends in the Knowledge Organization journal. Presented in *Knowledge Organization – making a difference: The impact of knowledge organization on society, scholarship and progress. ISKO UK biennial conference 13th – 14th July 2015*, London.
- Hjørland, B. (2002). Domain analysis in Information science: eleven approaches-traditional as well as innovative. *Journal of documentation*, 58(4), 422-462.
- Hjørland, B. (2008). What is Knowledge Organization (KO)? *Knowledge Organization*, 43(6), 475-484.
- Hjørland, B. (2013). Theories of Knowledge Organization-Theories of Knowledge. *Knowledge Organization*, 40(3), 169-181.
- Hjørland, B. (2014) Is Facet Analysis Based on Rationalism? *Knowledge Organization*, 41(5), 369-376.
- Hjørland, B., Albrechtsen, H. (1995). Toward a new horizon in information science: domain-analysis. *Journal of the American Society for Information Science*, 46(6), 400-425.
- López-Huertas, M.J., Contreras, E.J. (2004). Spanish research in knowledge organization. *Knowledge Organization* (1992-2001), 31(3), 136-150.
- Martínez-Ávila, D., Fox, M.J. (2015). The construction of ontology: a discourse analysis. In R.P. Smiraglia, & H-L Lee (Eds.) *Ontology for Knowledge Organization* (pp. 13-37). Würzburg: Ergon Verlag.
- Martínez-Ávila, D., Guimarães, J.A.C., Assis Pinho, F., Fox, M.J. (2015). The representation of ethics and knowledge organization in the WoS and LISTA databases. *Knowledge Organization*, 42(5), 269-275.
- Moneda Corrochano, M., López-Huertas, M.J., Jiménez Contreras, E. (2013). Spanish research in knowledge in knowledge organization (2002-2010). *Knowledge Organization*, 40(1), 28-41.

Newman, M.E.J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, 69(2), 026113.

Oliveira, E.F.T., Gracio, M.C.C., Silva, A.C.C. (2010). Investigadores de mayor visibilidad en organización y representación del conocimiento: un estudio desde el análisis de cocitaciones. *Scire*, 16(2), 39-45.

Shiffrin, R.M., & Börner, K. (2004). Mapping knowledge domains. *Proceedings of the National Academy of Sciences of the United States of America*, 101 (suppl. 1), 5183-5185.

Smiraglia, R.P. (2014). *Cultural synergy in information institutions*. New York, NY: Springer.

Smiraglia, R.P. (2015). *Domain analysis for knowledge organization: tools for ontology extraction*. Waltham, MA: Chandos Publishing.

Van Eck, N.J., Waltman, L. (2010) Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523-538.

White, H.D., McCain, K.W. (1998). Visualizing a discipline: An author co-citation analysis of information science, 1972-1995. *Journal of the American Society for Information Science* 49(4), 327-355.

Wolfram, D. (2016). The power to influence: an informetric analysis of the works of Hope Olson. *Knowledge Organization*, 43(5), 331-338.



LA INDIZACIÓN DE ARTÍCULOS CIENTÍFICOS CON EL SISTEMA DE INDIZACIÓN AUTOMÁTICA SISA COMPARADA CON LA INDIZACIÓN EN LAS BASES DE DATOS AGRICOLA, WOS Y SCOPUS

Isidoro Gil-Leiva

Facultad de Comunicación y Documentación, Universidad de Murcia, isgil@um.es

RESUMEN Desde hace unos años la generación de documentos digitales es enorme así como su incorporación masiva a los sistemas de información y ambas realidades parecen imparables. Del mismo modo, no hay duda de que la indización es uno de los procesos fundamentales ejecutados en las unidades documentales. Aunque las primeras investigaciones en automatización de la indización se iniciaron hace décadas este asunto sigue suscitando interés. Desde entonces diferentes propuestas y metodologías han sido planteadas. SISA es un sistema de indización automática multilingüe para artículos científicos fundamentado en principios heurísticos y estadísticos regido mediante reglas basadas en dichos principios. **Objetivo.** En este contexto descrito de incremento digital constante, se persigue conocer las capacidades de SISA en la indización automática de artículos en relación a cómo lo hacen en las bases de datos Agrícola, WOS y SCOPUS. **Material y método.** Se seleccionaron al azar cien artículos publicados en diferentes años por la revista de Agricultura *Agronomy for sustainable development*, se localizó la indización asignada a los artículos en las mencionadas bases de datos, se indizaron los documentos con SISA, se compararon las diferentes indizaciones y se calcularon los índices de consistencia entre Agrícola y SISA. **Conclusiones.** Las capacidades de indización de SISA en relación a las bases de datos de referencia han sido satisfactorias, si bien se precisan algunos ajustes. SISA ha producido un número medio de descriptores por documento similares a Agrícola y Scopus, si bien, los descriptores compuestos de SISA es menor que en estas dos bases de datos. Asimismo, el 21,61% de consistencia conseguido entre SISA y Agrícola se encuentra dentro de los porcentajes en este tipo de estudios. Por último, la propuesta de una fórmula integral para la evaluación de la indización automática denominada Evaluación Robusta de la Indización (ERI) permitiría estimar de una manera sólida la viabilidad de un sistema de indización automática.

PALAVRAS-CHAVE *Indización automática, Evaluación, SISA, Bases de datos, Agrícola, WoS, Scopus, ERI, Evaluación Robusta de la Indización.*

ABSTRACT Since some years the generation of digital documents is enormous as well as its massive incorporation to the information systems and both realities seem unstoppable. Likewise, there is no doubt that indexing is one of the fundamental processes executed in documentary units. Although the first investigations in automatic indexing began decades ago this subject continues to raise interest. Since then different proposals and methodologies have been presented. SISA is a multilingual automatic indexing system for scientific articles based on heuristic and statistical principles governed by rules based on these principles. **Objective.** In this described context of constant digital increase, it is sought to know the SISA capabilities in the automatic indexing of articles in relation to how they do it in the Agrícola, WOS and SCOPUS databases. **Material and method.** One hundred articles published in different years by the journal *Agronomy for sustainable development* were randomly selected, the indexing assigned to the articles in the mentioned

databases was located, the documents were indexed with SISA, the different indexing were compared and they were calculated the consistency between Agricola and SISA. **Conclusions.** The indexing capabilities of SISA in relation to the reference databases have been satisfactory, although some adjustments are needed. SISA has produced a mean number of descriptors per document similar to Agricola and Scopus, although the composite descriptors of SISA are smaller than in these two databases. Also, the 21.61% consistency achieved between SISA and Agricola is within the percentages in this type of studies. Finally, the proposal of a comprehensive formula for the evaluation of the automatic indexing called Robust Indexing Evaluation (RIE) would allow a solid estimation of the viability of an automatic indexing system.

KEYWORDS *Automatic indexing, Evaluation, SISA, databases, Agricola, WoS, Scopus, RIE, Robust Indexing Evaluation.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

1. INTRODUCCIÓN

La indización ha sido extensamente estudiada. Algunos trabajos valiosos sobre la teoría y práctica de la indización son Frohmann (1990), Lancaster (1991), Farrow (1991), Fugmann, (1993), Hjørland, B. (1997), Anderson y Perez-Carballo (2001) o Mai (2000), entre otros. La norma ISO 5963-1985 define la indización como “la acción de describir o identificar un documento en relación con su contenido.” Esto se puede completar señalando que en ocasiones los conceptos resultantes se normalizan y controlan por medio de un vocabulario controlado, de lo contrario sería un indización en lenguaje natural, y por otro lado, cabe añadir que la indización también se ejecuta –consciente o inconscientemente- sobre las necesidades de información de los usuarios para convertirla mediante lenguaje natural o controlado en una ecuación de búsqueda, de ahí que sea un proceso esencial para el almacenamiento de los documentos y puede serlo en la recuperación de información si su resultado (palabras clave o descriptors) es usado posteriormente en la recuperación. Las primeras propuestas de indización automática se fundamentaron en la Ley de Zipf a las que prosiguieron otras basadas en cálculos estadísticos para conseguir términos de indización como la frecuencia inversa del documento (Sparck Jones, 1972), las aportaciones de Gerard Salton con su modelo de discriminación del término o modelo espacio vectorial (Salton and Yang, 1973; Salton, Wong and Yang, 1975) o entre otros, Deerwester et al. (1988) con su indización semántica latente. Asimismo, prototipos de indización automática se fueron desarrollando en menor o mayor medida en grandes centros de información y documentación, como por ejemplo, en el *STN Internacional* de Karlsruhe, de Alemania; en la Biblioteca Nacional de Medicina de los Estados Unidos, en el Centro de Información Aeroespacial de la NASA; en la Biblioteca Nacional de Agricultura también de los Estados Unidos; o más recientemente en el *Data Archive* del Reino Unido. (Gil-Leiva, 2017, p. 140).

Como se ha señalado, desde el inicio de estas investigaciones sobre la automatización de la indización a finales de la década de 1950 se han realizado numerosas y variadas propuestas para acometer el proceso intelectual que supone la indización. La terminología utilizada en la literatura para referirse al proceso de la automatización de la indización es variada, pudiendo encontrar estas denominaciones, entre otras: «Automated assisted indexing», «Automated indexing», «Automated support to indexing», «Automatic support to indexing», «Computer aided indexing», «Computer assistance in indexing», si bien la más utilizada es «Automatic indexing». La definición de la automatización de la indización se debe acometer desde una triple perspectiva: a) Programas informáticos que asisten en el proceso de almacenamiento de los términos de indización, una vez obtenidos de modo intelectual (Indización

Asistida por Ordenador Durante el Almacenamiento); b) Sistemas que analizan los documentos de modo automático, pero los términos de indización propuestos los valida y edita -si es necesario- un profesional (Indización Semiautomática); y c) Programas sin ningún tipo de validación, es decir, los términos propuestos se almacenan directamente como descriptores de dicho documento (Indización Automática), (Gil Leiva, 2008).

Desde hace unos años la generación de documentos digitales es enorme así como su incorporación masiva a determinadas unidades documentales como por ejemplo, los libros electrónicos a las bibliotecas académicas, o los artículos científicos a las bases de datos (ver Tabla 1 como ejemplo ilustrativo). De ahí que más frecuentemente se recurra a sistemas de indización automática o semiautomática que ayuden a ejecutar esta tarea. En este contexto de crecimiento digital imparable y la necesidad de disponer de sistemas automáticos o semiautomáticos para llevar a cabo procesos técnicos documentales o que asistan en ellos, se enclava este trabajo. Concretamente se persigue conocer las capacidades de SISA en la indización automática de artículos en relación a cómo lo hacen en las bases de datos Agrícola, Web Of Science (en adelante WoS) y SCOPUS, bases de datos de referencia y prestigio de ámbito internacional. Así pues, para cada artículo indizado por SISA se dispondrá de indizaciones previas producidas por estas unidades, lo que nos permitirá responder a cuestiones como ¿cuál es el número medio de términos asignados a cada documento en cada sistema de información? ¿Cuáles son las características de los términos? o ¿Cuál es la semejanza de la indización de SISA con respecto a la indización en Agrícola?, entre otros aspectos.

2. MATERIAL Y MÉTODO

2.1 MATERIAL

SISA es un sistema de indización automática que ha sido desarrollado en JAVA, maneja diferentes librerías para extraer la información de los documentos en PDF, txt ó XML y también puede emplear un vocabulario controlado en formato txt o SKOS. SISA está diseñado para la indización de artículos de revista e implementado en plataforma web. Procesa documentos en español, portugués e inglés usando para ello, listas de palabras vacías (artículos, preposiciones, etc.) y vocabularios controlados en estos idiomas. Hace uso del stemming para contabilizar la aparición de raíces y no contar como diferentes los términos *economía*, *económico*, *economías* o *económicamente*, por ejemplo. Para la asignación de los descriptores usa un conjunto de reglas fundamentas en métodos heurísticos (posicionamiento) y estadísticos (frecuencia). En el Anexo 1 se muestran algunas de estas reglas. Las tareas sucesivas para la indización de un artículo con SISA son las siguientes: etiquetar los artículos, procesarlos (aplicar el stemming, calcular la frecuencia de aparición de los términos en los documentos y en la colección, calcular el TFIDF y registrar el lugar en el que aparecen palabras y frases y seguidamente se indizan de acuerdo a las reglas establecidas. Interrelacionado con las utilidades de recuperación disponibles, en la actualidad se está incorporando a SISA el módulo de evaluación mediante la recuperación para hallar índices de exhaustividad, precisión y f-measure. Por otro lado, se ha usado Cascading Style Sheets para el diseño de la aplicación y MySQL como base de datos para guardar las fuentes, los documentos y los resultados de la indización. Finalmente, señalar que SISA está instalado en un servidor Proliant ML310E con 32GB RAM y con un sistema operativo CentOS 7.0. (Gil Leiva, 2008; Rocha Souza y Gil-Leiva, 2016 y Gil-Leiva, 2017).

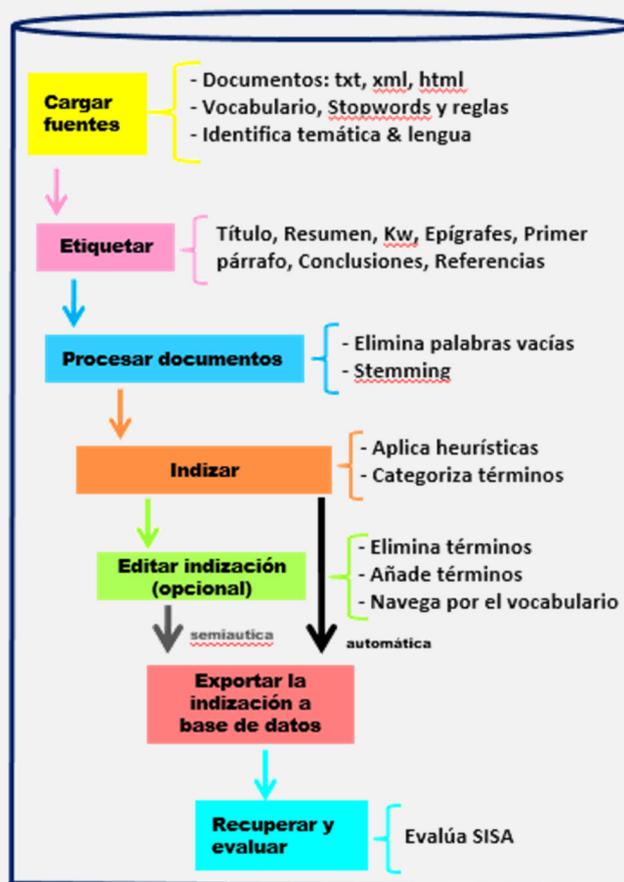


Figura 1. Esquema de los procesos en SISA.

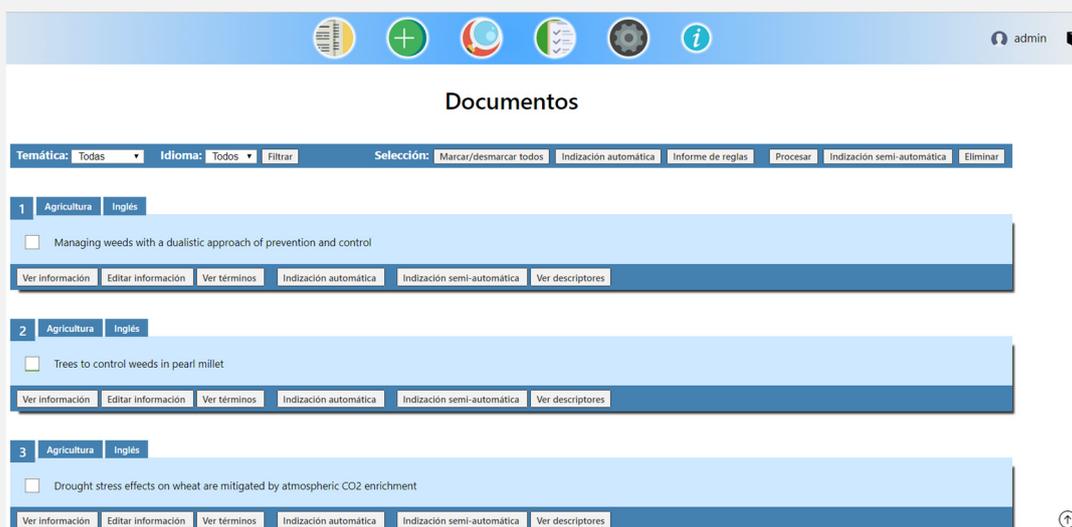


Figura 2. Pantalla principal de SISA.

En la siguiente figura aparecen las utilidades de los iconos que aparecen en cada momento en la parte superior de la interfaz de SISA.

	Retorno a la página principal
	Carga documentos
	Búsqueda en la base de datos una vez indizados los documentos
	Evaluación del sistema
	Configuración (sistema, reglas y usuarios)

Figura 3. Utilidades de los iconos de SISA

Agricola, la base de datos de la Biblioteca Nacional de Agricultura de los Estados Unidos contiene más de 5,2 millones de registros entre artículos, capítulos de libro, tesis, etc. Scopus más de 67 millones de artículos con cerca de 8 millones de “conference proceedings”, además de libros, revisiones o patentes. Y WoS supera los 100 millones de artículos procedentes de treinta y tres mil revistas, y cuenta además con más de 7,4 millones de “conference proceedings”.

Tabla 1. Artículos incorporados a las bases de datos por años

	2013	2014	2015	2016
Agricola*	128.804	185.006	332.680	430.126
Scopus**	1.831.731	1.915.250	1.922.662	1.877.243
WoS**	1.396.330	1.435.393	1.478.959	1.511.814

* Proporcionados por la responsable de indización de la base de datos

** Obtenidos por medio de sendas búsquedas en estas bases de datos

Como se observa en la Tabla 1 solo el número de artículos incorporados cada año a estas bases de datos es enorme. En Scopus la media de registros añadidos anualmente en el período 2013-2016 ha sido de 2.800.000 documentos y de 2.400.000 en WoS. Los artículos suponen aproximadamente el 65% del total de registros incorporados cada año en Scopus, mientras que en WoS representa el 60%. Los documentos procesados cada año en estas bases de datos hace casi inviable una indización manual de los mismos. Tanto Scopus como WoS intentan en un plazo máximo de cuatro o cinco semanas convertir los artículos publicados en registros accesibles de sus bases de datos obtenidos previamente en su mayoría vía electrónica en formatos XML, PDF o descargados de los sitios webs de las propias revistas. Desde 2012 la Biblioteca Nacional de Agricultura ha automatizado la indización en Agricola por medio de un software de indización semiautomática, denominado Luxid de la empresa Temis, de ahí el fuerte incremento en el número de artículos añadidos desde 2014 en adelante.

2.2 METODOLOGÍA

Para llevar a cabo este experimento se ha manejado SISA para obtener la indización automática de un corpus de cien artículos de Agricultura en inglés publicados en la revista *Agronomy for sustainable development* entre 2007 y 2016. SISA ha usado un vocabulario controlado compuesto por casi 128.000 términos, de los cuales más de 67.000 son descriptores y 60.000 son no descriptores. Estos términos

han sido extraídos del Tesouro de la Biblioteca Nacional de Agricultura disponible en la web de la institución. Asimismo SISA ha empleado un fichero de palabras vacías en inglés con casi 600 palabras.

Las principales tareas para la realización de este ensayo han sido las siguientes:

- Extraer los descriptores y no descriptores del Tesouro de la Biblioteca Nacional de Medicina, edición 2017 y adaptarlo al formato de SISA.
- Seleccionar al azar cien artículos de la revista *Agronomy for sustainable development* publicados en los años 2007, 2008, 2009, 2014 y 2015.
- Localizar la indización asignada a los cien artículos en Agricola, WoS y SCOPUS
- Indizar los 100 artículos con SISA
- Analizar y comparar las tres indizaciones
- Hallar los índices de consistencia entre la indización de Agricola y la indización de SISA, ya que en los dos sistemas de indización se ha usado el mismo vocabulario controlado.

Para valorar o evaluar la indización o la indización automática se viene recurriendo de manera mayoritaria al cálculo de los índices de exhaustividad y precisión en la recuperación, como hicimos por ejemplo en Rocha Souza y Gil-Leiva (2016) o Gil-Leiva (2017); o bien, al índice de consistencia entre indizaciones, como llevamos a cabo en (Gil-Leiva, 2001; 2002). Aquí de nuevo recurrimos a la consistencia para comparar la indización de Agricola con la de SISA, ya que se ha configurado y nos hemos servido del mismo vocabulario controlado empleado en Agricola. No calculamos los índices de consistencia entre SISA y WoS o Scopus porque no disponemos de los vocabularios controlados de estas bases de datos. Confrontar la indización de dos bases de datos que usan vocabularios controlados diferentes es desaconsejable para un inexperto en la materia, en este caso, en Agricultura, ya que un mismo concepto puede estar representado por descriptores diferentes en cada base de datos. Nos podemos encontrar que SISA proponga a un documento los descriptores *no-tillage* y Scopus *zero tillage*, o SISA *soil temperature* y Scopus *surface temperature*, pero en estos casos parece que estemos ante los mismos conceptos, por tanto, hay que tener un conocimiento profundo del ámbito terminológico para llevar a cabo comparaciones de este tipo.

Hooper (1965) introdujo una fórmula para hallar la consistencia entre dos indizaciones, y posteriormente, Rolling (1981) introdujo una variante de aquella. Estas dos fórmulas han sido extensamente usadas en numeros experimentos desde entonces. Nosotros venimos sirviéndonos de la siguiente variante de la fórmula de Hooper:

$$C_i = \frac{T_{co}}{(A+B) - T_{co}}$$

en donde,

T_{co} = Número de términos comunes en las dos indizaciones

A= Número de términos usados en la indización A

B= Número de términos empleados en la indización B

La aplicación de esta fórmula se puede realizar de manera “rígida” o “relajada”. El Anexo 2 proporciona la indización de Agricola y SISA para el documento 83. La coincidencia del descriptor *biodiversity*

sumaría 1, y la coincidencia del término *food*, sumaría en cambio 0,5, ya que solamente existe coincidencia en una parte del descriptor. A esto llamamos una aplicación “relajada” y si se aplicara una comparación “rígida”, en este caso de *food*, sumaría 0. En el experimento que nos ocupa, hemos aplicado la fórmula de manera “relajada”.

3. RESULTADOS Y DISCUSIÓN

3.1 LA INDIZACIÓN DE AGRICOLA, SCOPUS, WOS Y SISA

Tabla 2. Total de descriptores asignados a los 100 artículos

	nº total de descriptores asignados	nº medio de descriptores por documento	nº de descriptores simples	%	nº de descriptores compuestos	%
Agricola	1569	15,6	744	47,4	825	52,5
Scopus	1416	14,2	608	42,9	808	57,0
SISA	1446	14,4	891	61,6	555	38,3
WoS	801	8,01	450	56,1	351	43,8

El tiempo medio empleado por SISA para la indización de un documento es de diecisiete segundos aproximadamente, si bien en el preprocesamiento previo (localización del artículo en PDF, conversión a txt, etiquetado del texto que en esta ocasión se ha realizado de manera integral) se consumen de media unos siete minutos aproximadamente.

De acuerdo a la Tabla 2 en Agricola, Scopus y SISA parece existir cierta homogeneidad en el número de descriptores asignados por documento. En WoS el número es sensiblemente menor (Anexo 3), y aunque el corpus de este ensayo es pequeño parece observarse que no suelen asignar como descriptores conceptos ya incluidos en las palabras clave aportadas por los autores de los artículos, por tanto, en WoS parecen realizar una indización a partir de las palabras clave (Anexo 4), de ahí quizás el menor número de descriptores asignados, ya que WoS (al igual que Scopus) disponen de los campos “Palabras clave de autor”. En cambio, en Agricola y en Scopus se percibe que las palabras clave de los autores sí que terminan convirtiéndose en descriptores de dichos artículos. Las heurísticas de posicionamiento SISA conceden un peso destacado a las palabras clave de los autores, de ahí que SISA proponga descriptores que también aparecen como palabras clave.

En cuanto a los descriptores parece existir cierta similitud entre Agricola y Scopus porque tienen un número semejante de descriptores simples y compuestos; mientras que parece alinearse SISA con WoS, si bien, SISA presenta claramente el menor número de descriptores compuestos.

En la indización de los cien artículos por parte de Agricola hemos encontrado dos documentos con cuatro y cinco descriptores respectivamente, en cambio con SISA el número menor de descriptores asignados ha sido de siete, concretamente a dos documentos. En cambio, tanto Agricola como SISA a la mayor parte de los documentos se han asignado entre diez y diecinueve descriptores; y curioamente al documento 36 es al que más descriptores han asignados ambos sistemas, Agricola cuarenta y tres y SISA ha asignado cuarenta y cinco. El artículo 36 cuenta con 28 páginas, casi más del doble que la media de páginas de los artículos del corpus, y aunque solamente se trata de un caso coincidente en Agricola y SISA, (Anexo 5), ¿será que el tamaño de los documentos y sus características interviene

directamente en el resultado de la indización automática o semiautomática?, porque en la indización manual, no parece existir una correlación directa entre el tamaño y el número de descriptores asignados Gil-Leiva y Rodríguez Muñoz (1997, p. 162).

3.2 LA CONSISTENCIA EN LA INDIZACIÓN ENTRE AGRICOLA Y SISA

La consistencia media entre Agrícola y SISA ha sido del 21,61%. El índice de consistencia más bajo obtenido ha sido en el artículo 83 con el 6,98%, mientras que el más alto se ha dado en el artículo 2 con un 44,44% (Anexo 2). En este tipo de ensayos la consistencia media oscila entre el 20 y 60%, por tanto, aunque en la parte más baja, se encuentra dentro de esta horquilla.

A la hora de la evaluación de la indización, un elemento a considerar es la corrección, que implica la a la vez la ausencia de errores de inclusión (no asignar un descriptor incorrecto) y de omisión (no dejar de asignar un descriptor que corresponda). En la indización de SISA en este ensayo se han detectado tanto errores de omisión como de inclusión. Errores de inclusión como por ejemplo proponer en un mismo documento “grain yield”, “grains” y “yields”; o bien en otro documento “habitat preferences” y “habitats”. De igual modo, se ha detectado un error al adaptar el tesoro de la Biblioteca Nacional de Agricultura al formato de SISA que ha provocado que más de mil quinientos descriptores erróneos aparezcan en el vocabulario controlado configurado y empleado, lo que ha podido interferir en los resultados.

Desde que estamos desarrollando SISA, ya se ha señalado que hemos tenido oportunidad de aplicar formas para valorar o evaluar su indización. En Gil-Leiva (2017, p. 150) tomando como base el oportuno e interesante trabajo de Golub et al. (2016) señalamos la necesidad de trabajar en la búsqueda de fórmulas robustas para evaluar la indización automática. Ya hemos dicho que para la evaluación se recurre a la exhaustividad y precisión en la recuperación y a la consistencia entre indizaciones, pero su aplicación de manera independiente se nos antoja insuficiente, lo que nos lleva a presentar la siguiente manera integral para evaluar la indización automática. Aunque será expuesta con más detalle y espacio en un futuro próximo, cabe decir ahora que se fundamenta en tres elementos: la f-measure (que es la media armónica que combina la exhaustividad y precisión), la consistencia (aplicación de la fórmula de la consistencia a dos indizaciones) y la valoración de expertos (juicio de un experto sobre una indización). Esta propuesta la hemos denominado Evaluación Robusta de la Indización (ERI) y se formula de la siguiente manera:

$$ERI = (f\text{-measure} + \text{consistencia} + \text{valoración de expertos}) / 3$$

en donde,

f-measure = media entre los índices exhaustividad y precisión

consistencia = índice de consistencia medio

valoración de expertos = índice medio de la valoración de los expertos

Para la aplicación de ERI se precisa un corpus documental (compuesto por n documentos), la indización perfecta o ideal del corpus (gold indexing), un experto indizador en la temática a evaluar y un conjunto de n necesidades de información para calcular la f-measure. Cada elemento que conforma ERI aporta un valor numérico de 0 a 1, por tanto, cuanto más se acerque ERI a 1 mejor funciona u opera el sistema evaluado.

4. CONCLUSIONES Y TRABAJOS FUTUROS

Los resultados logrados nos han permitido aproximarnos a las capacidades de SISA frente a otros sistemas de indización relevantes. A modo de resumen, y a falta de análisis más profundos de los aquí mostrados y próximas mejoras en la herramienta, parece que la indización de SISA puede llegar a ser tan válida como la de Agricola, WoS y Scopus. La configuración de SISA para este ensayo ha producido de media un número de descriptores por documento similares a Agricola y Scopus, si bien quizás habría que ajustar su reglas para intentar alcanzar una media de 8-12 descriptores por documento. Asimismo, el número de descriptores compuestos de SISA es sensiblemente menor que en estas dos bases de datos, aspecto que requiere un estudio detallado. El 21,61% de consistencia conseguido entre SISA y Agricola se encuentra dentro de los porcentajes en este tipo de estudios, si bien en la parte baja de la horquilla que oscila entre el 20 y 60% aproximadamente. Por último, se ha propuesto una fórmula integral para la evaluación de la indización automática denominada ERI, Evaluación Robusta de la Indización que podrá servir para estimar de una manera sólida la viabilidad de un sistema de indización automática.

Esta investigación ha abierto líneas de trabajo para el futuro como la revisión del algoritmo de SISA para su ajuste, un análisis pormenorizado descriptor a descriptor en relación al contenido de los artículos procesados para seguir mejorando nuestra herramienta, así como ampliar este estudio a otras bases de datos internacionales. Y por último, profundizar en la concepción teórica y práctica de ERI, que aquí solamente ha sido apuntada.

REFERENCIAS BIBLIOGRÁFICAS

Anderson, James D. & Perez-Carballo, José. (2001). The nature of indexing: How humans and machines analyze messages and texts for retrieval. Part I: Research and the nature of human indexing. *Information Processing & Management*, 37(2), 231-54.

Farrow, John F. (1991), A cognitive process model of document indexing. *Journal of Documentation*, 47(2), 149-166.

Frohmann, Bernd. (1990). Rules of indexing: a critique of mentalism in information retrieval theory. *Journal of Documentation*, 46(2), 81-101.

Fugmann, Robert. (1993). *Subject analysis and indexing: Theoretical foundation and practical advice*. Frankfurt/Main: Indeks Verlag.

Gil-Leiva, Isidoro & Rodríguez Muñoz, José Vicente (1997). Análisis de los descriptores de diferentes áreas de conocimiento indizadas en bases de datos del CSIC. Aplicación a la indización automática. *Revista Española de Documentación Científica*, 20, 150-60.

Gil-Leiva, Isidoro. (2001). Consistencia en la asignación de materias en Bibliotecas Públicas del Estado. *Boletín de la Asociación Andaluza de Bibliotecarios*, 63, 69-86.

Gil-Leiva, Isidoro. (2002). Consistencia en la indización de documentos entre indizadores noveles. *Anales de Documentación*, 5, 99-111.

Gil-Leiva, Isidoro. (2008). *Manual de indización. Teoría y práctica*. Gijón: Trea.

Gil-Leiva, Isidoro (2017). SISA: Automatic indexing system for scientific articles. Experiments with location heuristics rules versus TF-IDF rules. *Knowledge Organization*, 43(3), 139-162.

Golub, Koraljka, Soergel, Dagobert, Buchanan, George, Tudhope, Douglas, Likke, Marianne and Hiom, Debra. (2016). A framework for evaluating automatic indexing or classification in the context of retrieval. *Journal of the Association for Information Science and Technology*, 67(1), 3-16.

Hjørland, Biger. (1997). *Information seeking and subject representation: An activity-theoretical approach to information science*. Westport, CT: Greenwood Press.

Hooper, Robert S. (1965). *Indexer Consistency Tests: Origin, Measurement, Results, and Utilization*. Bethesda: IBM Corporation.

ISO 5963:1985 : *Documentation -- Methods for examining documents, determining their subjects, and selecting indexing terms*. Geneva: ISO.

Lancaster, Frederick W. (1991). *Indexing and abstracting in theory and practice*. Champaign: University of Illinois.

Mai, Jens-Erik. (2000). Deconstructing the Indexing Process. *Advances in Librarianship*, 23, 269-298.

Rolling, Loll N. (1981). Indexing Consistency, Quality and Efficiency. *Information Processing & Management*, 17, 69-76.

Souza, Renato Rocha & Gil-Leiva, Isidoro. (2016). Automatic Indexing of Scientific Texts: A Methodological Comparison. In Chaves Guimarães, José Augusto, Oliveira Milani, Suelen & Dodebei, Vera. *Knowledge Organization for a Sustainable World: Challenges and Perspectives for Cultural, Scientific, and Technological Sharing in a Connected Society: Proceedings of the Fourteenth International ISKO Conference 27-29 September 2016, Rio de Janeiro, Brazil*, Advances in Knowledge Organization, 2016 (pp. 243-250). Würzburg: Ergon Verlag.

ANEXO 1: EJEMPLO DE ALGUNAS REGLAS USADAS POR SISA EN EL ENSAYO.

	ID	Título	Resumen	Palabras clave	Epígrafe	Primer párrafo	Conclusiones	Referencias	Frecuencia documento (DF)	TF-IDF	Voc. Controlado
<input type="checkbox"/>	R1	1	1	1	1	1	1	1	-	-	S
<input type="checkbox"/>	R2	1	1	1	1	1	1	-	-	-	S
<input type="checkbox"/>	R3	1	1	1	1	1	-	1	-	-	S
<input type="checkbox"/>	R16	1	1	-	1	1	-	1	-	-	S
<input type="checkbox"/>	R78	1	-	-	-	-	1	-	-	-	S
<input type="checkbox"/>	R79	-	-	1	-	-	-	-	-	-	S
<input type="checkbox"/>	R80	1	-	-	-	-	-	-	-	0.025	S

ANEXO 2 : ARTÍCULOS CON EL ÍNDICE DE CONSISTENCIA MENOR Y MAYOR.

	Descriptorios en Agrícola	Descriptorios en SISA	Consistencia
<p>Artículo 83:</p> <p>Título: Using our agrobiodiversity: plant- based solutions to feed the world</p>	<ol style="list-style-type: none"> 1. agricultural land 2. arable soils 3. biodiversity 4. food availability 5. gene pool 6. people 7. politics 8. population growth 9. risk 10. sustainable agriculture 	<ol style="list-style-type: none"> 1. andean crops 2. biodiversity 3. breeding 4. climate change 5. feeds 6. food production 7. foods 8. nutrition 9. planting 10. solutes 11. utilities 12. yields 13. zoning 	6,98%
<p>Artículo 2:</p> <p>Título: Trees to control weeds in pearl millet</p>	<ol style="list-style-type: none"> 1. agroforestry 2. canopy 3. crop yield 4. <i>Faidherbia albida</i> 5. millets 6. parasitic plants 7. <i>Pennisetum glaucum</i> 8. plant growth 9. semiarid zones 10. <i>Striga hermonthica</i> 11. weed control 12. Nigeria 	<ol style="list-style-type: none"> 1. canopy 2. controllers 3. <i>Faidherbia albida</i> 4. inflorescences 5. millets 6. pearls 7. <i>Pennisetum glaucum</i> 8. plant growth 9. planting 10. <i>Striga hermonthica</i> 11. surveys 12. trees 13. weed control 14. Nigeria 	44,44%

ANEXO 3 : EJEMPLOS DE INDIZACIÓN EN WOS VERSUS SCOPUS Y SISA.

Documento	Descriptorios WoS	Descriptorios Scopus	Descriptorios SISA
25	<ol style="list-style-type: none"> 1. EUTROPHICATION 2. GROWTH 3. STRATEGIES 	<ol style="list-style-type: none"> 1. concentration (composition) 2. crop yield 3. growing season 4. irrigation 5. leaching 6. life cycle analysis 7. Lycopersicon esculentum 8. Mediterranean environment 9. nitrate 10. sustainability 11. water use efficiency 	<ol style="list-style-type: none"> 1. climate 2. concentrates 3. controllers 4. decrease in nitrate leaching 5. drainage 6. environmental impact 7. eutrophication 8. fertigation 9. greenhouses 10. hydroponics 11. nitrification 12. nutrient solutions 13. physics 14. reduction 15. solanum lycopersicum var.lycopersicum 16. solutes 17. tomatoes 18. water-use efficiency 19. yields
41	<ol style="list-style-type: none"> 1. GROWTH 2. YIELD 	<ol style="list-style-type: none"> 1. aboveground biomass 2. agronomy 3. alternative agriculture 4. Citrus 5. crop yield 6. environmental impact 7. fertilizer application 8. growth rate 9. industrial waste 10. leaf area index 11. Mediterranean environment 12. nutrient cycling 13. soil fertility 14. soil organic matter 15. Triticum aestivum 16. Triticum turgidum subsp. durum 17. waste disposal 18. wheat 	<ol style="list-style-type: none"> 1. crops 2. durum wheat 3. industrialization 4. mineral fertilizers 5. mineralization 6. oranges 7. organic fertilizers 8. organic soil fertility 9. organic soils 10. organisms 11. reduction 12. wastes 13. yields
78	<ol style="list-style-type: none"> 1. FRANCE 2. SYSTEMS 	<ol style="list-style-type: none"> 1. agricultural intensification 2. agricultural modeling 3. agricultural practice 4. agronomy 5. alternative agriculture 6. automation 7. biophysics 8. climate change 9. environmental factor 10. farm 11. farming system 12. grass 13. grassland 14. grazing 15. management practice 16. model test 17. prediction 18. remote sensing 19. surface area 20. France 	<ol style="list-style-type: none"> 1. administrative management 2. climate change 3. design 4. farming systems 5. grasses 6. grasslands 7. models 8. range management 9. remote sensing

ANEXO 4 : DESCRIPTORES PROPUESTOS EN LAS BASES DE DATOS PARA EL DOCUMENTO 14.

Título: High efficacy of extracts of Cameroon plants against tomato late blight disease.

Agricola	WoS	Scopus	SISA	Palabras clave del autor
1. agar 2. biodegradability 3. biopesticides 4. Chrysopogon zizanioides 5. Cupressus 6. disease control 7. disease severity 8. ecosystems 9. food contamination 10. fungi 11. fungicide resistance 12. fungicides 13. germination 14. greenhouse experimentation 15. greenhouses 16. indigenous species 17. pathogens 18. pests 19. Phytophthora infestans 20. plant diseases and disorders 21. plant protection 22. sporangia 23. tomatoes 24. toxic substances 25. toxicity 26. Cameroon	1. field 2. pathogenicity 3. phytophthora-infestans 4. potato 5. resistance	1. disease control 2. disease severity 3. experimental study 4. fungal disease 5. fungicide 6. fungus 7. germination 8. inhibition 9. plant extract 10. toxic substance 11. Africa 12. Cameroon 13. Sub-Saharan Africa 14. West Africa 15. Cupressus 16. Cupressus benthamii 17. Fungi 18. Lycopersicon esculentum 19. Phytophthora infestans 20. Vetiveria 21. Vetiveria zizanioides	1. biopesticides 2. controllers 3. disease control 4. extraction 5. fungicides 6. greenhouses 7. late blight disease 8. lates 9. pathogenicity 10. phytophthora infestans 11. plant extracts 12. planting 13. sporangial germination 14. tomatoes 15. Cameroon	1. antigungal activity 2. biopesticide 3. disease suppression 4. late blight 5. Phytophthora infestans 6. plant extracts 7. tomato

Anexo 5: Relación entre el nº de descriptores y el nº de documentos con dicho nº de descriptores.

	Agricola	SISA
Con 4 descriptores	1	0
Con 5 descriptores	1	0
6	1	0
7	1	2
8	4	2
9	3	4
10	6	8
11	5	7
12	11	8
13	4	15
14	8	8
15	5	11
16	10	10
17	9	7
18	6	5
19	5	5
20	4	3
21	3	0
22	2	1
23	1	0
24	5	2
25	3	1
26	1	0
43	1	0
45	0	1
Total	100	100



ORGANIZAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO EM CONTEXTO EMPRESARIAL: A APLICAÇÃO DA ANÁLISE POR FACETAS

Olívia Pestana

Faculdade de Letras da Universidade do Porto, 0000-0002-5485-3143, opestana@letras.up.pt

RESUMO A análise por facetas em contextos empresariais e o seu uso pelos utilizadores de informação em toda a organização pode ser de grande importância para o desenvolvimento de sistemas de informação, mas é necessário entender como esses processos estão sendo desenvolvidos e se há diferentes abordagens a esses conceitos e práticas em contextos diferentes do domínio bibliográfico. Através de uma abordagem exploratória, realizou-se a revisão da literatura indexada nas bases de dados Library and Information Science and Technology Abstracts, Scopus e, ainda, ISI Web of Science. Este trabalho apresenta uma comparação dos artigos analisados de acordo com os seguintes critérios: fundamentos teóricos utilizados nos trabalhos, definições dos conceitos de análise por facetas e de classificação facetada, processos de formulação de facetas, princípios para o arranjo e ordenação de facetas e, também, as soluções tecnológicas para a criação e utilização de classificações facetadas nos contextos analisados, como os produtos financeiros, os documentos eletrónicos em pequenas, médias e microempresas, os documentos de engenharia e, ainda, de processos em indústrias de manufactura. A literatura analisada no presente trabalho evidencia o potencial da análise por facetas, apesar das diferentes abordagens e sustentação teórica que os artigos expressam.

PALAVRAS-CHAVE *análise por facetas, classificações facetadas, organização da informação e do conhecimento*

ABSTRACT Facet analysis in business contexts and its use by information users throughout the organization may be of great importance for the development of information systems, but it is necessary to understand how these processes are being developed and if there are different approaches to this concepts and practices in contexts other than the bibliographical domain. An exploratory approach was carried out to review the literature indexed in the Library and Information Science and Technology Abstracts, Scopus and ISI Web of Science. This paper presents a comparison of the articles analysed according to the following criteria: theoretical basis used in the works, definitions of facet analysis and faceted classification, facet formulation processes, principles for the arrangement and ordering of facets, and also the technological solutions for the creation and use of faceted classifications in the analysed contexts, such as financial products, electronic documents in small, medium and micro enterprises, as well as engineering documents, and processes in manufacturing industries. The literature analysed in the present work highlights the potential of facet analysis, despite the different approaches and theoretical support expressed in the articles.

KEYWORDS *facet analysis, faceted classifications, information and knowledge organization*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A análise por facetas e a utilização de esquemas facetados não são técnicas novas para o campo científico da ciência da informação e para alguns serviços tradicionais de informação como as bibliotecas. Mas as crescentes aplicações tecnológicas e as necessidades de organização da informação e do conhecimento nas empresas promoveram o aparecimento desses termos na produção científica diretamente relacionada com a organização da informação em contexto empresarial.

Alguns autores indicam que a gestão de conteúdos no contexto das empresas constitui um desafio para os trabalhadores da informação e do conhecimento (Brocke & Simons, 2014), considerando que uma organização de conhecimento eficaz nas empresas pode ser determinante para atingir os objetivos a que a entidade se propõe. Choo (2003) identifica um ciclo contínuo de seis actividades distintas, mas relacionadas entre si, no ciclo de gestão da informação: identificação das necessidades de informação, aquisição de informação, organização e armazenamento de informação, desenvolvimento de produtos e serviços de informação, distribuição de informação, e, ainda, utilização de informação (também referidos por Davenport, 1993; McGee e Prusak, 1993). Se a informação e o conhecimento estiverem incorrectamente ou ineficientemente organizados, os processos subsequentes de gestão da informação serão comprometidos. A análise temática da informação e do conhecimento em empresas permaneceu distante dos procedimentos regulares das bibliotecas e pode considerar-se uma prática relativamente recente. No entanto, alguns trabalhos teóricos e práticos relacionados com a análise por facetas e a aplicação de classificações facetadas (tais como, entre outros, os trabalhos dos autores Ranganathan, 1962, 1967; Vickery, 1960; Broughton, 2004, 2006; La Barre, 2010) constituem a base dos estudos atuais que exploram a análise por facetas no âmbito da organização de informação nas empresas. Porém, isso não significa que as interpretações, os processos e os resultados sejam análogos aos referidos e utilizados em contextos tradicionais de organização da informação e do conhecimento.

METODOLOGIA

Considerando a visão integrada, sistémica e dinâmica da informação, expressa no paradigma pós-custódia, informacional e científico da ciência da informação, que transmite uma visão holística da informação nas organizações, ou seja, uma visão integrada que não separa artificialmente a informação tradicionalmente atribuída ao arquivo (e o *records management*), a informação tradicionalmente atribuída às bibliotecas ou, ainda, o suporte digital, este trabalho apresenta uma análise da literatura científica relacionada com diferentes contextos empresariais, no que diz respeito à análise por facetas e à utilização de esquemas facetados nesses contextos.

A análise por facetas em contextos empresariais e o seu uso pelos utilizadores de informação em toda a organização pode ser de grande importância para o desenvolvimento de sistemas de informação, mas é necessário entender como esses processos estão sendo desenvolvidos e se há diferentes abordagens a esses conceitos e práticas em contextos diferentes do domínio bibliográfico.

Através de uma abordagem exploratória, realizou-se a revisão da literatura indexada nas bases de dados Library and Information Science and Technology Abstracts, Scopus e, ainda, ISI Web of Science. Estas bases de dados são utilizadas não apenas para a pesquisa de literatura, mas também como fonte para a avaliação da qualidade da produção científica de unidades de investigação nacionais e internacionais, o que pode ser um indicador do interesse do mundo empresarial no desenvolvimento de trabalhos de

investigação na área analisada. Os dados coletados referem-se a documentos indexados de 2000 até 2016 e esses limites foram estabelecidos no sentido de se recuperar artigos relacionados com soluções tecnológicas mais recentes ou passíveis de estarem, ainda, atualizadas. Por outro lado, foi considerado como critério de inclusão o artigo que ilustrasse aplicações concretas e em contexto real do objeto de estudo.

Para encontrar pontos de similaridade e diferença entre os artigos resultantes da pesquisa de literatura, este trabalho apresenta uma comparação dos artigos analisados de acordo com os seguintes critérios: base teórica utilizada nos trabalhos, definições dos conceitos de análise por facetas e de classificação facetada, processos de formulação de facetas, princípios para o arranjo e ordenação de facetas e, também, as soluções tecnológicas para a criação e utilização de classificações facetadas nos contextos analisados. Foram analisados diferentes contextos empresariais como os sistemas de classificação de produtos financeiros, de documentos eletrônicos em pequenas, médias e microempresas, de documentos de engenharia e, ainda, de processos em indústrias de manufatura.

RESULTADOS

As pesquisas realizadas e a aplicação do critério de inclusão resultaram na recuperação de um reduzido volume de artigos, sendo, no entanto, de realçar a sua importância e contributo para a compreensão do tema em causa. Deste modo, foram selecionados seis artigos, retratando quatro tipos diferentes de contextos de trabalho. A utilização do termo "faceta" é comum em muitos artigos, os quais foram excluídos por não se referirem ao desenvolvimento de análise facetada ou à construção de uma classificação deste cariz. Por outro lado, é visivelmente crescente o relato de experiências em contexto laboratorial e relativo a diminutos conjuntos de matérias quanto ao desenvolvimento da análise por facetas, com vista à representação de conteúdos no contexto dos mais diversos tipos de negócio.

Os artigos selecionados para este estudo referem-se, como já mencionado, a quatro contextos diferentes, sendo que a apresentação do trabalho realizado em dois dos casos se divide em dois artigos, cada caso. Apresentando os artigos por ordem de publicação, são de realçar os trabalhos relativos a um estudo desenvolvido na África do Sul em 24 pequenas, médias e micro empresas referente à organização de documentos eletrônicos, tendo em consideração a informação interna e externa. Denner & Van der Walt (2004) e Van der Walt (2004) apresentam em detalhe o estudo levado a cabo em 2002, sendo necessário a leitura dos dois artigos para a compreensão de todo o processo que permitiu desenvolver a classificação facetada. De salientar que os autores são, respetivamente, investigadora e docente no Department of Information Science, Stellenbosch University, África do Sul. O estudo desenvolvido partiu da revisão da literatura sobre organização da informação, informação empresarial, inteligência competitiva e sistemas de informação, constituindo uma base para o estudo do sector em análise e para a recolha de termos com vista à criação da classificação. A revisão da literatura foi completada com o estudo das práticas de organização da informação nas 24 empresas, tendo sido recolhidos termos dos sistemas de pastas relativos aos documentos, ao e-mail e às pastas de favoritos da Internet. Este estudo foi desenvolvido através de entrevistas estruturadas, auxiliadas por um questionário, e da observação dos sistemas existentes. Para além destas fases do estudo, e no sentido de reunir conceitos e ideias para a classificação, conforme expresso pelos autores, foi, ainda, consultada a *Dewey Decimal Classification*.

Para a abordagem inicial à organização da informação e do conhecimento em contexto empresarial, os autores alicerçaram as suas reflexões em Choo (2002), Taylor (1999) e, ainda, em Rowley and Farrow

(2000), constituindo, portanto, uma base de conhecimento transferida por autores que abordam a temática em todos os seus contextos, desde os arquivos, às bibliotecas, às empresas e aos museus. Quanto aos procedimentos e técnicas relacionados com a construção da classificação de acordo com a abordagem facetada, os autores basearam a sua análise partindo das obras de autores mais diretamente relacionados com a definição terminológica neste âmbito e com a construção de classificações: Buchanan (1976) Foskett (1996), Hunter (2002), Marcella & Newton (1994), tendo citado, novamente, Rowley and Farrow (2000), apesar de ser uma obra de temática mais alargada. É, no entanto, de referir que van der Walt (2004) assume que, em alguns momentos, as práticas padronizadas na construção de classificações não foram seguidas, como, por exemplo, na atribuição de preferência à ordenação do tipo de documento em prioridade perante o assunto. Não são apresentadas pelos autores deste estudo as definições de análise facetada e de classificação facetada, muito embora se refiram em detalhe aos procedimentos desenvolvidos, ilustrando com as citações que serviram de base ao seu trabalho.

Como resultado do estudo, e por razões de ordem prática, foi criada uma classificação com apenas dois níveis de subdivisão, após as classes principais. Os conceitos foram agrupados de acordo com processos de negócio/áreas funcionais, originando as facetas e foi atribuída a notação de 0 a 9 para as classes principais, por forma a transmitir a ordem de processos de negócio, correspondendo às seguintes classes: 0 General documents, 1 External environment, 2 Management (General), 3 Finance, 4 Human resources, 5 Products & Services, 6 Marketing & Sales, 7 Customers, 8 Special collections, 9 Other subjects. Van der Walt (2004) disponibiliza no seu artigo o quadro classificatório completo, resultante do estudo desenvolvido e os dois autores (Denner & Van der Walt, 2004; Van der Walt, 2004) afirmam que foi possível conceber um esquema de classificação genérico para informação empresarial. Num conjunto de recomendações a seguir, os autores indicam que as linguagens de indexação podem ser utilizadas para a padronização do vocabulário a utilizar na concepção dos esquemas de classificação a desenvolver nas empresas.

Giess, Wild, P. J. e McMahon (2008), investigadores do Innovative Manufacturing Research Centre, University of Bath, Reino Unido, relatam a tentativa de criação de esquemas de classificação facetada aplicados a documentos de engenharia e as dificuldades em atingir esse propósito (Wild, Giess, & McMahon, 2009). A sua principal motivação em alcançar o objectivo resulta da identificada necessidade de um esquema que não só reflecta as perspectivas dos utilizadores envolvidos na sua criação, mas também seja aplicável a diferentes grupos de engenheiros com diferentes perspectivas e necessidades de informação. Após uma revisão de literatura sobre as origens das classificações e sua construção, os autores debruçam-se sobre dois estudos de caso que servem de ilustração à aplicabilidade das classificações facetadas na organização de documentos de engenharia. A revisão de literatura efetuada pelos autores assenta predominantemente na revisão de trabalhos de autores sobre o domínio que designam de Library and Information Science, sendo que, especificamente relacionados com os esquemas facetados, e, para além dos princípios estabelecidos por Ranganathan (1967), Vickery (1960) e Spiteri (1998), referem-se à importância dos autores Broughton (2004), Ellis & Vasconcelos (1999) e Rowley (1992) quanto à discussão da aplicação da análise por facetas. É, no entanto, de salientar que o trabalho de Spiteri é citado como sendo um modelo simplificado e pragmático relativamente às facetas a considerar como fundamentais. A menção à falta de orientações metodológicas quanto à implementação da análise por facetas, é, no entanto, uma constante no trabalho de Giess, Wild, P. J. e McMahon (2008). Relativamente à construção da estrutura facetada, salientam que o reconhecimento dos conceitos é visto como não sendo problemático, mas é justamente nessa parte do processo que podem surgir diferenças de perspectiva por parte de quem desenvolve a análise. Estes autores apresentam, por outro lado, duas aplicações informáticas que podem servir de apoio à construção da

classificação facetada: Flamenco e FacetMap. O sistema Flamenco, desenvolvido na Universidade de Berkeley, foi utilizado, entre outros casos na montagem e navegação de uma classificação dos vencedores de prémios Nobel e de uma classificação da coleção de um museu de belas artes. O esquema de classificação é gerado à parte e são colocados em causa os métodos de recuperação da informação em casos em que o domínio não é familiar ao pesquisador. O FacetMap, por seu lado, apresenta a limitação de que cada entidade deve aparecer apenas uma vez em cada faceta, o que, de alguma forma, contraria os princípios estabelecidos por Ranganathan.

O trabalho desenvolvido por estes autores (Wild, Giess, & McMahon, 2009) revela-se de acentuada importância, dado que apresentam uma aprofundada e sustentada perspectiva crítica relativamente a estes assuntos. Desde logo, identificam as diferentes definições de faceta, bem como a sobreposição das definições de análise por facetas e classificação facetada, colocando em confronto as definições apresentadas por Broughton (2004), em que a autora sugere classificação facetada como um método de fazer uma classificação e não uma classificação em si, o que os autores contrapõem designando análise por facetas como o processo e classificação como o produto da análise. Mas a sua reflexão vai mais longe, ao equacionar os métodos utilizados para a análise por facetas e a verdadeira aplicabilidade das classificações facetadas, tal como estão sendo desenvolvidas em vários contextos.

A criação de uma classificação para aplicação a produtos financeiros motivou a reflexão levada a cabo pelos autores Loehrlein, Lemieux & Bennett (2014), sendo os primeiros docentes na University of British Columbia, E.U.A., mais especificamente na iSchool@UBC: School of Library, Archival, and Information Studies, e o último especialista em semântica e normalização no Enterprise Data Management Council, Reino Unido. Não apresentando uma proposta de classificação, mas sim alguns exemplos de análises de categorias, os autores desenvolvem no seu estudo uma revisão das várias abordagens possíveis na classificação de produtos financeiros, concentrando-se nos de risco financeiro sistémico, recomendando o uso da análise por facetas, pela sua flexibilidade em abarcar vários pontos de vista e ser suficientemente rigorosa para facilitar inferências baseadas na estrutura hierárquica. Apresentam, então, os esquemas facetados como um instrumento para a criação de uma hierarquia genérica, com a vantagem de não ser necessário aplicar a uma taxonomia completa. De modo a caracterizar os esquemas facetados, estes autores referem-se aos autores que serviram de ponto de partida para o seu raciocínio, como La Barre (2010), Broughton (2006) e Cheti & Paradisi (2008). Consideram, então, que os esquemas facetados permitem assegurar que o princípio da divisão das categorias nas hierarquias genéricas é aplicado de forma consistente. Os esquemas não organizam os itens, mas identificam as características dos itens, organizando-as em facetas, fazendo com que as características de múltiplas facetas possam ser combinadas para representar um item. Por outro lado, salientam que a ordem das facetas facilmente poderá ser alterada, por forma a ir de encontro às prioridades. Loehrlein, Lemieux & Bennett (2014) chamam, ainda, a atenção para a necessidade de diferenciar as facetas dos outros tipos de categorias, pelo que remetem para a definição de Taylor (2006): "clearly defined, mutually exclusive, and collectively exhaustive aspects, properties, or characteristics of a class or specific subject". Complementando com a definição de Ranganathan (1962), afirmam que as facetas "can be understood to be sets of characteristics that are ordered according to a single principle".

Com o objectivo de avaliar o desempenho da sustentabilidade dos processos de fabricação, Kumaraguru, Rachuri e Lechevalier (2014) propõem uma classificação facetada dos processos em indústrias de manufactura. Estes autores, membros associados da Systems Integration Division, sector do laboratório de engenharia do National Institute of Standards and Technology (NIST), E.U.A., apresentam o ciclo

de vida dos processos, bem como algumas propostas anteriores de taxonomias, referindo-se criticamente à limitada flexibilidade das suas estruturas hierárquicas. É de referir que os autores abordam a análise por facetas e a classificação dentro do que designam por domínio da organização do conhecimento, diferenciando do que é veiculado no domínio da "Library and Information Science".

Os autores apontados elencam os tipos de taxonomias frequentemente usadas na seleção e classificação de processos de manufactura:

- listas (vocabulários controlados) - resultante de uma lista predefinida de processos;
- listas de sinónimos;
- taxonomias hierárquica - estabelecimento de relações genérico/específico, parte de, tipo de;
- taxonomias facetadas - hierarquias com uma categoria etiquetada designada de facetas;
- ontologias - taxonomias facetadas com toda a ambiguidade retirada e todos os conceitos completamente descritos.

Esta identificação e suas descrições requerem alguma reflexão, dado que não expressam com rigor as diversas linguagens passíveis de utilização, bem como as suas características. A definição de faceta expressa pelos autores cita Taylor (1992), autora referida em quase todos os artigos analisados no presente trabalho, no entanto apontam uma edição anterior à existente à data do trabalho desenvolvido, apesar da definição não ter sofrido alterações ao longo das edições da obra citada. É, também, a partir da citação desta obra que os autores se referem a Ranganathan como tendo desenvolvido a análise por facetas. Como referências para a elaboração dos processos de formulação de facetas e sua ordenação, não foram referidos autores em que tivessem alicerçado a fundamentação teórica. Na construção da classificação facetada, incluindo o estabelecimento das relações entre os elementos, e para a visualização da taxonomia foi utilizada a NIST Ontological Visualization Interface for Standards, tecnologia desenvolvida, como o nome indica, pela NIST. A modelação dos termos da classificação foi feita com base numa lista de processos reunida manualmente a partir de um conjunto de livros, normas e bases de dados. Esses processos podem ser classificados em diferentes categorias baseadas no tipo de energia, mudança de estado material, mecanismo, função, tipo de material, precedência e estado material inicial.

CONCLUSÕES

Partilhando as observações de Hjørland (2013), é raro encontrar-se discussões sobre as limitações da abordagem da análise por facetas. Mesmo no contexto das empresas, os autores são quase unânimes em destacar a força dessa abordagem na organização da informação e do conhecimento e poucos apresentam comparações com outras abordagens. No trabalho aqui apresentado, apenas Wild, P. J., Giess, M. D. & McMahon, C. A. (2009) questionam a validade dos esquemas facetados para representar idealmente o conteúdo da informação usando algumas categorias específicas do domínio e apontam as diferentes interpretações da noção de faceta, bem como alguma confusão existente entre o processo de análise por facetas e a construção de uma classificação facetada.

A literatura analisada no presente trabalho evidencia, no entanto, o potencial da análise por facetas nos processos de organização da informação e do conhecimento em vários contextos empresariais e áreas de negócio. Esses diferentes contextos originarão diferentes formas de interpretação ou criação de

facetas e essa exploração poderá enriquecer as visões conceptuais da análise por facetas e ampliar as dimensões da sua aplicabilidade e uso, todavia é, ainda, necessário o aprofundamento destes estudos para que as suas perspectivas possam consolidar-se.

Por outro lado, alguns procedimentos práticos e técnicos não foram claramente revelados, principalmente no que diz respeito à análise por facetas. Além disso, algumas referências teóricas da literatura referenciada são distintas nos diversos artigos analisados, apesar das referências fundamentais mencionadas serem coincidentes com as frequentemente citadas nas abordagens dos estudos em contexto bibliográfico.

Alguns artigos foram elaborados em colaboração entre académicos e centros parceiros de investigação, o que pode contribuir para a disseminação de abordagens mais amplas sobre a análise por facetas e para uma exploração multidisciplinar, apesar das diferentes abordagens e sustentação teórica.

O presente trabalho explorou a literatura nas bases de dados indicadas na metodologia, as quais comportam a indexação de artigos e atas de conferências das mais diversas áreas científicas. No entanto, é de realçar a eventual utilidade da exploração da literatura indexada em bases de dados especializadas de áreas científicas mais tecnológicas, bem como das áreas da gestão de empresas.

REFERÊNCIAS BIBLIOGRÁFICAS

Broughton, V. (2004). *Essential classification*. London: Facet Publishing.

Broughton, V. (2006). The need for a faceted classification as the basis of all methods of information retrieval. *Aslib Proceedings*, 58, 49-72.

Brocke, J. vom & Simons, A. (Ed.) (2014). *Enterprise content management in information systems research : foundations, methods and cases*. Berlin [etc.]: Springer.

Cheti, A. & Paradisi, F. (2008). Facet analysis in the development of a general controlled vocabulary. *Axiomathes*, 18, 223-241.

Choo, C. W. (2002). *Information management for the intelligent organization: the art of scanning the environment*. Medford: Information Today.

Choo, C. W. (2003). *Gestão de informação para a organização inteligente: a arte de explorar o meio ambiente*. Lisboa: Editorial Caminho.

Davenport, T. H. (1993). *Process Innovation: Reengineering Work Through Information Technology*. Boston, MA: Harvard Business School Press.

Denner, L. & Van der Walt, M. S. (2004). The organization of electronic information in small, medium and micro enterprises (SMMEs) in South Africa. *Knowledge Organization*, 31, 4-25.

Ellis, D. & Vasconcelos, A. (1999). Ranganathan and the Net: Using facet analysis to search and organise the World Wide Web. *ASLIB Proceedings*, 51, 3-10.

Foskett, A. C. (1996). *The subject approach to information*. London: Library Association.

- Giess, M. D, Wild, P. J. & McMahon, C. A. (2008). The generation of faceted classification schemes for use in the organization of engineering design documents. *International Journal of Information Management*, 28, 379-390. doi:10.1016/j.ijinfomgt.2007.10.001
- Hjørland, B. (2013). Facet analysis: The logical approach to knowledge organization. *Information Processing and Management*, 49, 545–557.
- Hunter, E. J. (2002). *Classification made simple*. Aldershot: Ashgate.
- Kumaraguru, S., Rachuri, S. & Lechevalier D. (2014). Faceted classification of manufacturing processes for sustainability performance evaluation. *International Journal of Advanced Manufacturing Technology*, 75, 1309-1320. doi:10.1007/s00170-014-6184-x
- La Barre, K. (2010). Facet analysis. *Annual Review of Information Science and Technology*, 44, 243–284.
- Loherlein, A. J., Lemieux, V. L. & Bennett, M. (2014). The classification of financial products. *Journal of the Association for Information Science and Technology*, 65, 263-280.
- Marcella, R & Newton R. (1994). *A new manual of classification*. Aldershot: Gower.
- McGee, J. V. & Prusak, L. (1993). *Managing Information Strategically*. New York, NY: John Wiley & Sons.
- Ranganathan, S. R. (1962). *Elements of library classification*. Bombay: Asia Publishing House.
- Ranganathan, S. R. (1967). *Prolegomena to library classification*. Bangalore: SRELS.
- Taylor, A.G. (1992). *Introduction to cataloging and classification*. Englewood: Libraries Unlimited.
- Taylor, A.G. (2006). *Introduction to cataloging and classification*. Westport: Libraries Unlimited.
- Rowley, J. (1992). *Organizing knowledge*. Aldershot: Gower.
- Rowley, J. & Farrow, J. (2000). *Organizing knowledge: an introduction to managing access to information*. Aldershot: Gower.
- Spitery, L. (1998). A simplified model for facet analysis. *Canadian Journal of Information and Library Science*, 23, 1-30.
- Van der Walt, M. S. (2004). A classification scheme for the organization of electronic documents in small, medium and micro enterprises (SMMEs). *Knowledge Organization*, 31, 26-38.
- Vickery, B. C. (1960). *Faceted classification: a guide to construction and use of special schemes*. London: ASLIB.
- Wild, P. J., Giess, M. D. & McMahon, C. A. (2009). Describing engineering documents with faceted approaches: Observations and reflections. *Journal of Documentation*, 65, 420-445. doi: 10.1108/00220410910952410



PERSPETIVAS DE HARMONIZAÇÃO DOS SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO NAS ORGANIZAÇÕES

Ana Rita Ferreira¹, Maria da Graça Simões²

¹Universidade de Coimbra, 0000-0002-0881-643X, ana.ferreira@student.fl.uc.pt

²Universidade de Coimbra, gsimoes@fl.uc.pt

ABSTRACT In contemporary society there is a multiplicity of Organizations with different functions, subject to constant changes. Immersed in a dynamic and complex network, the organization and management of information has become a strategic asset of greater value. In this context, the focus of the interdisciplinarity conferred by Information Science (CI), fits the present study, with relevance to the Knowledge Organization Systems (SOC). In order to achieve greater viability among the different SOCs in the Organizations, it is proposed to observe and reflect on the possibility of harmonization of the SOCs, in a way transversal to the Organizations to facilitate the transition of data between the different systems and guarantee, simultaneously, the semantic integration and interoperability of information, internally and externally, without constraining work processes and organizational dynamics in their different contexts. In order to understand the models, patterns and structures of contemporary organizations, as well as to know what SOCs are and how they evolved, we started with an exploratory qualitative paradigm using literature review, centered on the research of the conjugated terms: "interoperability" "Organization of Knowledge" and "Organizations", as well as "harmonization" and "Knowledge Organization Systems", in the Web of Science (WoS) database, for the period between 2007 and 2017, in order to contribute for the development and maturation of the harmonization of SOCs in Organizations. Although projects and initiatives of this nature are still insufficient, which prevented us from arriving at substantial conclusions, the trends indicate new ways of organizing information in Organizations, based on several SOCs, but without however attending to their interoperability.

KEYWORDS *Organizations; Interoperability; Harmonization; Knowledge Organization Systems.*

RESUMO Na sociedade contemporânea prolifera uma multiplicidade de Organizações com diversas funções, sujeitas a constantes alterações. Imbrincada numa rede dinâmica e complexa, a organização e gestão da informação tornou-se um ativo estratégico de valor maior. Neste contexto, pelo foco da interdisciplinaridade conferida pela Ciência da Informação (CI), enquadra-se o presente estudo, com relevância para os Sistemas de Organização do Conhecimento (SOC). Com vista a atingir uma maior viabilidade entre os diferentes SOC nas Organizações, propõe-se como objetivo observar e refletir sobre a possibilidade de harmonização dos SOC, de modo transversal às Organizações para facilitar a transição de dados entre os diferentes sistemas e garantir, em simultâneo, a integração e interoperabilidade semântica da informação, interna e externamente, sem constranger os processos de trabalho e a dinâmica organizacional nos seus diferentes contextos. Para compreender os modelos, padrões e estruturas das Organizações contemporâneas, bem como saber o que são e como evoluíram os SOC, partiu-se de um paradigma qualitativo exploratório com recurso à revisão da literatura, centrada na pesquisa dos termos conjugados: "interoperabilidade", "Sistemas de Organização do Conhecimento" e "Organizações", bem como "harmonização" e "Sistemas de Organização do Conhecimento", na base de dados Web of Science (WoS), para o período compreendido

entre 2007 e 2017, de modo a contribuir para o desenvolvimento e maturação da harmonização dos SOC nas Organizações. Embora os projetos e iniciativas desta natureza sejam ainda insuficientes, o que nos impediu de chegar a conclusões substanciais, as tendências indicam novas formas de organizar a informação nas Organizações, assentes em vários SOC, mas sem deixar de atender, porém, à sua interoperabilidade.

PALAVRAS-CHAVE *Organizações; Interoperabilidade; Harmonização; Sistemas de Organização do Conhecimento.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

De acordo com Teixeira (2005) a Organização é entendida como uma atividade desenvolvida por um conjunto de indivíduos que interagem entre si, combinando diversos recursos e estabelecendo relações recíprocas para melhor atingirem objetivos comuns, independentemente da sua tipologia, quer tenham ou não fins lucrativos. Segundo o mesmo autor, atualmente assiste-se a uma tendência significativa para o *outsourcing*, em que a Organização apenas faz o essencial do seu negócio, subcontratando o que for necessário e possível. Em cadência, presencia-se a valorização do emprego flexível, temporário, de equipas, móvel e à distância, atribuindo ênfase à democracia participativa e conferindo maior atenção às relações humanas e à ética nos negócios num contexto de responsabilidade social em que se promove simultaneamente a aprendizagem contínua e o trabalho mais qualificado.

Ao nível da tomada de decisões assiste-se a uma acentuada descentralização e achatamento das estruturas (redução de hierarquias), bem como da redução da burocracia, com a substituição dos canais verticais de comunicação por relações horizontais entre pares (Teixeira, 2005). Uma consequência direta desta nova realidade, de acordo com Teixeira (2005), referindo-se, este, às reflexões de Peter Druker, é o destaque para o papel de relevo dos trabalhos desenvolvidos na esfera da informação e do conhecimento nas novas formas de Organização. Se, por um lado, as Tecnologias da Informação permitiram que a informação suplantasse as barreiras do tempo e do espaço, por outro lado, a própria Informação adquiriu elevado valor estratégico, sendo necessário garantir que os diversos membros da cadeia hierárquica acedem, em tempo útil, a informação completa, fiável e segura.

Entende-se a Organização como um todo orgânico, em que cada parte desempenha um determinado conjunto de tarefas e funções, focadas em áreas do conhecimento específicas e operando num ambiente próprio. Cada departamento desenvolve um contexto específico, representado pelo conjunto de forças, variáveis ou instituíveis que, de algum modo, afetam o seu desempenho (Teixeira, 2005) e no seio do qual opera e se (re)organiza, internamente. Por sua vez, os diferentes departamentos comunicam e interagem entre si, trocando dados e informações que, em conjunto, completam os processos documentais.

Tendo em conta que “contextos diferentes revelam signos diferentes” (Angelos, 2013, p.12), a linguagem que se desenvolve naturalmente em cada um destes contextos “transpõe o objetivo de uma simples representação, agregando valores de significados e experiências culturais e sociais” (Angelos, 2013, p.12), podendo dificultar o processo de comunicação. Se o recetor não atribuir à informação recebida o mesmo significado que o emissor, tal resultará em falhas à transferência correta da informação e do conhecimento ao longo da estrutura (Teixeira, 2005).

Nesta medida, o papel da interoperabilidade revela-se fulcral, não como um fim, mas como um meio “para atingir um fim. Os cidadãos não demandam interoperabilidade; entretanto, os sistemas precisam interoperar, na maioria dos casos, para que o governo entregue o que o cidadão demanda” (Pardo e Burke, 2009, como referido em Araújo, 2012, p.19).

De acordo com (Navarra y Usero, 2011), entende-se por interoperabilidade a capacidade de um sistema para utilizar partes ou equipamento de outro sistema. Segundo os mesmos autores, fala-se de interoperabilidade quando se refere o conjunto dos “Procesos, tecnologías y protocolos requeridos para asegurar la integridad de los datos cuando se transfieren de un sistema a otro, así como la transmisión de resultados correctos y con significado para el usuario final” (Navarra y Usero, 2011, p.11).

Não basta fazer o intercâmbio dos dados. É necessário que os sistemas participantes consigam interpretar os dados transmitidos. Como no processo de comunicação, não chega garantir que o canal de transmissão faça chegar a mensagem do emissor ao recetor, é necessário que os recetores entendam a mensagem da mesma maneira.

Os métodos e procedimentos usados por cada Sistema de Informação (SI) para interpretar o significado do conteúdo das mensagens enviadas baseiam-se na descrição dos recursos da informação e no desenvolvimento de ferramentas para a representação do conhecimento contido nesses recursos (Navarra y Usero, 2011). Esta representação do conhecimento é operacionalizada através dos Sistemas de Organização do Conhecimento (SOC), ferramentas “que consistem em palavras, conceitos e relações semânticas, definidas e seleccionadas” (Carlan, 2010, p.30) para tratar a informação, de modo a poder recuperá-la, quer em ambientes mais tradicionais, quer em ambientes informatizados.

Procurando representar os conceitos de um domínio, através da construção organizada de modelos abstratos do mundo real, os SOC são entendidos na literatura como “instrumentos que fazem a tradução dos conteúdos dos documentos para um esquema estruturado, que representa esse conteúdo, com a finalidade de organizar a informação e o conhecimento e facilitar a recuperação das informações” (Carlan, 2010, p.28). Para Lara (2015), a sua operacionalização torna-se tão mais complexa quanto mais nebulosos forem os limites entre os SOC, motivo que tem estimulado a evolução de uma lógica clássica para uma lógica que permita contemplar a complexidade do mundo (Lara, 2015).

Para Hodge (2000, como referido em Carlan, 2010) os SOC constituem o ‘coração’ dos Sistemas de Recuperação da Informação (SRI) no ambiente físico e, principalmente, no ambiente web, cumprindo um objetivo de padronização terminológica.

É neste propósito que a interoperabilidade se contextualiza e, neste sentido, a principal motivação da norma ISO 25964-1/2 é a promoção da interoperabilidade. Na segunda parte da norma, a interoperabilidade com outros vocabulários é concretizada através de modelos de mapeamento. Todos os mapeamentos são feitos a partir de conceitos representados de diferente modo conforme o tipo de vocabulário procurando as equivalências, as hierarquias e as associações (Lara, 2015).

Os SOC devem funcionar como código inteligível e como fonte para a interpretação do sentido, mas também como metalinguagem ao incorporar o utilizador como parte integrante do processo. Para o efeito deve utilizar referências de linguagem e de significado partilhadas pela comunidade de utilizadores (Carlan, 2010), as quais são dadas mediante a criação de “consensos na Organização, (...) expressos mediante acordos semânticos” (Assis e Moura, 2015, p.2).

Para que possamos trabalhar com a interoperabilidade entre as instituições, é fundamental que estas tenham o mesmo entendimento sobre aquilo que desejam compartilhar, sendo necessária a definição de um conceito compartilhado (Araújo, 2012). Este será expresso pela sua semântica. São estes acordos de sentido, formalizados a partir de redes de colaboração, que garantem a interoperabilidade semântica.

De um modo geral, um SOC é escolhido de acordo com a cultura, a coleção ou um domínio, sendo que uns se adaptam melhor a uns contextos do que outros, dadas as especificidades desses mesmos contextos. Na revisão da literatura observou-se, contudo, uma tendência para a adoção de estruturas multidimensionais (tesauros, taxonomias e ontologias), assim como para a tendência (embora ainda pouco expressiva) da harmonização entre os diferentes SOC, a nível interno e externo à Organização.

Como explicita Lara (2015, p.105), mais do que problematizar ou reorganizar os diferentes SOC, trata-se de “tirar deles insumos para a construção de novas perspectivas de organização.”

Procurando contribuir para a compreensão desta complexidade estabeleceu-se como objetivo geral proceder a uma reflexão crítica sobre a possibilidade de harmonização dos SOC nas Organizações, de modo a facilitar a transição de dados entre os diferentes sistemas e garantir, em simultâneo, a integração e interoperabilidade semântica da informação, interna e externamente, sem constranger os processos de trabalho e a dinâmica organizacional nos seus diferentes contextos. Para tal, desenharam-se como objetivos específicos: i) contextualizar os ambientes organizacionais atuais (suas realidades e conjunturas) e a relevância da interoperabilidade nestes ambientes; ii) definir e compreender o conceito dos SOC e quais os tipos mais utilizados atualmente nas organizações, bem como a importância do seu uso para a recuperação do conhecimento; iii) refletir sobre a possibilidade da harmonização dos SOC na Organização, partindo dos diferentes contextos e das novas realidades.

Partiu-se de um paradigma qualitativo exploratório com recurso à revisão da literatura sistemática para compreender o propósito, modelos, padrões e estruturas das Organizações contemporâneas, bem como entender a evolução dos SOC neste contexto.

METODOLOGIA

Partindo da pergunta de investigação: qual a possibilidade de harmonização dos SOC nas Organizações, dados os diferentes contextos e as novas realidades e, com base nos objetivos enunciados, optou-se por uma abordagem qualitativa exploratória, baseada na revisão sistemática da literatura, de modo a apresentar um leque de ideias, de teorias, de argumentos fundamentados e reflexões defendidas por cada autor considerado, sobre o tema. O mapeamento destes trabalhos e a sua análise, baseada em seis categorias, permitir-nos-á identificar os estudos que têm vindo a ser desenvolvidos nesta área, ao mesmo tempo, que permitirá também realizar uma reflexão crítica sobre os mesmos. A adoção desta metodologia poderá evoluir ou ser completada, *à posteriori*, em outros estudos mais exaustivos, com outras técnicas metodológicas como o estudo de caso, o qual poderá ou não vir a corroborar o que se inferiu da revisão da literatura efetuada.

Para o efeito, a revisão da literatura subdividiu-se em duas fases: i) recolha e seleção dos textos, particularmente artigos científicos (Gil, 2002) e, ii) análise e interpretação dos mesmos (*corpus* de estudo) com base na análise de conteúdo, na qual foram privilegiadas seis categorias de análise.

1ª FASE

Num primeiro momento, realizou-se uma revisão bibliográfica geral focada nas seguintes temáticas: Organizações, Interoperabilidade e Sistemas de Organização do Conhecimento (SOC). Desta revisão foram privilegiadas três obras: *Gestão das Organizações* (Teixeira, 2005); *La interoperabilidad de la Información* (Usero & Navarra, 2007) e *Sistemas de Organización do Conhecimento: uma reflexão no contexto da Ciência da Informação* (Carlan, 2010). A sua leitura permitiu construir e fixar uma base de trabalho mais segura e rigorosa para se desenvolver posteriormente uma revisão de literatura sistemática. A seleção resultou do facto destas obras apresentarem, *per se*, um estudo abrangente das áreas consideradas, sem, contudo, caírem na generalidade; acresce ao facto as citações que referem serem uma mais-valia para a fundamentação das ideias deste estudo.

Num segundo momento procedeu-se à pesquisa na base de dados de referência *Web of Science* (WoS). Para a área da ciência da informação, consideraram-se os artigos publicados entre 2007-2017 que pudessem complementar e substanciar as noções adquiridas inicialmente. Partiu-se de uma pesquisa onde se consideraram os termos «Sistemas de organização do conhecimento», «Organizações», «Interoperabilidade» e «Harmonização» em inglês, português e espanhol. Os artigos considerados neste estudo encontravam-se em acesso aberto.

Dos textos recuperados na base de dados considerada, foram selecionados os 108 que se relacionavam com o tema, após o que se seguiu uma segunda seleção, baseada na leitura dos *abstract*, numa abordagem indutiva e de acordo com a pertinência percecionada em relação aos objetivos propostos, de modo a conduzir a resultados pertinentes a sólidas conclusões. Desta segunda seleção resultaram 38 textos que foram ainda triados de acordo com as palavras-chave apresentadas nos mesmos, de modo a irem ao encontro dos objetivos propostos. Desta 3ª triagem resultou um *corpus* de estudo de 18 artigos, que se apresentam no quadro 1.

2.ª FASE

Na análise do *corpus* procuraram-se as tendências atuais, semelhanças e diferenças nas teorias estudadas, o que nos permitiu realizar um cruzamento entre as principais ideias. Assim, se por um lado se consolidaram determinadas noções adquiridas; por outro, excluíram-se hipóteses e caminhos já abandonados por pesquisas realizadas.

Para a interpretação dos textos propriamente dita recorreu-se à análise de conteúdo, mediante a definição de 6 categorias de análise: Interoperabilidade, Organizações, Sistemas de Organização do Conhecimento, Linguagem, Comunicação e Harmonização, de forma a chegar a inferências (Bardin, 2003).

Quadro 1. *Corpus de Estudo*

Autor, data	Título do Artigo	Título do Periódico de Publicação do Artigo	Palavras-chave
Berité, et. al. 2010	Garantia literária: elementos para uma revisão crítica após um século	TransInformação	Garantia literária. Sistemas de organização do conhecimento. Vocabulário controlado
Brant e Medeiros, 2010	Folksonomia: esquema de representação do conhecimento?	TransInformação	Folksonomia. Organização do conhecimento. Organização da informação. Representação do conhecimento. Web.
Cavalcante e Bräscher, 2014	Taxonomias navegacionais em sítios de comércio eletrônico: critérios para avaliação	TransInformação	Comércio eletrônico. Navegação Web. Organização da informação. Sistemas de organização do conhecimento. Taxonomia navegacional.
Choo, 2007	Information seeking in organizations: Epistemic contexts and contests	Information Research	Não indica (retirados da leitura do artigo: Information; information behaviour; Knowledge; Organizations)
Feinberg, Bullard e Carter, 2013	Using design experiments to investigate conceptual issues in knowledge organization: an ongoing study	Information Research	Não indica (retirados da leitura do artigo: category systems; classification; residuality; organização do conhecimento)
Ganzha, et. al., 2012	Information resource management in na agent-based virtual organization—initial implementation	ComSIS	agent systems, ontologies, virtual organization, resource management, Information Resources, e-learning, adaptability.
Lara, 2013	Documentary languages and knowledge organization systems in the context of the semantic web	TransInformação	Linguagens documentárias. Interoperabilidade. Sistemas de organização do conhecimento. Web semântica. Tesouros. Vocabulários
LópezHuertas, 2013	Transcultural categorization in contextualized domains	Information Research	Não indica (retirados da leitura do artigo: classifications, knowledge organization models, cultural integration)
Orna, 2007	Keynote address: Collaboration between library and information science and information design disciplines. On what? Why? Potential benefits?	Information Research	Não indica (retirados da leitura do artigo: organizational information products, Organizations)
Ortega, 2008	Fundamentos da organização da informação frente à produção de documentos	TransInformação	organização da informação documentária; produção de documentos; produção de documentos eletrônicos; documento
Pastor-Sanchez, Mendez e Reodriguez-Muñoz, 2009	Advantages of thesaurus representation using the Simple Knowledge Organization System (SKOS) compared with proposed alternatives	Information Research	Não indica (retirados da leitura do artigo: Simple Knowledge Organization System, thesaurus representation, semantic web, RDF vocabularies)
Ramos et. al., 2011	Imagem e percepção humana: alternativa aplicada na classificação da literatura infantojuvenil de uma biblioteca escolar	Perspectivas em Ciência da Informação	Biblioteca escolar; Organização do conhecimento; percepção humana
Rodrigues, 2011	A seleção conceitual na organização de domínios de conhecimento nas ciências humanas e sociais: o caso da cultura	Perspectivas em Ciência da Informação	Human Science; Social Science; Conceptual selection; system of knowledge organization
Senso, Leiva-Mederos e Domínguez-Velasco, 2010	Modelo para la evaluación de ontologías. Aplicación en Onto-Satcol	Revista Española de Documentación Científica	Ontologías, evaluación de ontologías, ontologías de dominio, análisis, evaluación de ontologías, Ingeniería de puertos y costas
Sladić, et. al., 2015	The Use of Ontologies in Cadastral Systems	Computer Science and Information Systems	ontologies, cadastre, LADM, SDI
Tennis, 2013	Metaphors of time and installed knowledge organization systems: Ouroboros, Architectonics, or Lachesis?	Information Research	Não indica (retirados da leitura do artigo: knowledge organization systems, knowledge organization, architectonic, ouroboric, lachesis)
Widén-Wulff e Davenport, 2007	Activity systems, information sharing and the development of organizational knowledge in two Finnish firms: an exploratory study using Activity Theory	Information Research	Não indica (retirados da leitura do artigo: information sharing, organizational knowledge production, organizations)
Zirak, 2015	Implementation of Learning Organization Components in Ardabil Social Security Hospital	International Journal of Organizational Leadership	Learning Organization, Organizational Learning, Organizational Learning; Management, Knowledgebased; Learning System, Knowledge Management

RESULTADOS

Com base nas leituras efetuadas e de uma reflexão crítica acerca das perspetivas transversais e interoperáveis de harmonização dos SOC nas Organizações, chegou-se aos resultados que se apresentam no quadro 2.

A categoria mais frequente é a *Organização* (34,17%), o que decorre do facto de a nossa pesquisa ter sido focada nas Organizações, ficando assim garantido o foco. Para a reflexão em torno da harmonização dos SOC nas Organizações, os resultados revelam uma maior percentagem de incidência

nas categorias relacionadas com a *linguagem* (26,02%) e com os *SOC* (27,59%); observa-se ainda uma reduzida percentagem no que respeita às categorias *comunicação* (1,88%) e *harmonização* (0,63%).

Quadro 2. Resultado da análise dos textos com base nas categorias consideradas

Título	Autor, data	Categorias de análise						Total
		Interoperabilidade (interoperability)	Organizações (organizations)	Sistemas de Organização do Conhecimento (knowledge organization systems)	Linguagem (language)	Comunicação (communication)	Harmonização (harmonization)	
Garantia literária: elementos para uma revisão crítica após um século	Berité, et. al. 2010	1	1	19	8	2	0	31
Folksonomia: esquema de representação do conhecimento?	Brant e Medeiros, 2010	0	0	6	1	0	0	7
Taxonomias navegacionais em sítios de comércio eletrónico: critérios para avaliação	Cavalcante e Bräscher, 2014	0	0	3	2	0	0	5
Information seeking in organizations: Epistemic contexts and contests	Choo, 2007	0	40	0	0	0	0	40
Using design experiments to investigate conceptual issues in knowledge organization: an ongoing study	Feinberg, Bullard e Carter, 2013	1	0	0	2	0	0	3
Information resource management in na agent-based virtual organization—initial implementation	Ganzha, et. al., 2012	0	0	0	1	0	0	1
Documentary languages and knowledge organization systems in the context of the semantic web	Lara, 2013	16	0	12	11	0	0	39
Transcultural categorization in contextualized domains	LópezHuertas, 2013	0	0	5	0	0	0	5
Keynote address: Collaboration between library and information science and information design disciplines. On what? Why? Potential benefits?	Oma, 2007	0	25	0	5	0	0	30
Fundamentos da organização da informação frente à produção de documentos	Ortega, 2008	0	0	0	1	0	0	1
Advantages of thesaurus representation using the Simple Knowledge Organization System (SKOS) compared with proposed alternatives	Pastor-Sanchez, Mendez e Reodríguez-Muñoz, 2009	1	3	7	32	0	0	43
Imagem e percepção humana: alternativa aplicada na classificação da literatura infantojuvenil de uma biblioteca escolar	Ramos et. al., 2011	0	0	0	0	1	0	1
A seleção conceitual na organização de domínios de conhecimento nas ciências humanas e sociais: o caso da cultura	Rodrigues, 2011	0	0	5	8	3	0	16
Modelo para la evaluación de ontologías. Aplicación en Onto-Satcol	Senso, Leiva-Mederos e Domínguez-Velasco, 2010	0	0	0	3	0	0	3
The Use of Ontologies in Cadastral Systems	Sladić, et. al., 2015	12	1	0	6	0	2	21
Metaphors of time and installed knowledge organization systems: Ouroboros, Architectonics, or Lachesis?	Tennis, 2013	0	0	31	1	0	0	32
Activity systems, information sharing and the development of organizational knowledge in two Finnish firms: an exploratory study using Activity Theory	Widén-Wulf e Davenport, 2007	0	7	0	1	0	0	8
Implementation of Learning Organization Components in Ardabil Social Security Hospital	Zirak, 2015	0	32	0	1	0	0	33
Inferência Absoluta		31	109	88	83	6	2	319
Inferência Percentual		9,72%	34,17%	27,59%	26,02%	1,88%	0,63%	100,00%

Incidência com alguma expressividade é o lugar que a *interoperabilidade* (9,72%) ocupa neste conjunto, o que transmite a garantia que os diferentes SOC estão aptos a comunicar ou que os diferentes projetos em curso estão a encontrar uma forma de capacitar os SOC no sentido de se poder estabelecer uma comunicação efetiva entre si.

Assim, e na esteira de Carlan (2010), dos 3 elementos que os SOC devem integrar: o léxico (listas de elementos descritores devidamente filtrados e depurados), a rede paradigmática (importante para traduzir as relações essenciais e estáveis que existem entre os descritores) e a rede sintagmática (a expressão das relações contingentes que existem entre os descritores, ou seja, aquelas que são válidas no contexto particular em que aparecem), detivemo-nos nas questões do léxico por ser um elemento complexo de lidar ao nível organizacional devido às diferenças de visão, abordagem e paradigma dos diferentes domínios do conhecimento. “Em todas essas áreas é necessário definir como o conhecimento vai ser representado para que as representações permitam sua manipulação” (Brant e Medeiros, 2010).

Porém, a expressividade da categoria *linguagem* no *corpus* analisado aparece mais como sinónimo de *SOC* do que como o reflexo de uma análise das propriedades linguísticas ou da complexidade dessas propriedades no processo de comunicação, o que esgota o resultado esperado. Como Rodrigues (2011, p.138) alerta, “quando se trabalha com domínios de conhecimentos específicos, a seleção inadequada dos conceitos pode refletir na qualidade da sua recuperação da informação”, sobretudo num espaço inter e multidisciplinar, onde o conhecimento é dinâmico e progressivo (Rodrigues, 2011), e que é o espaço das Organizações atuais.

“Sobretudo quanto às áreas interdisciplinares com uma forte participação das ciências sociais, não é fácil encontrar conceituações ou denominações mais ou menos homogêneas dos tópicos, de tal modo que o alcance semântico de um termo pode expandir-se ou restringir-se em função de autores, correntes ou contextos concretos nos quais esse termo ganha valor de referência” (Barité, 2006, como referido em Barité et al., 2010, p.132).

Por outro lado, neste espaço de confluência, é possível entrever uma necessidade de *harmonizar* os diferentes SOC, ou seja, encontrar uma forma de entendimento e equilíbrio que advém das relações que se estabelecem entre os termos utilizados em cada um dos diferentes SOC em uso na Organização e, mais especificamente, entre os termos utilizados pelos diferentes SOC com relação ao mesmo objeto, o que se torna cada vez mais emergente no contexto interdisciplinar e colaborativo das Organizações atuais.

A este propósito Sladic *et al.* (2015, p.1049) procura apresentar um modelo de ontologia que permite obter interoperabilidade semântica no contexto nacional e internacional, mas também uma *harmonização* (de conceitos) que reforce a garantia da comunicação efetiva entre os diferentes SOC, tal como se pode inferir: “Ontologies are useful for data integration and harmonization with other cadastral systems using standard based domain ontology for cadastre on national and international level”.

Em Andrade (2015) também nos é possível observar a tendência de reflexão em torno da *harmonização*. Para a investigadora, e reportando-se aos mentores do projeto Dryad, numa Organização cada departamento poderá ter a sua própria solução de SOC, uma vez que um só SOC não poderia representar adequadamente a diversidade de disciplinas e necessidades conceptuais, mas desde que essas soluções sejam compatíveis (em *harmonia*), quer na ótica da Ciência da Informação, quer na ótica da Ciência da Computação.

Na tentativa de desenvolver um protótipo híbrido para o domínio da informação legislativa, os trabalhos de Torres *et al* (2015) denotam igualmente um processo de *harmonização* dos SOC, mas partindo da premissa de que, dentro de um contexto social, os utilizadores devem ser vistos como indivíduos em situações concretas, dentro de organizações sociais e de domínios do conhecimento.

É ainda neste sentido que o projecto *Linked Data* de Tim Berners-Lee tem desenvolvido a iniciativa *Linked Open Vocabularies (LOV)*, a qual se constitui em recomendações do *W3Consortium*, cujo objetivo é estimular e fundamentar a publicação da abertura dos dados relativos aos vocabulários controlados (Andrade, 2015). Esta tem vindo a ser operacionalizada através dos SKOS (*Simple Knowledge Organization System*) que permitem “estabelecer pontes entre conceitos de SOC distintos com estruturas diferentes e relações semânticas distintas, mantendo os termos na forma original dos SOC fontes” (Andrade, 2015, p.281).

Porém, pela insuficiência de trabalhos onde esta propriedade é argumentada, não foi possível obter resultados mais concretos a este nível, inviabilizando conclusões mais sustentadas e pragmáticas.

Do *corpus* de estudo analisado foi ainda possível observar uma evolução entre a era da pré-coordenação em que os SOC eram estruturas estáticas e atendiam às necessidades dos sistemas manuais de organização e recuperação da informação, e a era da web semântica em que os SOC se diferenciam “dos demais por serem projetados para uso por agentes inteligentes” (Vickery, 2008, como referido em Angelos, 2013, p.41), ou seja, a era das Organizações atuais que, dado o grande volume de informação decorrente da explosão informacional, continuam à procura de novas formas de organizar a informação (Angelos, 2013).

CONCLUSÕES

Dos resultados, infere-se que não há necessidade de desenvolver um único SOC na Organização, desde que os existentes garantam a interoperabilidade e a comunicação efetiva entre os diferentes sistemas. Desta inferência, intui-se que, ao desenvolver um só SOC para a Organização, tal poderá comprometer a completude da representação conceptual resultante dos diferentes contextos que aí coabitam.

Ainda, da observação das inferências percentuais é possível concluir a relevância que os SOC têm vindo a adquirir nas Organizações, embora as preocupações com a interoperabilidade ou a harmonização dos referidos SOC no contexto organizacional tenham ainda pouca expressão, sobretudo no caso da harmonização que se revela residual.

Tal circunstância releva do facto de os trabalhos que têm vindo a ser realizados nos últimos 10 anos se centrarem sobretudo no desenvolvimento e uso dos SOC em si mesmos, individualizados, ou seja, em que o foco se centra no SOC que está a ser desenvolvido para uma determinada área do conhecimento e não nas (inter)relações que terão de ser estabelecidas com os restantes SOC comuns a toda a Organização e em que operam conceitos como a interoperabilidade ou a harmonização.

A literatura aponta também para a construção de modelos de organização de conhecimento híbridos cuja conceção se baseiam na bidirecionalidade, resultando, desta orientação, novos modelos sustentados, à partida, em características de dois ou mais tipos de SOCs. De tal circunstância infere-se que a questão da interoperabilidade, no contexto considerado, não tenha evoluído substancialmente nos últimos 10 anos, pois centra-se em modelos individuais cuja construção se baseia na unidirecionalidade.

Intui-se que a harmonização, entre os diferentes SOC, garante a convergência entre as especificidades de cada contexto e as necessidades dos diferentes utilizadores, como beneficia o desenvolvimento de novas formas de organizar a informação e o conhecimento nas Organizações.

Por fim, espera-se que o interesse em torno do tema possa contribuir para o aperfeiçoamento do desenvolvimento de métodos e instrumentos cada vez mais eficientes e eficazes na Organização do Conhecimento, em particular no que se refere aos SOC, promovendo a realização de trabalhos futuros que ajudem a desenvolver soluções de SOC semanticamente viáveis, interoperáveis, que garantam a comunicação efetiva entre os diferentes sistemas, e a sua harmonização, quer ao nível da internalização, quer ao nível da externalização da informação.

O contrário poderá inviabilizar o desenvolvimento de Sistemas de Informação cada vez mais seguros, confiáveis e de qualidade que respondam com eficiência e eficácia às necessidades e interesses de cada Organização. Como Albert Einstein já refletia no início do séc. XX: *Things should be made as simple as possible, but no simpler.*

REFERÊNCIAS BIBLIOGRÁFICAS

Andrade, J. (2015). *Interoperabilidade e mapeamentos entre Sistemas de Organização do Conhecimento na Busca e Recuperação de Informações em Saúde: estudo de caso em Ortopedia e Traumatologia*. (Tese de Doutoramento). Universidade de S. Paulo, S. Paulo.

Angelos, L. (2013). *Evolução das Linguagens Documentárias até os Sistemas de Organização do Conhecimento*. (Monografia de Bacharel). Universidade de Brasília, Brasília.

Araújo, C. (2012). *Um Modelo para Interoperabilidade entre Instituições Heterogêneas*. (Tese de Doutoramento). Universidade de S. Paulo, S. Paulo.

Assis, J., Moura, M. A. (2015). Consenso e garantia no Contexto Digital: entre as Interoperabilidade Efetivas e Possíveis. In *XVI ENANCIB, Informação, memória e Património: do documento às redes*, Universidade Federal da Paraíba, João Pessoa/PB, 26-30 Outubro 2015.

Bardin, L. (2003). *Análise de conteúdo*. Lisboa: Edições 70.

Barité, M., et al. (2010). Garantia literária: elementos para uma revisão crítica após um século. *Revista TransInformação*, Campinas, 22(2):123-138, maio/ago., 2010.

Bocato, V. (2011). Os Sistemas de Organização do Conhecimento nas perspectivas actuais das normas internacionais de construção. *Revista de Ciência da Informação e Documentação*, 2 (1), 165-192.

Bocato, V., Torquetti, M. (2012). Interoperabilidad entre lenguajes de indización como recurso de construcción de herramientas de representación temática de recortes de noticias de periódicos de coordinadorias de comunicación social en ambientes universitarios: una propuesta metodológica. *Inf.*, 17 (3), 76-101.

Brant, M., Medeiros, M. (2010). Folksonomia: esquema de representação do conhecimento? *Revista TransInformação*, Campinas, 22(2):111-121, maio/ago., 2010.

Brascher, M., & Café, L. (2008). Organização da Informação ou Organização do Conhecimento? *ANCIB*, 1-14.

Carlan, E. (2010). *Sistemas de Organização do Conhecimento: uma reflexão no contexto da Ciência da Informação*. (Dissertação de Mestrado). Universidade de Brasília, Brasília.

Carlan, E., Medeiros, M. (2011). Sistemas de Organização do Conhecimento na visão da Ciência da Informação. *Revista Ibero-americana de Ciência da Informação*, 4 (2), 53-73.

Cavalcante, R., Bräscher, M (2014). Taxonomias navegacionais em sítios de comércio eletrônico: critérios para avaliação. *Revista TransInformação*, Campinas, 26(2): 191-201, maio/ago., 2014.

Choo, C. (2007). Information seeking in organizations: Epistemic contexts and contests. *Revista IR Information Research*, Vol. 12 No. 2, January 2007.

Feinberg, M., Bullard, J., Carter, D. (2013). Using design experiments to investigate conceptual issues in knowledge organization: an ongoing study. *Revista IR Information Research*, vol.18, n.3, September, 2013.

Ganzha, M. et al. (2012). Information resource management in an agent-based virtual organization – initial Implementation. *Revista ComSIS* Vol. 9, No. 3, Special Issue, September 2012.

Gil, A. C. (2002). Como elaborar projetos de pesquisa. São Paulo: Atlas.

Lara, M. (2013). Documentary languages and knowledge organization systems in the context of the semantic web. *Revista TransInformação*, Campinas, 25(2):145-150, maio/ago., 2013

Lara, M. (2015). Propostas de tipologias de KOS: uma análise das referências de formas dominantes de organização do conhecimento. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, 20 (esp. 1), 89-107.

López-Huertas, M. (2013). Transcultural categorization in contextualized domains. *Revista IR Information Research*, 18(3) paper C16.

Martínez Usero, J. Á., & Lara Navarra, P. (2007). La interoperabilidad de la información. Barcelona: UOC.

Neto, R.; Neves, J. (2012). Gestão da informação e do conhecimento nas organizações: análise de casos relatados em organizações públicas e privadas. *ENANCIB - Encontro Nacional de Pesquisa em Ciência da Informação*.

Oliveira, E. L.; Oliveira, E. A.; Carniello, M. F. (2011). Comunicação nas organizações: um estudo exploratório das tecnologias de comunicação utilizadas em empresas do Vale do Paraíba. *LAJBM*, 2 (2), 82-102.

Orna, L. (2007). Keynote address: Collaboration between library and information science and information design disciplines. On what? Why? Potential benefits? *Revista IR Information Research*, 12(4) paper colis02.

Ortega, C. (2008). Fundamentos da organização da informação frente à produção de documentos. *Revista TransInformação*, Campinas, 20(1): 7-15, jan./abr., 2008.

Pastor-Sanchez, J., Mendez, F. e Reodríguez-Muñoz, J. (2009). Advantages of thesaurus representation using the Simple Knowledge Organization System (SKOS) compared with proposed alternatives. *Revista IR Information Research*, vol.14, n.4, December, 2009.

Ramos, C. et al. (2011). Imagem e percepção humana: alternativa aplicada na classificação da literatura infantojuvenil de uma biblioteca escolar. *Revista Perspectivas em Ciência da Informação*, v.16, n.4, p.55-72, out./dez. 2011.

Rodrigues, A. (2011). A seleção conceitual na organização de domínios de conhecimento nas ciências humanas e sociais: o caso da cultura. *Revista Perspectivas em Ciência da Informação*, v.16, n.2, p.131-152, abr./jun. 2011.

Senso, J., Leiva-Mederos, A. e Domínguez-Velasco, S. (2011). Modelo para la evaluación de ontologías. Aplicación en Onto-Satcol. *Revista Española de Documentación Científica*, 34, 3, julio-septiembre, 334-356, 2011.

Sladić, D. et al. (2015). The Use of Ontologies in Cadastral Systems. *Revista Computer Science and Information Systems* 12(3):1033–1053.

Teixeira, S. (2005). *Gestão das Organizações*. Lisboa: Mc Graw-Hill.

Tennis, J. (2013). Metaphors of time and installed knowledge organization systems: Ouroboros, Architectonics, or Lachesis? *Revista IR Information Research*, vol.18, n.3, December, 2013.

Torres, S. et al. (2015). Relações semânticas em sistemas de organização do conhecimento: uma investigação no domínio da informação legislativa. In *Desafios y oportunidades de las Ciencias de la Información y la Documentación en la era digital: actas del VII Encuentro Ibérico EDICIC 2015*, Madrid, 16-17 Novembro 2015.

Vignoli, R., et al. (2013). Sistemas de Organização do Conhecimento com foco em Ontologias e Taxonomias. *Inf. & Soc.:Est.*, 23 (2), 59-72.

Widén-Wulff, G., Davenport, E. (2007). Activity systems, information sharing and the development of organizational knowledge in two Finnish firms: an exploratory study using Activity Theory. *Revista IR Information Research*, vol. 12, n.3, April, 2007.

Zirak, A. (2015). Implementation of Learning Organization Components in Ardabil Social Security Hospital. *International Journal of Organizational Leadership* 4(2015) 200-212.



ORGANIZACIÓN AUTOMÁTICA DE DOCUMENTOS MEDIANTE TÉCNICAS DE ANÁLISIS DE REDES SOCIALES

Carlos G. Figuerola

Universidad de Salamanca, 0000-0001-6799-2874, figue@usal.es

RESUMEN La organización automática de documentos permite conocer la estructura temática de grandes colecciones documentales. En este trabajo se plantea modelar la colección de documentos mediante una red o grafo de nodos y enlaces y aplicar técnicas de Análisis de Redes Sociales. Se describe un experimento llevado a cabo con una colección de noticias de prensa, analizando la estructura temática resultante tras aplicar técnicas de detección de comunidades de nodos en dos niveles de agrupamiento. Los resultados parecen bastante prometedores, planteando como trabajo futuro la implementación y comparación de diferentes algoritmos de detección de comunidades

PALAVRAS CLAVE *clasificación automática, clustering, Analisis de Redes Sociales*

ABSTRACT Automatic organization of documents allows to know the semantic structure of broad collections of documents. This paper proposes to model a document's collection by means of a graph or network and then apply the so called Social Networks Analysis techniques. We describe a practical experiment carried out with a collection of newspaper articles, and then we analyze the topic's structure resulting after applying communities discovery techniques. Results look enough promising; we envisage as future work the application and comparison of diverse communities discovery algorithms .

KEYWORDS *Automatic classification, clustering, Social Network Analysis.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUCCIÓN

El desarrollo de las Tecnologías de la Información y las Comunicaciones ha propiciado no sólo la producción de una cantidad cada vez mayor de documentos, sino también la aparición de una serie de cuestiones que han alterado en forma importante todo el proceso documental. Así, factores como la volatilidad, la falta de tipología y la aparición de tipologías documentales nuevas, las nuevas formas colaborativas de autoría, las estructuras documentales nuevas, etc. hacen que las técnicas aplicadas en entornos clásicos se muestren ineficaces. La descripción a partir de campos, más o menos explícitos, y el uso de lenguajes documentales (controlados o no) que se han aplicado durante décadas en el terreno de bibliotecas y centros de documentación; están orientados a la aplicación con documentos con soporte físico. Su aplicación en el panorama documental que ha producido el desarrollo de la tecnología es problemática, aunque parece claro que las técnicas y procedimientos que podríamos denominar clásicos, pensados para aplicarse en procesos documentales basados en documentos con soporte físico; son incapaces de gestionar eficazmente el contexto documental digital.

Una vía de abordar esta situación es la automatización de procesos; sin embargo, está lejos de poder enfrentarse de forma eficaz a estos problemas. Una de las cuestiones que tiene planteadas es la organización automática de los documentos en función de su contenido temático. Compatible con otros modelos de organización y acceso, la organización automática puede favorecer la recuperación, al colocar documentos de contenidos similares juntos; o el análisis cuantitativo de las colecciones documentales; o el desarrollo y perfeccionamiento de aplicaciones de extracción de información basados en *machine learning*, al poder ofrecer grandes cantidades de contextos similares para el entrenamiento de dichas aplicaciones.

En este trabajo se describe un sistema de clasificación automática de documentos a partir de la utilización de técnicas de Análisis de Redes Sociales; igualmente, se describe su aplicación en un caso concreto y se analizan los resultados. Este trabajo está organizado como sigue: en la sección siguiente se introducen las metodologías utilizadas: la clasificación automática de documentos y algunos de los sistemas clásicos más utilizados, así como el Análisis de Redes. En la sección siguiente se describe la colección de documentos utilizada, el procesamiento a que ha sido sometida, así como los algoritmos aplicados. A continuación, se describen los resultados obtenidos y se evalúan y discuten. Finalmente, se ofrecen unas conclusiones.

LA CLASIFICACION AUTOMATICA

Es frecuente distinguir entre sistemas de clasificación supervisada (a veces denominada también categorización) y clasificación no supervisada o *clustering* (Campos Ibáñez y Romero López, 2011; Ares Brea et al., 2011). En el primero de los casos los documentos se clasifican en una estructura creada *ad-hoc* por personas; el cometido de estos sistemas consiste en encontrar el lugar dentro de esa estructura donde debe colocarse cada documento a clasificar. La mayor parte de estos sistemas se basan en técnicas de *machine learning* y consisten, de una forma u otra, en construir patrones o modelos de las diferentes categorías que conforman la estructura organizativa; y medir después el parecido o similitud de cada documento con cada uno de esos patrones (Baharudin y otros, 2010). En función de esa similitud, cada documento es asignado a la categoría correspondiente.

Existen diferentes sistemas de clasificación supervisada, como los probabilísticos (Langley y otros, 1992; McCallum y Nigam, 1998), o el llamado *nearest neighbour* con sus variantes (Yang, 1999) o los *Support Vector Machines* (SVM) (Joachims, 2002). Estos sistemas se vienen empleando con éxito en diferentes aplicaciones; esto incluye, obviamente, el campo documental, en el que se han aplicado con el fin de clasificar documentos o texto (Eyheramendy, Lewis y Madigan, 2003; Kim et al., 2006; Joachims, T., 1998). Por nuestra parte, hemos podido documentar tasas de éxito superiores al 94 % (Figuerola, 2013; Quintanilla, Figuerola y Groves, 2015; Figuerola et al. 2017).

La clasificación no supervisada o *clustering*, de otro lado, carece de estructuras clasificatorias predefinidas y es el mismo sistema el que, en función de las características de la colección documental, las crea. Parte de cero, por así decir, construyendo una estructura organizativa en la que va encuadrando cada uno de los documentos. Así pues, es la base de la Organización Automática de Documentos.

Existen diversos sistemas para clasificar de forma automática colecciones amplias de documentos en función del contenido temático de esos documentos. Así, dentro de lo que se conoce como clasificación no supervisada, es relativamente frecuente el uso de algoritmos de *clustering* como *k-means* (Jain, 2010).

La versión clásica de éste tiene el inconveniente de producir *clusters* planos, de un solo nivel, además de necesitar fijar de antemano el número de *clusters* deseados, lo cual no siempre es fácil y requiere, con frecuencia, de un proceso de prueba y error.

Diferentes sistemas de modelado de temas (*topic modelling*) (Hidayat et al., 2015; Griffiths and Steyvers, 2004) se utilizan también para descubrir de forma automática los temas tratados en una colección de documentos. Obviamente, delimitados los temas, es posible agrupar o clasificar los documentos en función de esos temas. El procedimiento más conocido es el denominado *Latent Dirichlet Allocation* (LDA) (Blei, Ng and Jordan, 2003), que goza de cierto prestigio en el terreno de las *Digital Humanities* (Shawn and Milligan, 2012). Aunque existen implementaciones fáciles de utilizar (de ahí, tal vez, su relativa popularidad), tiene el inconveniente de que es preciso elegir de antemano el número de temas o topics deseados, lo cual no siempre es sencillo (Arun et al., 2010). Además, el etiquetado o identificación de los temas detectados suele requerir una fase de análisis por parte de expertos, no exenta de subjetividad e inconsistencias.

En este trabajo se plantea abordar la clasificación automática de documentos mediante la aplicación de técnicas de Análisis de Redes Sociales (Otte and Rousseau, 2002), efectuando una aplicación sobre una colección de documentos y analizando sus resultados. La aplicación de estas técnicas no requiere predefinir el número de *clusters* deseados, puede producir estructuras jerárquicas y, como se verá, los resultados de su aplicación parecen bastante prometedores.

TECNICAS DE ANALISIS DE REDES SOCIALES

El Análisis de Redes Sociales tiene su origen en la teoría matemática de Redes o Grafos y debe su nombre al hecho de ser aplicada, en los años 60 del siglo pasado, por sociólogos intentando modelar las relaciones entre personas y grupos sociales (Scott, 2013). Brevemente, una red es un conjunto de nodos o vértices conectados por arcos o enlaces. Los nodos pueden tener una serie de características o atributos arbitrarios, definidos por quien aplica este artefacto. Los enlaces o arcos conectan dos nodos entre sí; los enlaces pueden tener dirección (parten de un nodo y apuntan a otro) o no (simplemente conectan dos nodos en una relación ambivalente); pueden existir también arcos reflexivos (parten y llegan al mismo nodo). Los arcos o enlaces también pueden tener atributos arbitrarios a gusto del usuario, pero uno de los más habituales es el peso: un valor numérico que intenta expresar la fortaleza de la relación que representa ese arco.

Lo interesante de las Redes es que se han desarrollado métodos y procedimientos para analizar la estructura interna de una red. De manera que, si conseguimos modelar un determinado fenómeno mediante una red, podemos utilizar esas técnicas de análisis para estudiar la estructura interna de ese fenómeno. Una de las cosas que es posible hacer con una red es descubrir o detectar las posibles comunidades de nodos que haya en ella (Plantíé and Crampes, 2013).

Es posible modelar una colección de documentos como una red, en la cual los documentos pueden ser representados por nodos. Dos documentos o nodos pueden estar conectados entre sí por un arco o enlace si ambos tienen un contenido parecido o similar; y el peso de ese arco podría ser al grado de similitud entre esos dos documentos.

Desde finales de los años 60 del pasado siglo disponemos de formas para medir la similitud entre dos documentos, gracias a la formulación del modelo vectorial de recuperación de la información por G.

Salton (Salton, 1983). Naturalmente, ha habido desde entonces nuevos modelos y nuevas propuestas y podemos aplicar cualquiera de ellas; pero las más utilizadas hoy día siguen estando basadas en el modelo vectorial de Salton. Así que, de un modo u outro, es posible calcular la similitud entre cada pareja de documentos de nuestra colección, trazando una red de nodos y arcos. Sin alejarnos del modelo vectorial, podemos representar cada documento como un vector de términos con pesos. El cálculo del peso está basado en la frecuencia de cada término en cada documento y en la frecuencia inversa (*Inverse Document Frequency*) o *tf x idf*. Y después es posible aplicar cualquiera de las funciones habituales para estimar la similitud entre los vectores de los documentos.

A esta red es posible aplicar alguno de los varios sistemas disponibles de detección de comunidades de nodos. Una comunidad de nodos es un conjunto de éstos que enlazan fuertemente entre sí, y débilmente con los no pertenecientes a esa comunidad. Dado que los enlaces de nuestra red están basados en las similitudes de contenido entre los nodos, una comunidad debería agrupar a documentos de temática similar (Lee and Cunningham, 2014; Pons and Latapy, 2005).

LA COLECCION DE DOCUMENTOS

Hemos aplicado estas ideas a una colección de noticias de prensa extraídas de algunos periódicos. Esta colección está formada por 50.000 noticias, todas sobre Ciencia y Tecnología y es conocida como Spanish Corpus of Scientific Culture (SCSC) (Figuerola, Quintanilla et al., 2017). Se trata de noticias de tres diarios españoles de ámbito nacional en su versión digital, publicadas entre 2002 y 2011, a texto completo. A la totalidad de noticias extraídas de la hemeroteca digital de cada uno de los periódicos se aplicó un sistema de categorización para filtrar u obtener todas las noticias relacionadas con la Ciencia y/o la Tecnología.

El sistema de categorización automática aplicado, basado en SVM (*Support Vector Machine*), consiguió, en este caso, una precisión bastante elevada (94.5 % de aciertos), dejando la colección con una cantidad muy pequeña de ruido (Groves, Figuerola y Quintanilla, 2015). Las noticias se convirtieron a texto plano, desde su formato web original.

Pese a ser noticias sobre Ciencia y/o Tecnología, es obvio que, dentro de ese amplio campo, es posible preguntarse por su estructura temática: qué campos o disciplinas científicas están presentes, y en qué medida (con qué número de documentos o noticias); aquí es donde entra la aplicación de las técnicas de Análisis de Redes.

Se calculó la similitud entre cada par de noticias o documentos aplicando un sistema clásico de *tf-idf* para estimar los pesos de cada palabra y el bien conocido coeficiente del coseno (como el que aplican muchos sistemas de recuperación de la información) para medir la similitud. El resultado es una semi-matriz cuadrada de 50,730 por 50,730 elementos. Esta matriz es bastante densa, puesto que casi todos los documentos tienen alguna palabra en común, aunque sean palabras muy poco significativas, de peso muy bajo.

Esta matriz de similitudes nos permite construir una red de 50.730 nodos (cada documento). Y si enlazamos cada documento con los que mantiene alguna similitud, por baja que ésta sea, obtenemos una cantidad de enlaces cercano a 1.200 millones; una cantidad intratable debido a su tamaño. Se impone podar enlaces aplicando un umbral de similitud; tras varias pruebas, se aplicó un umbral de

0.09, eliminando todos los enlaces inferiores a este valor. En resultado es una red de 50,730 nodos y 23,612,49 enlaces.

Una representación visual de esa red puede observarse en la figura 1.

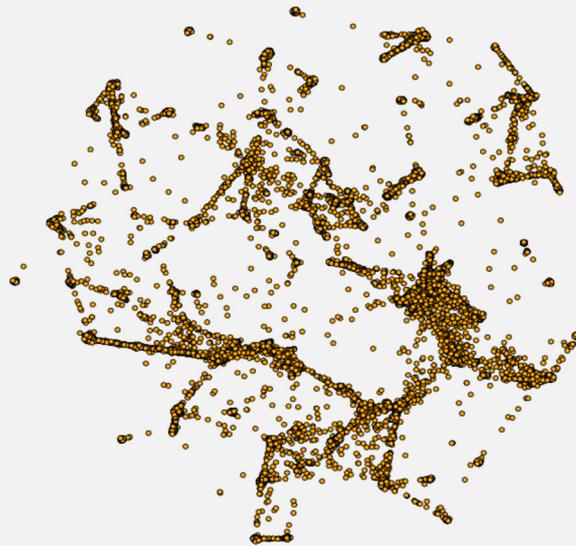


Figura 1. Red de documentos

Esta representación está realizada mediante un algoritmo *force directed* conocido como *Open Ord*, (Martin et al., 2011), que intenta colocar los nodos más fuertemente enlazados más próximos entre sí. Sin entrar en las dificultades de conjugar un número tan alto de nodos y enlaces, claramente podemos observar como hay conjuntos de nodos (documentos) más o menos agrupados. Estos grupos es lo que conocemos como comunidades de nodos de una red y, dado que la relación entre los nodos, en nuestro caso, está basada en la similitud semántica; podemos pensar que esas comunidades agrupan documentos que tratan sobre los mismos o parecidos temas.

Existen diversos algoritmos de detección de comunidades de nodos en redes; uno de los inconvenientes de muchos de estos algoritmos es la elevada capacidad de procesamiento que requieren (Lancichinetti and Fortunato, 2009). Por este motivo hemos aplicado el algoritmo de detección de comunidades conocido como *Infomap* (Rosvall, Axelsson and Bergstrom, 2009; Bohlin et al., 2014; Edler and Rosvall, 2015), una de cuyas características es la de ser bastante más rápido que sus competidores.

RESULTADOS

Como resultado obtenemos 23 comunidades, muchas de ellas con un segundo nivel de subcomunidades cada una. Sin embargo, de ellas solamente 13 tienen un tamaño, en número de documentos, apreciable, abarcando el 90 % de toda la colección.

Un análisis a simple vista de las noticias componentes de cada comunidad significativa muestra la consistencia temática de las mismas. Para tratar con estas comunidades con mayor comodidad podemos etiquetarlas manualmente para describir su contenido temático; obviamente el detector de comunidades simplemente las identifica con un número. Así, podemos ver como comunidades más significativas las que contienen noticias sobre Salud Pública, Biomedicina, Energía, Desarrollo Humano, Recursos

Naturales, Investigación Aeroespacial, Biodiversidad, Astronomía y Cosmología, Tecnologías de la Información, Política Científica, Especies Animales y Vegetales Protegidas, Contaminación.

Pero muestra también la existencia de subcomunidades claramente definidas dentro de cada grupo de primer nivel. Por ejemplo, dentro de la comunidad o grupo temático relativo a la Salud Pública, las noticias referentes a epidemias en sus diversos aspectos aparecen juntas, formando una subcomunidad diferenciada de otras noticias referidas a otros asuntos, dentro de lo que podemos entender como Salud Pública. Pero dentro de esa misma subcomunidad sobre epidemias, las noticias referentes a enfermedades concretas aparecen agrupadas en función de la enfermedad que sea: las de la malaria por un lado, las de la gripe aviar por otro, las relativas al VIH por otro, etc.

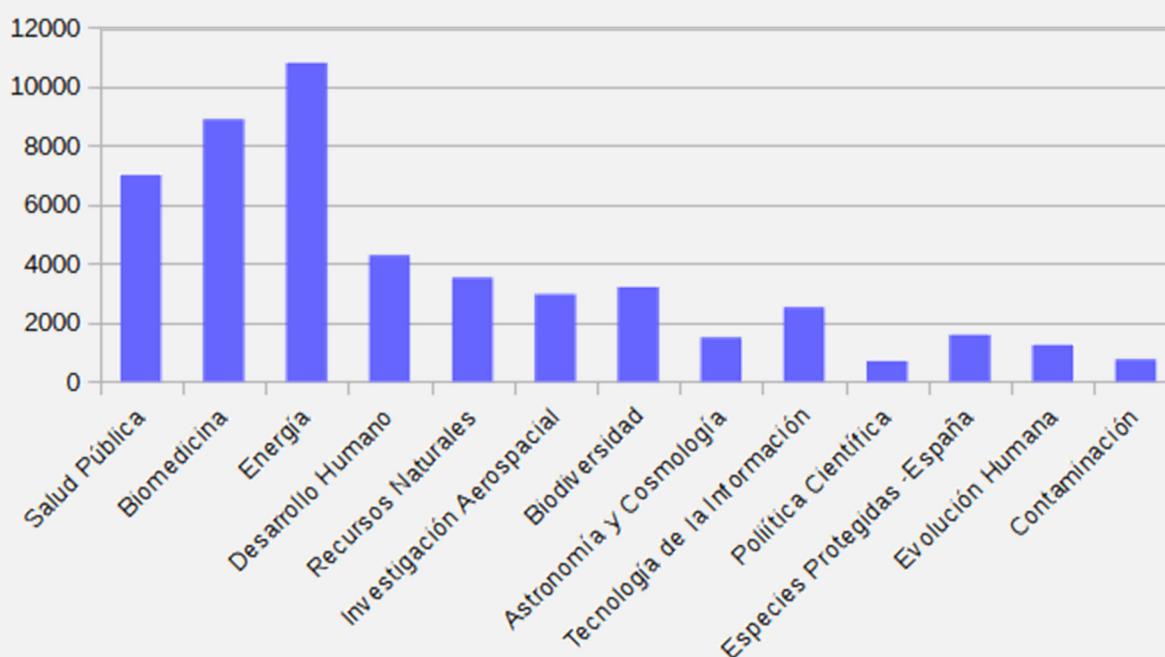


Figura 2. Comunidades o temas más importantes

Una evaluación algo más rigurosa que la simple observación plantea algunas cuestiones, al no disponer de puntos de referencia. En efecto, en la evaluación de la clasificación automática suele distinguirse entre sistemas externos e internos. Los primeros se basan en contrastar los resultados obtenidos en una colección de documentos construida *ad-hoc*, de la cual se conoce previamente su estructura temática. Obviamente éste no es nuestro caso, por lo que debemos recurrir a un sistema interno de evaluación, es decir, medir la similitud entre los componentes de cada comunidad y la diferencia o distancia con los otros *clusters*.

Una de las medidas más difundidas es la conocida como *silueta* (Rousseau, 1987; Rendón et al., 2011) de un documento. Éste es un coeficiente que se aplica a cada documento individual e intenta conjugar ambas cuestiones: la similitud con todos los integrantes de su misma comunidad y la separación con los documentos de la comunidad más próxima. Su valor oscila entre -1 y 1; los valores más alejados de -1 son mejores. Además, cuanto más se acercan a 1 indican una buena separación entre comunidades, algo que depende mucho de los campos temáticos y de las características semánticas de los documentos.

Tabla 1. Algunos titulares de la subcomunidad 1.6

Comunidad	Titular
1.6	Mosquitos transgénicos contra la malaria
1.6	Malaria, un recorrido desde la quinina hasta la esperada vacuna
1.6	Acabar con la malaria costará 4.400 millones durante 50 años - Público.es
1.6	Males importados pero no exóticos
1.6	Los más ricos, entre los más pobres
1.6	Un tratamiento de 0,20 euros al año por niño evita un 20% de la malaria infantil
1.6	Optimismo ante un tratamiento para prevenir la malaria en niños
1.6	La malaria regresa a Europa
1.6	Noventa millones de niños africanos siguen sin dormir bajo una mosquitera
1.6	La terapia contra la malaria sólo alcanza al 16% de los enfermos
1.6	La malaria mata cada año a más de un millón de personas en África
1.6	Una vacuna ejemplar
1.6	Pedro Alonso: 'La vacuna contra la malaria estará lista en 2011'
1.6	Las iniciativas antimalaria salvaron a 750.000 niños

La silueta es un valor individual para cada documento, por lo que para una valoración de conjunto tal vez es más útil calcular la silueta media. La Tabla I muestra los valores medios para cada comunidad importante (en número de documentos). Es preciso indicar que la separación entre comunidades se ha medido tomando como referencia la comunidad de primer nivel.

Tabla 2. Valores de silueta para los temas más importantes

Com.	Etiqueta	Silue.
1	Salud Pública	0.62
2	Biomedicina	0.52
3	Energía	0.53
4	Desarrollo Humano	0.56
5	Recursos naturales	0.55
6	Investigación Aeroespacial	0.62
7	Biodiversidad	0.55
8	Astronomía y Cosmología	0.57
9	Tecnología de la Información	0.59
10	Política Científica	0.56
11	Especies protegidas-España	0.66
12	Evolución Humana	0.58
13	Contaminación	0.56

Los valores de silueta son bastante elevados; indican claramente una firme cohesión interna de cada comunidad o cluster, y una separación con el resto de los clusters bien definida.

El cálculo de la silueta, de otro lado, se ha efectuado teniendo en cuenta solamente los ítems dentro de cada subcomunidad. Por ello los valores de silueta a nivel de subcomunidad son relativamente bajos, incluso por debajo de 0 en algunos casos. La silueta media calculada por subcomunidades (nivel 2) es de 0.21, un valor algo bajo que puede explicarse por la relativamente baja separación entre clusters, dado que las subcomunidades más cercanas, que sirven como base para el cálculo de la silueta, forman parte de la misma comunidad y existe, por tanto, afinidad temática aunque de un nivel más amplio. Recordemos, de otro lado, que todas las noticias o documentos tratan sobre Ciencia y/o Tecnología, lo cual introduce ya un elemento de proximidad de contenido entre todas ellas.

Tabla 3. Subtemas del tema 3 (Energía)

Subc.	Palabras clave	Subc.	Palabras clave
3.1	energía, nuclear, solar	3.10	gas natural
3.2	cambio climático, emisiones CO2, kyoto	3.11	seísmos, volcanes
3.3	energía nuclear, Fukushima	3.12	nuclear, reactor fusión
3.4	energía nuclear, centrales nucleares en España	3.13	medio ambiente, capa de ozono
3.5	cambio climático, calentamiento global, ecologismo	3.14	energía, petróleo
3.6	seísmos, tsunamis	3.15	automóvil, motor, motor y medio ambiente
3.7	medio ambiente, residuos, contaminación	3.16	automóvil, tráfico, seguridad vial
3.8	medio ambiente, contaminación	3.18	arquitectura, rascacielos
3.9	medio ambiente, cambio climático edificación	3.19	casas solares

CONCLUSIONES

La clasificación automática de documentos nos permite conocer la estructura temática y su distribución en una colección de documentos. El Análisis de Redes Sociales nos permite modelar una colección de documentos como una red de afinidades temáticas, de manera que la aplicación en esa red de sistemas de detección de comunidades descubre los temas tratados en esos documentos y la estructura temática de la colección. A diferencia de otros sistemas de organización automática, el número de temas es establecido por el propio sistema en función de las características de la colección documental; además, es posible establecer al menos un subnivel de clasificación.

Los experimentos efectuados sobre una colección de noticias de prensa muestran una gran precisión en los resultados; como trabajo futuro, se plantea la aplicación de otros sistemas de detección de comunidades y la comparación entre los resultados obtenidos.

REFERENCIAS BIBLIOGRÁFICAS

- Aggarwal, C. C. y Zhai, C. (2012). A survey of text clustering algorithms, en Aggarwal y Zhai, eds.: *Mining Text Data*. Springer US: Boston MA. 77--128
- Ares Brea, M.E.; Parapar López, J.; Barreiro García, A. (2011). Agrupamiento Documental // Casheda Seijo, F.; Fernández Luna, J.M. ; Huete Guadix, J.F. Eds. (2011). *Recuperación de Información. Un enfoque práctico y multidisciplinar*. Madrid; Ra-Ma, 2011. 392-416.
- Arun, R.; Suresh, V.; Veni Madhavan, C. E.; Narasimha Murthy, M. N.; Zaki, M. J.; Yu, J. X.; Ravindran, B.; Pudi, V. (2010). On Finding the Natural Number of Topics with Latent Dirichlet Allocation: Some Observations. // *Advances in Knowledge Discovery and Data Mining: 14th Pacific-Asia Conference*, PAKDD 2010. Hyderabad, India. 391-402. http://dx.doi.org/10.1007/978-3-642-13657-3_43 (2017-01-12).
- Blei, D., Ng, A.; Jordan, M. (2003). Latent dirichlet allocation. // *The Journal of Machine Learning Research*, 3 993-1022.
- Baharudin, B; Lee, L. H.; Khan, K. (2010). A review of machine learning algorithms for text-documents classification. // *Journal of Advances in Information Technology*. 1:1 4-20.
- Bohlin, L., Edler, D., Lancichinetti, A.; Rosvall, M. (2014). Community detection and visualization of networks with the map equation framework. // *Measuring Scholarly Impact* (pp. 3-34). Springer International Publishing.
- Campos Ibáñez, L.M.; Romero López, A.E. (2011). Clasificación documental. // Casheda Seijo, F.; Fernández Luna, J.M. ; Huete Guadix, J.F. Eds. (2011). *Recuperación de Información. Un enfoque práctico y multidisciplinar*. Madrid; Ra-Ma, 2011. 359-392.
- Edler, D.; Rosvall, M. (2015). *The infomap software package*. <http://www.mapequation.org/code.html> (2017-02-16).
- Eyheramendy, S., Lewis, D. D., & Madigan, D. (2003). *On the naive bayes model for text categorization*. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.20.4949> (2017-02-16).
- Figuerola, C.G. (2013). *Clasificación automática de documentos. Un caso práctico*. <http://grulla.usal.es/figuerola2013clasificacion.pdf> (2017-02-16).
- Figuerola, C.G.; Quintanilla Fisac, M.A. et al. (2017): *Sistema de Indicadores para el SCSC (Spanish Corpus of Scientific Culture)*. <http://grulla.usal.es/figuerola2017sistema.pdf> (2017-03-28).
- Griffiths, T. L.; Steyvers, M. (2004). Finding scientific topics. // *Proceedings of the National Academy of Sciences*, 101:1, 5228-5235.
- Groves, T; Figuerola, C.G. y Quintanilla, M.A (2015). Ten years of science news: A longitudinal analysis of scientific culture in the Spanish digital press. *Public Understanding of Science*. 25:6, 691 – 705. <https://gredos.usal.es/jspui/handle/10366/127539> (2017-02-16)
- Jain, A. K. (2010). Data clustering: 50 years beyond K-means. // *Pattern recognition letters*, 31:8, 651-666. http://www.ppgia.pucpr.br/~fabricio/ftp/Roges/JainClustering_PRL10.pdf (2017-02-18)

- Joachims, T. (1998, April). Text categorization with support vector machines: Learning with many relevant features. In *European conference on machine learning* (pp. 137-142). Springer Berlin Heidelberg. https://eldorado.tu-dortmund.de/bitstream/2003/2595/1/report23_ps.pdf (2017-02-16).
- Joachims T. (2002) *Learning to Classify Text Using Support Vector Machines – Methods, Theory and Algorithms*. Boston, MA: Kluwer Academic Publishers.
- Kim, S. B., Han, K. S., Rim, H. C., & Myaeng, S. H. (2006). Some effective techniques for naive bayes text classification. *IEEE transactions on knowledge and data engineering*, 18(11), 1457-1466. <http://ir.kaist.ac.kr/papers/2006/some%20effective%20techniques%20for%20naive%20bayes%20text%20classification.pdf> (2017-02-16).
- Hidayat, E. Y.; Firdausillah, F.; Hastuti, K.; Dewi, I. N.; Azhari, A. (2015). Automatic Text Summarization Using Latent Dirichlet Allocation (LDA) for Document Clustering. // *International Journal of Advances in Intelligent Informatics*, 1:3 132-139.
- Lancichinetti, A.; Fortunato, S. (2009) Community detection algorithms: A comparative analysis.// *Physical Review E*. 80:5. <http://arxiv.org/pdf/0908.1062v2.pdf> (2017-02-18)
- Langley, P.; Iba, W.; Thompson, K. (1992). An analysis of bayesian classifiers. // *Proceedings of National Conference on Artificial Intelligence*. San Antonio, CA: AAAI Press and MIT Press. 223–228
- Lee, C.; Cunningham, P. (2014) Community detection: Effective on large social networks. *Journal of Complex Networks*. // 2:1 19–37. <http://comnet.oxfordjournals.org/content/2/1/19.full.pdf+html> (2017-02-18)
- Leydesdorff, L. (2008). On the normalization and visualization of author co-citation data: Salton's Cosine versus the Jaccard index. *Journal of the American Society for Information Science and Technology*, 59(1), 77-85.
- Martin, S.; Brown, M.W.; Klavans, R.; Boyack K.W.(2011). OpenOrd: an open-source toolbox for large graph layout. // *Proc. SPIE 7868, Visualization and Data Analysis 2011*. doi:10.1117/12.871402
- Martin-Pozuelo Campillos, M. P. (1996). *La construcción teórica en archivística: el principio de procedencia*. Madrid: Universidad Carlos III de Madrid.
- McCallum, A.; Nigam, K. (1998) A comparison of event models for naive bayes text classification. // *AAAI-98 workshop on learning for text categorization*. 41-48. <http://www.kamalnigam.com/papers/multinomial-aaaiws98.pdf> (2016-12-14)
- Otte, E.; Rousseau, R. (2002). Social network analysis: a powerful strategy, also for the information sciences. // *Journal of information Science*. 28:6 441-453. http://www.academia.edu/download/42254790/Social_Network_Analysis_A_Powerful_Strat20160206-25456-1pc1lcl.pdf (2017-02-18)
- Plantíe, M. ; Crampes, M. (2013) Survey on social community detection. // *Social media retrieval*, 65–85. <http://hal.archives-ouvertes.fr/docs/00/80/42/34/PDF/Survey-on-Social-Community-Detection-V2.pdf> (2017-02-18)

- Pons, P.; Latapy, M. (2005). Computing communities in large networks using random walks. // *Computer and information sciences (ISCIS)* 284–293. <http://arxiv.org/abs/physics/0512106> (2017-02-18)
- Rendón, E.; Abundez, I.; Arizmendi, A.; Quiroz, E. (2011). Internal versus external cluster validation indexes. // *International Journal of computers and communications*. 5:1 27-34.
- Rosvall, M.; Axelsson, D.; Bergstrom, C. (2009). The map equation. // *European Physical Journal Special Topics*. 178 13–23.
- Rousseeuw, P. J. (1987). Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis. *Computational and Applied Mathematics*. 20 53–65. doi:10.1016/0377-0427(87)90125-7.
- Salton, G.; McGill, M.J. (1983) *Introduction to Modern Information Retrieval*. New York, NY: McGraw-Hill.
- Scott, J. (2013). *Social network analysis*. Thousand Oaks, CA, US: Sage Publications, Inc
- Shawn, G.; Milligan, I. (2012). Review of MALLET, produced by Andrew Kachites McCallum. // *Journal of Digital Humanities*, 2:1, <http://journalofdigitalhumanities.org/2-1/review-mallet-by-ian-milligan-and-shawn-graham/> (2017-03-15)
- Yang, Y. (1999). An evaluation of statistical approaches to text categorization. // *Information retrieval*. 1:1-2 69-90.



VISIBILIDADE DE PESQUISADORES NA TEMÁTICA POLÍTICA DE INDEXAÇÃO: ANÁLISE DE CITAÇÃO E COCITAÇÃO

Roberta Cristina Dal' Evedove Tartarotti¹, Paula Regina Dal' Evedove²,
Mariângela Spotti Lopes Fujita³

¹Universidade Estadual Paulista (UNESP), 0000-0002-0246-4971, roberta_tartarotti@yahoo.com.br

²Universidade Federal de São Carlos (UFSCar), 0000-0003-1521-2201, dalevedove@ufscar.br

³Universidade Estadual Paulista (UNESP), 0000-0002-8239-7114, mariangelaslf57@gmail.com

RESUMO Na contemporaneidade, um campo científico se consolida e adquire visibilidade por meio de seus atores e produção científica. O trabalho teve por objetivo caracterizar a comunidade científica em Política de Indexação no âmbito da Organização e Representação do Conhecimento, a partir da publicação de artigos em periódicos nas bases de dados *Web of Science* e *Scopus* por meio de indicadores bibliométricos de citação e cocitação. Os resultados indicam a existência de três subredes de cocitação sobre a temática Política de Indexação, sendo que a frente de pesquisa é composta principalmente por autores brasileiros. Como considerações finais, observa-se pouca produção científica na temática neste universo de pesquisa.

PALAVRAS-CHAVE *Indexação, Política de indexação, Análise de citação, Análise de cocitação, Rede social de cocitação.*

ABSTRACT In contemporary times, a scientific field is consolidated and acquires visibility through its actors and scientific production. The objective of this work was to characterize the scientific community in Indexing Policy within the scope of Organization and Representation of Knowledge, from the publication of articles in periodicals in the *Web of Science* and *Scopus* databases through bibliometric citation and co-citation indicators. The results indicate the existence of three co-citation subnetworks on the subject Indexing Policy, and the research front is represented mainly by Brazilian authors. As final considerations, there is little scientific production on the subject in this universe of research.

KEYWORDS *Indexing, Indexing policy, Citation analysis, Co-citation analysis, Co-citation social network.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

1 INTRODUÇÃO

O processo de comunicação na ciência é vital, pois sua legitimidade somente se processa quando analisada e revista por seus pares. Desse modo, o ciclo de produção do conhecimento científico, institucionaliza-se por meio da publicação de seus resultados à comunidade científica mediante processo de recuperação da informação e publicação dos resultados de pesquisa.

Neste cenário, o campo da Bibliometria visa quantificar a medição de índices de produção do conhecimento científico, mapeando um determinado campo científico por meio de “ferramentas e

modelos matemáticos, estatísticos e computacionais para realizar análises quantitativas e comparativas das pesquisas científicas” (Tartarotti & Fujita, 2016, p. 140).

Segundo Cabrini, Tannuri e Matos (2009), “toda produção na ciência, qualquer que seja o campo, tem como premissa o fato de que o cientista constrói seu trabalho a partir de trabalhos anteriores, e mostra isso mencionando-os em uma lista de referências”. Ao incorporar uma citação em seu trabalho, “o pesquisador identifica os pesquisadores anteriores, cujos conceitos, métodos ou teorias serviram de paradigma e referência para seu trabalho atual, gerando uma listagem de citações”.

Por meio da frequência de citações e cocitações, é possível visualizar as relações estabelecidas entre os pesquisadores e a frente de pesquisa sobre um determinado campo, área ou temática do conhecimento, partindo-se da premissa de que “as mesmas pressupõem pesquisas com similaridades de assunto e associação ou cocorrência de ideias”. (Cabrini, Tannuri & Matos, 2009). Nesse sentido, os estudos bibliométricos em torno da área de Organização e Representação do Conhecimento da Ciência da Informação têm contribuído para traçar um panorama do conhecimento científico produzido e da atuação de autores e seus pares, viabilizando seu mapeamento por meio da avaliação e mensuração com uso de indicadores bibliométricos de produção, citação ou ligação.

De forma particular, no âmbito da Organização e Representação do Conhecimento a política de indexação é apropriada e útil para o contexto de bibliotecas universitárias quando reflete seus fundamentos teórico-metodológicos “e, ao mesmo tempo, as necessidades profissionais, contribuindo com o processo de melhorias contínuas no fazer do bibliotecário” (Dal’Evedove, Tartarotti & Fujita, 2015).

O trabalho teve por objetivo identificar a comunidade científica na temática “Política de Indexação”, a partir da publicação de artigos em periódicos nas bases de dados *Web of Science* e *Scopus* por meio de indicadores bibliométricos de citação e cocitação.

2 POLÍTICA DE INDEXAÇÃO NO ÂMBITO DA ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO

Em meio à crescente produção de conhecimento e inserção das tecnologias de comunicação e informação em bibliotecas, características de uma sociedade contemporânea, um dos desafios teórico-práticos da área de Organização e Representação do Conhecimento da Ciência da Informação centra-se no adequado tratamento, recuperação e disseminação do ciclo de produção de informações. Neste cenário, as bibliotecas buscam otimizar seus produtos e serviços com base no estabelecimento de instrumentos e diretrizes que norteiem as atividades de organização para a socialização das informações.

O sistema de recuperação da informação de bibliotecas abriga um conjunto de documentos – impressos, virtuais ou híbridos – reunidos em função dos interesses e necessidades de sua comunidade usuária que também possui características particulares e variáveis de acordo com o domínio do conhecimento e nível de especialidade. Com isso, o profissional que lida com estas informações precisa dominar técnicas adequadas para organizá-las, possibilitando sua futura recuperação.

Na área de Organização e Representação do Conhecimento da Ciência da Informação, a abordagem teórico-prática da indexação é definida como um processo formado por etapas que buscam identificar e descrever o conteúdo de um documento por meio de uma metalinguagem construída – a linguagem

documental – para a efetiva recuperação da informação. Durante a indexação, os conceitos são extraídos do documento por meio da análise de assunto, e são traduzidos para os termos de instrumentos de indexação.

Contudo, cada uma das etapas realizadas no processo de indexação, assim como as demais atividades de tratamento da informação documental, não são neutras e sofrem influências de diversos variáveis. Nesse sentido, algumas problemáticas subjacentes se apresentam como fatores interferentes no processo de indexação: o próprio sistema de recuperação da informação adotado; o perfil dos usuários; o conhecimento prévio do indexador, sua experiência profissional e sua formação em análise de assunto e domínio das ferramentas utilizadas (Gil Leiva, 2008); as diretrizes profissionais (Mai, 1997); as diretrizes ou a política de indexação da biblioteca (Fujita & Rubi, 2006) e, em um contexto mais amplo, a política de ciência e tecnologia da universidade (Tartarotti, 2014), dentre outros (grifo nosso).

Conceitualmente, a política de indexação “é um guia de tomada de decisões” (Carneiro, 1985) e “a forma de realizar a indexação de uma determinada instituição, isto é, o esforço por concretizar, sistematizar e reunir em manuais os processos da indexação” (Gil Leiva, 2008). Caracteriza-se, ainda, como

[...] um conjunto de diretrizes que favoreça o aprimoramento e a consistência na recuperação da informação. Por ser uma sistematização que integra a gestão da informação registrada, a política de indexação atua como recurso base para o estabelecimento de condutas teóricas e práticas dos profissionais envolvidos no tratamento da informação documentária (Tartarotti, Dal’ Evedove & Fujita, 2016, p. 9).

O objetivo principal do estabelecimento de uma política de indexação reside na padronização dos procedimentos para a realização do processo de tratamento temático da informação (Dal’ Evedove, Tartarotti & Fujita, 2015). Esta sistematização das diretrizes em manuais de indexação está diretamente relacionada com o número de profissionais que se dedicam à atividade de indexação e ao nível de qualidade que o sistema de informação pretende alcançar (Gil Leiva, 2008). Por outro lado, a política de indexação não se limita a uma simples lista de procedimentos adotados pelo sistema de recuperação da informação, visto que sua elaboração deve ser conduzida âmbito da administração da biblioteca, representada por uma filosofia que reflita os seus objetivos e que sirva de guia para os bibliotecários (Rubi, 2012, p. 92).

Nesse sentido, como um fator-chave que influencia cada uma das etapas do processo de indexação, a elaboração, implantação e avaliação de uma política de indexação torna-se fundamental para a melhoria dos produtos e serviços em unidades de informação, particularmente em bibliotecas universitárias.

A política de indexação, que consiste em uma decisão gerencial, divide as atividades informativas em gerenciais, de processamento e de interface com o usuário, passando a ver a gestão da unidade ou serviço de informação como um todo, no qual se inserem atividades de organização, disponibilização e uso da informação, de forma integrada e interdependente (Cesarino & Pinto, 1978). Para o estabelecimento de uma política direcionada ao tratamento documental, como um guia para tomada de decisões, diversos são os fatores a serem considerados, visto que cada biblioteca universitária possui suas próprias especificidades e características. De fato, a política de indexação está inserida em dois contextos complementares: o contexto sociocognitivo do profissional – a política de indexação, as regras e os procedimentos do manual de indexação, a linguagem documental para representação e mediação da linguagem natural do usuário e seus interesses de busca e; o contexto físico de trabalho (Fujita, 2003).

3 OS ESTUDOS DE ANÁLISE DE CITAÇÃO E COCITAÇÃO NA CIÊNCIA

A comunicação científica é um elemento vital para a institucionalização social da ciência na contemporaneidade. Com o objetivo de contribuir com subsídios teórico-metodológicos quantitativos em torno do ciclo de produção da ciência, o campo de estudos da Bibliometria é definido como “conjunto de conhecimentos metodológicos para a aplicação de técnicas quantitativas, destinadas ao estudo dos processos de produção, comunicação e uso da informação científica” (Carrizo Sainero, 2000, p. 9). Para a avaliação do comportamento científico, são diversos os indicadores bibliométricos empregados, dependendo do tipo de análise e ponto de vista sobre a produção científica e seus respectivos atores. De forma geral, os indicadores bibliométricos são divididos em:

Indicadores de produção científica – construídos a partir da contagem do número de publicações por tipo de documentos, por instituição, área do conhecimento, país, etc. São os produtos da ciência;

Indicadores de citação – construídos pela contagem do número de citações recebidas por uma publicação de artigo, sendo essa a forma mais reconhecida de atribuir crédito ao autor. Reflete, potencialmente, o impacto dos artigos ou assuntos citados;

Indicadores de ligação – construídos pela ocorrência de coautorias, citações e palavras, sendo aplicados na elaboração de mapas de estrutura de conhecimento e redes de relacionamento entre pesquisadores, instituições e países. Emprega técnicas de análise estatística de agrupamentos (Marcelo & Hayashi, 2013, p. 221).

No que tange aos *indicadores de citação*, a análise de citação trata do número de citações recebidas por um documento, autor ou instituição. Na visão de Fujino (2009, p. 216), “a comunidade científica, como manifesto social, sempre atuou em rede, por meio dos vários recursos que foram estabelecendo-se ao longo de sua história”.

Nesse sentido, as citações e cocitações nas publicações científicas caracterizam uma rede social – uma *rede de citações*, em que são conectados autores, correntes teóricas e pontos de vista sobre um determinado campo, área ou temática científica. Para Población, Mugnaini e Ramos (2009, p. 626), trata-se de um “conjunto de nós e laços com relações ilimitadas e híbridas articuladas entre sujeitos, objetos e discursos, que interagem no mundo real e no virtual”. De forma complementar, Currás (2009, p. 61) a percebe como “um conjunto de pontos (atores sociais) vinculados por uma série de relações que cumprem determinadas propriedades. Por possuir uma estrutura e morfologia próprias, permite quantificar as relações e seu consequente tratamento matemático, evidenciando importantes aplicações para a análise e interpretação das condutas sociais”.

Entretanto, as citações não são medidas de qualidade de trabalhos, mas medidas de qualidade definidas socialmente em um determinado contexto, ou seja, por uma determinada comunidade científica (Cole & Cole, 1973 citados por Vanz & Caregnato, 2003, p. 253). Outro ponto a ser considerado refere-se à questão da subjetividade decorrente do elemento humano presente no processo de citação (Bavelas, 1978, p. 159). Apesar destas limitações quanto ao seu uso, a aplicabilidade dos indicadores de citação possibilita tanto a “mensuração das fontes de informação utilizadas, como o tipo de documento, o idioma, e os trabalhos mais citados” (Vanz & Caregnato, 2003, p. 251) como a “avaliação da visibilidade e do impacto de um documento, autor ou instituição, na produção do conhecimento científico da área ou tema sob análise, permitindo a visualização da sua frente de pesquisa” (Grácio,

Oliveira & Matos, 2009). Além disso, os estudos de citação e cocitação no campo da Bibliometria são relevantes,

para o entendimento dos processos de comunicação científica nas diferentes áreas do conhecimento humano. Eles nos permitem mapear um campo emergente ou consolidado, identificar seus principais atores e as relações que estabelecem entre eles e identificar uma série de características do comportamento de uso da informação recuperada. Constituem um importante indicador da atividade científica, pois contribuem para entender a estrutura e o desenvolvimento da ciência e também identificam as regularidades básicas de seu funcionamento (Vanz & Caregnato, 2003, p. 255).

Nesse sentido, por meio das referências bibliográficas da publicação científica, é possível identificar os pesquisadores “cujos conceitos, métodos ou teoria serviram de inspiração, ou foram utilizados pelo autor no desenvolvimento de seu próprio trabalho, estabelecendo assim um processo de referência e citação” (Vanz & Caregnato, 2003, p. 248).

4 PROCEDIMENTOS METODOLÓGICOS

O procedimento de pesquisa adotado para identificação, análise e avaliação dos dados foi pesquisa bibliográfica realizada nas bases de dados *Web of Science* e *Scopus*, sem limitação temporal. Empregou-se como estratégia de busca o termo “*indexing policy*” e sua variante “*indexing policies*”, sendo considerado como tipologia documental apenas artigo de periódico como veículo científico que retrata estudos em estágio avançado.

Na *Web of Science* a busca foi realizada no campo Tópico, no qual são agrupados os campos de título, palavras-chave e resumo. A busca na *Scopus* ocorreu nos campos Título, Resumo e Palavras-chave. Todos os registros foram limitados à área de Ciência da Informação, sendo utilizados como filtros “*Information Science & Library Science*” na *Web of Science* e “*Social Sciences*” na *Scopus*.

O *corpus* documental foi composto por 23 registros, que abarcam o período de 1979 a 2016. Para cada trabalho, foi levantado o rol de referências e desdobradas as autorias múltiplas para contemplar todos os autores citados, ordenando as citações em ordem alfabética. O levantamento dos autores mais citados, realizado por meio das referências bibliográficas dos artigos, totalizou 721 pesquisadores citados para um total de 415 referências.

Para a análise de cocitações, foram considerados os autores que foram citados em pelo menos 3 trabalhos, correspondendo a um total de 24 autores. Para a elaboração da rede de cocitação entre os autores, construiu-se uma matriz quadrada (simétrica) de 24x24, isto é, com os 24 autores e seus relacionamentos, no *software* Excel. Os nomes dos mesmos foram padronizados, utilizando-se primeiro o sobrenome e os demais nomes de forma abreviada.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
2	Fujita, M. S. L.	0	6	3	3	0	4	0	0	0	5	0	4	0	0	0	0	0	4	2	0	0	0	0	0
3	Rubi, M. P.	6	0	3	3	0	4	0	0	0	5	0	4	0	0	0	0	0	4	2	0	0	0	0	0
4	Bocato, V. R. C.	3	3	0	2	0	2	0	0	0	2	0	2	0	0	0	0	0	0	1	0	0	0	0	0
5	GilLeiva, I.	3	3	2	0	0	2	0	0	0	3	0	3	0	0	0	0	0	1	1	0	0	0	0	0
6	Delgado López-Cózar, E.	0	0	0	0	0	0	1	0	1	0	1	0	1	0	0	0	0	0	0	1	0	0	1	0
7	Lancaster, F. W.	4	4	2	2	0	0	0	0	0	3	0	3	0	0	0	0	0	3	2	0	0	0	0	0
8	Jacso, P.	0	0	0	0	1	0	0	0	1	0	1	0	2	0	0	0	0	0	0	1	1	0	1	0
9	Lu, Z.	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	1	0	0	0	0	1	0
10	Cabezas-Clavijo, Á.	0	0	0	0	1	0	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	1	0
11	Carneiro, M. V.	5	5	2	3	0	3	0	0	0	0	0	4	0	0	0	0	0	3	0	0	0	0	0	0
12	Garfeld, E.	0	0	0	0	1	0	1	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0
13	Guimarães, J. A. C.	4	4	2	3	0	3	0	0	0	4	0	0	0	0	0	0	0	2	0	0	0	0	0	0
14	Harzing, A.	0	0	0	0	1	0	2	0	1	0	1	0	0	0	0	0	0	0	1	1	0	1	0	1
15	Kim, W.	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	1	0	1	0
16	Névéol, A.	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	1
17	Pidley, D. D.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	Darmoni, S.	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1
19	Foskett, A. C.	4	4	1	1	0	3	0	0	0	3	0	2	0	0	0	0	0	0	2	0	0	0	0	0
20	Hjerland, B.	2	2	1	1	0	2	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
21	Jiménez Contreras, E.	0	0	0	0	1	0	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0
22	Schreiber, M.	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
23	Thirion, B.	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	1
24	Torres-Salinas, D.	0	0	0	0	1	0	1	0	1	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0
25	Wilbur, W. J.	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0

Figura 1. Recorte da matriz quadrada 23x23 para elaboração da rede de cocitação

Cabe ressaltar que, embora a autocitação seja uma prática para “conectar um trabalho a outro, principalmente trabalhos recentes a trabalhos mais antigos” (Vanz & Caregnato, 2003, p. 253), optamos nesta investigação por retirar a autocitação. Dentre o rol de todas as citações utilizadas para o desenvolvimento das pesquisas, a autocitação foi da ordem de 72 (5,8%).

Na matriz, o valor de cada ocorrência foi registrado, utilizando-se a opção do *software* Excel “Localizar e selecionar”, “Ir para” “especial”, “em branco”, para o preenchimento de todas as células em branco na matriz com zeros. Por meio da matriz padronizada, foi possível gerar a representação gráfica da rede de cocitação dos autores no *software* UCINET, versão 6.627, disponível gratuitamente para *download* na internet. Como parte do pacote UCINET, o *software* NetDraw possibilitou a visualização da rede em duas dimensões.

5 APRESENTAÇÃO E ANÁLISE DOS DADOS

A origem dos pesquisadores mais produtivos na temática política de indexação, com base nos dados coletados, são Brasil e Estados Unidos, com 35% e 26% respectivamente. Os outros pesquisadores são da Espanha, França, Reino Unido, Holanda, Dinamarca, Grécia, Índia e Suíça. Dentre os 23 artigos coletados, o maior número de trabalhos publicados foi no ano de 2012, com 17% do total.

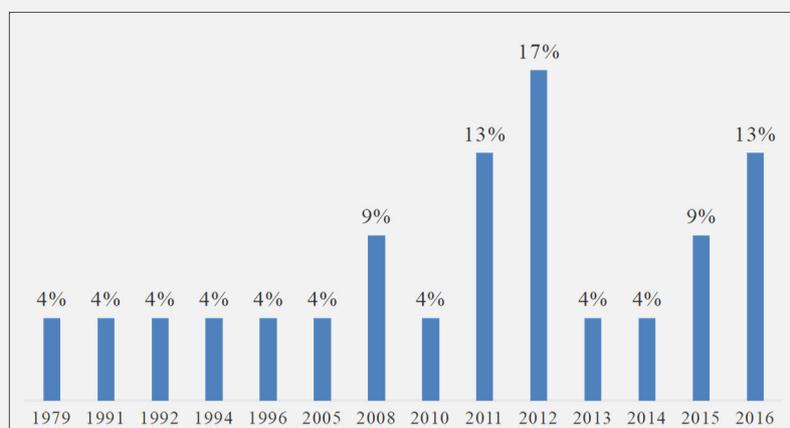


Gráfico 1. Percentual de trabalhos por ano de publicação

Houve uma recuperação bastante limitada de artigos nas bases de dados selecionadas nesta pesquisa. Do conjunto de 46 pesquisadores, os mais produtivos são Fujita, M. S. L. com 7 e Boccato, V R. C. com 3 publicações durante o período analisado (1979-2016). Estas autoras brasileiras figuram também no rol dos autores mais citados, o que indica forte influência como referentes teóricos na temática.

Mediante análise das fontes citadas, obteve-se um conjunto de 415 referências bibliográficas, o que significa uma média de 18 referências/citações por artigo. A análise das referências não contemplou aquelas que se caracterizavam como normas, pessoas jurídicas e instituições. As referências de autoria múltipla foram desdobradas para contemplar todos os autores do documento. Deste grupo, foram citados 464 autores, obtendo-se uma média de 1,1 citação por autor. Evidencia-se que 24 autores foram citados pelo menos 4 vezes em um total de 203 citações, sendo que 379 autores são citados apenas uma vez. Na Tabela 1 apresenta-se os autores mais citados, com os respectivos dados:

Tabela 1. Autores mais citados (até 4 citações)

Autores	N. de citações recebidas	N. de artigos em que foi citado
Fujita, M. S. L. (Brasil)	46	7
Rubi, M. P. (Brasil)	32	6
Boccato, V. R. C. (Brasil)	12	3
Gil leiva, I. (Espanha)	10	4
Lancaster, F. W. (Inglaterra)	9	5
Delgado López-Cózar, E. (Espanha)	9	1
Jacso, P. (Estados Unidos)	7	3
Lu, Z. (Estados Unidos)	6	1
Carneiro, M. V. (Brasil)	5	5
Guimarães, J. A. C. (Brasil)	5	4
Garfield, E. (Estados Unidos)	5	2
Harzing, A.-W. (Inglaterra)	5	2
Ridley, D. D. (Austrália)	5	2
Cabezas-Clavijo, Á. (Espanha)	5	1
Kim, W. (Estados Unidos)	5	1
Névéol, A. (França)	5	1
Foskett, A. C. (Inglaterra)	4	4
Hjørland, B. (Dinamarca)	4	2
Darmoni, S. (França)	4	1
Jiménez Contreras, E. (Espanha)	4	1
Schreiber, M. (Alemanha)	4	1
Thirion, B. (França)	4	1
Torres-Salinas, D. (Espanha)	4	1
Wilbur, W. J. (Estados Unidos)	4	1

Na rede de cocitação, elaborada com os mesmos 23 autores mais citados, os círculos correspondem à frequência com que os autores foram citados e a espessura dos segmentos de reta, à intensidade de cocitação entre os pares (Figura 1):

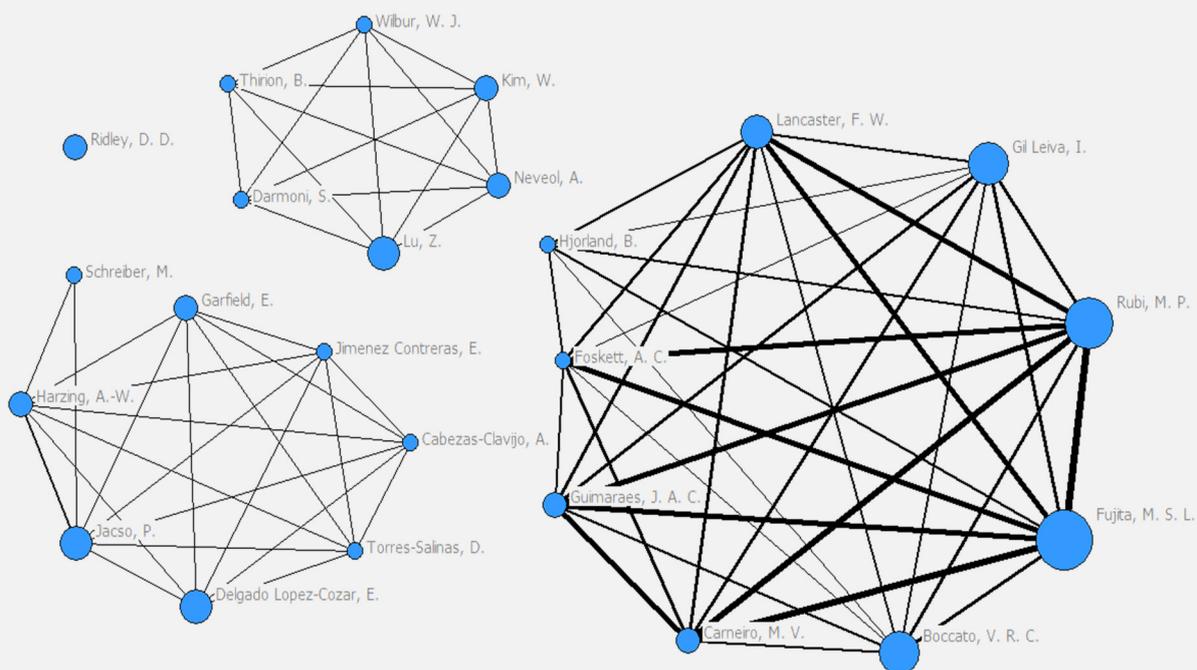


Figura 2: Representação gráfica da rede de cocitação dos autores com até 3 citações

Fonte: Dados da pesquisa.

Em um primeiro momento, percebe-se três distintas sub-redes. Na sub-rede de cocitação mais densa, destacam-se os seguintes autores: Fujita, M. S. L.; Rubi, M. P.; Boccato, V. R. C. e Gil Leiva, I., tanto com alta frequência de citações como de número de cocitações entre seus pares. Desse modo, neste universo de pesquisa, estes autores constituem a frente de pesquisa na temática “Política de indexação”. Outros autores compõem a rede: Lancaster, F. W., Hjørland, B.; Foskett, A. C., Guimarães, J. A. C., Carneiro, M. V. Nesta sub-rede, não houve cocitação apenas entre Hjørland, B. com outros dois autores: Carneiro, M. V. e Guimarães, J. A. C.

Em uma segunda sub-rede aparecem os autores: Delgado López-Cózar, E.; Jacso, P., Harzing, A.-W.; Garfield, E.; Schreiber, M.; Jiménez Contreras, E.; Cabezas-Clavijo e Á Torres-Salinas, D. Nesta, o autor Schreiber, M. aparece cocitado apenas com outros dois autores: Harzing, A.-W. e Jacso, P.

Por fim, em outra sub-rede estão: Lu, Z.; Néveol A.; Kim, W.; Wilbur, W. J.; Thirion, B. e Darmoni, S. Nesta sub-rede, todos os autores tiveram artigos em cocitação, ou seja, estes autores aparecem juntos em um mesmo artigo em todos os trabalhos levantados. Já sem nenhuma cocitação com outros autores, aparece Ridley, D. D., que recebeu 5 citações, porém, em 2 trabalhos que nenhum outro destes 23 autores apareceram.

De forma geral, observa-se três grupos de pesquisadores sobre a temática “Política de Indexação”, pois, embora exista um número significativo de cocitações entre os autores das sub-redes, seus autores não são citados com outros autores além daqueles que compõem a sub-rede das quais fazem parte. Ao calcularmos a densidade da rede para verificarmos o nível de conectividade da rede por meio da medida de coesão entre as conexões presentes na rede (138) e o número total de conexões possíveis (552), obteve-se 0,25, ou seja, a rede apresenta-se com 25% das conexões possíveis, configurando uma rede não muito densa.

6 CONSIDERAÇÕES FINAIS

Os estudos bibliométricos têm sido utilizados em diversos campos, áreas ou temáticas do conhecimento como procedimentos metodológicos que visam a obtenção de indicadores confiáveis de avaliação da atividade científica. A partir da execução desta pesquisa, nota-se pouca produção científica na temática “Política de indexação”, embora a frente de pesquisa seja composta principalmente por autores brasileiros.

Por meio do levantamento e análise de indicadores de citação e cocitação foi possível realizar um mapeamento da temática “Política de indexação” nas bases de dados *Web of Science* e *Scopus*, cujos resultados contribuem para um panorama e avaliação do comportamento deste universo de pesquisa. Em complementação, mostra-se oportuno o avanço de pesquisas direcionadas à política de indexação, com vistas ao avanço e consolidação teórico-metodológico da temática na literatura especializada da Ciência da Informação.

REFERÊNCIAS BIBLIOGRÁFICAS

- Bavelas, J. B. (1978). The social psychology of citations. *Canadian Psychological Review*, Calgary, 19(2), 158-163.
- Carneiro, M. V. (1985). Diretrizes para uma política de indexação. *Revista da Escola de Biblioteconomia da UFMG*, 14(2), 221-241.
- Carrizo Saineo, G. (2000). Hacia um concepto de bibliometría. *Revista Iberoamericana em Ciência de la Información y Documentación*, 1(2), 1-10.
- Cesarino, M. A. N. & Pinto, M. C. M. (1978). Cabeçalho de assunto como linguagem de indexação. *Revista da Escola de Biblioteconomia da UFMG*, Belo Horizonte, 7(2), 268-288.
- Currás, E. Integración vertical de las ciencias aplicada a redes sociales – sociedade de la Información en sus relaciones sistémicas. (2009). In: Población, D. A., Mugnaini, R., & Ramos, L. M. S. V. C. (Orgs.). *Redes sociais e colaborativas: em informação científica*. São Paulo: Angellara.
- Dal’ Evedove, P. R., Tartarotti, R. C. D. E., & Fujita, M. S. L. (2015). *Concepções de política de indexação na visão de bibliotecários brasileiros*. In: Desafios y oportunidades de las Ciencias de la Información y la Documentación en la era digital: actas del VII Encuentro Ibérico EDICIC 2015 (Madrid, 16 y 17 de noviembre de 2015).
- Fujino, A. (et al.). (2009). Políticas públicas de incentivo à formação de redes sociais e colaborativas em ciência e tecnologia. In: Población, D. A., Mugnaini, R., & Ramos, L. M. S. V. C. (Orgs.). *Redes sociais e colaborativas: em informação científica*. São Paulo: Angellara.
- Fujita, M. S. L. (2003). A leitura documentária do indexador: aspectos cognitivos e lingüísticos influentes na formação do leitor profissional. 2003. 321f. Tese (Livre-Docência em Análise Documentária e Linguagens Documentárias Alfabéticas). Marília: Faculdade de Filosofia e Ciências, Universidade Estadual Paulista.

Fujita, M. S. L. & Rubi, M. P. (2006). O ensino de procedimentos de Política de Indexação na perspectiva do conhecimento organizacional: uma proposta de programa para a educação à distância do bibliotecário. *Perspectivas em Ciência da Informação*, Belo Horizonte, 11(1).

Gil Leiva, I. (2008). *Manual de indización: teoría y práctica*. Gijón: Trea.

Gracio, M. C. C., Oliveira, E. F. T., & Matos, G. I. (2009). *Visibilidade dos pesquisadores no tema Estudos Métricos: análise de citação e co-citação nos periódicos do SciELO*. Ibersid, p. 81-86.

Mai, J-E. (1997). The concept of subject: on problems in indexing. *Knowledge Organization for Information Retrieval*. Proceedings of the 6th International Study Conference on Classification Research, 6, 60-67.

Marcelo, J. F., Hayashi, M. C. P. I. (2013). Estudo bibliométrico sobre a produção científica no campo da Sociologia da Ciência. *Informação & Informação*, 18(3), 138-153.

Población, D. A., Mugnaini, R., & Ramos, L. M. S. V. C. (Orgs.). *Redes sociais e colaborativas: em informação científica*. São Paulo: Angellara.

Rubi, M. P. (2012). Os princípios da política de indexação na análise de assunto para catalogação: especificidade, exaustividade, revocação e precisão na perspectiva dos catalogadores e usuários. In: Gil Leiva, I. & Fujita, M. S. L. (Org.). *Política de indexação*. São Paulo: Cultura Acadêmica, 81-93.

Tartarotti, R. C. D. (2014). *Atuação bibliotecária no tratamento temático da informação em unidades informacionais: um estudo comparativo qualitativo-quantitativo*. 2014. 277 f. Dissertação (Mestrado em Ciência, Tecnologia e Sociedade) – Universidade Federal de São Carlos - UFSCar, São Carlos.

Tartarotti, R. C. D. E., & Fujita, M. S. L. (2016). Produção e colaboração científica em Organização e Representação do Conhecimento: análise bibliométrica do GT2 do ENANCIB no período de 2009 a 2014. *Em Questão*, Porto Alegre, 22(3), 136-160.

Tartarotti, R. C. D. E., Dal' Evedove, P. R. & Fujita, M. S. L. (2016). *Política de indexação em ambiente colaborativo das folksonomias: uma reflexão preliminar*. In: XVII ENANCIB - Encontro Nacional de Pesquisa em Ciência da Informação (Salvador, 20 a 25 de novembro de 2016).

Vanz, S. A. S., & Caregnato, S. E. (2003). Estudos de citação: uma ferramenta para entender a comunicação científica. *Em Questão*, Porto Alegre, 9(2), 295-307.



INFORMAÇÃO E COMUNICAÇÃO DOS CONTROLES INTERNOS DAS UNIVERSIDADES FEDERAIS DO NORDESTE DO BRASIL

Wagner Miranda Gomes¹, Zeny Duarte², Dulcinéia Assunção³, Adriana Rosa⁴

¹Instituto de Ciência da Informação da Universidade Federal da Bahia, Brasil (ICI/UFBA),
Doutorando em Ciência da Informação (ICI/UFBA)

²Instituto de Ciência da Informação da Universidade Federal da Bahia, Brasil (ICI/UFBA), Professora Titular (ICI/UFBA) e Pós-doutora em Ciência da Informação (U.Porto, FCT-PT), Zeny.duarte@gmail.com

³Universidade Estadual do Sudoeste da Bahia, Brasil, Especialista em Saúde Coletiva UESB/BRASIL, ducenf@hotmail.com

⁴Instituto de Ciência da Informação da Universidade Federal da Bahia, Brasil (ICI/UFBA), Mestranda em Ciência da Informação (PPGCI)/UFBA, dricasrosa@hotmail.com

RESUMO Este trabalho apresenta os resultados da informação e comunicação dos controles internos das Universidades Federais da Região Nordeste do Brasil, sob a perspectiva do modelo emitido pelo *Committee of Sponsoring Organization (COSO)*. O método consistiu em análise dos relatórios de gestão de 2010 das quatorze Universidades Federais da Região Nordeste. Concluiu-se, dentre outros, que as informações disponíveis dessas universidades não são muito apropriadas, tempestivas, atualizadas, precisas e acessíveis, provocando, assim, uma qualidade pouco satisfatória para permitir ao seu gestor tomar as decisões adequadas.

PALAVRAS-CHAVE *Universidades, Informação e Comunicação, Controles Internos, Controle, Organização da Informação.*

ABSTRACT This paper presents the results of the information and communication of the internal controls of the Federal Universities of the Northeast of Brazil, from the perspective of the model issued by the Committee of Sponsoring Organization (COSO). The method consisted of analysis of the 2010 management reports of the fourteen Federal Universities of the Northeast Region. It was concluded, among others, that the information available from these universities is not very appropriate, timely, up-to-date, accurate and accessible, thus provoking an unsatisfactory quality to allow the manager to make the appropriate decisions.

KEY WORDS *Universities, Information and Communication, Internal Controls, Control, Information Organization.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

No cotidiano das organizações públicas ou privadas, informação, comunicação e conhecimento são termos que têm bastante convergência. Os colaboradores dessas organizações diariamente lidam com fluxos informacionais no seu ambiente de trabalho, que podem apresentar incompletudes de conhecimento, assimetria de informações, e falta de organização e de controle.

A temática do controle (controles internos, governança, gestão de risco, auditoria, correição, corregedoria, prestação de contas, fiscalização, *accountability*) pareceria ser campo de estudos das Ciências Econômicas, Ciências Contábeis, Direito ou Ciência da Administração, mas também se faz presente na Ciência da Informação (CI).

Saracevic (1996) salienta que a Ciência da Informação é um campo interdisciplinar e permeia não só por questões científicas, mas atua de forma prática. Volta-se para os problemas do registro da informação e da efetiva comunicação do conhecimento entre os humanos. Nesse sentido, é possível perceber o caráter convergente entre informação, comunicação e conhecimento.

Ainda na seara da CI, Lopes, Valentim e Almeida Junior (2011) trouxeram como tema a governança corporativa, que pode ser entendida como o conjunto de mecanismos de gestão, que através de controles internos e externos visam reduzir a distância entre acionistas minoritários e o poder de controle da empresa. Nesse contexto, a gestão da informação, a organização do conhecimento, a mediação da informação e a sua disseminação se fazem necessárias, visto que o estabelecimento da boa comunicação, rápida e clara, possibilita desenvolver um ambiente de controle favorável a um clima de confiança na relação da empresa com a sociedade, credores, colaboradores e investidores. E isso é extremamente importante para os controles internos.

Para D'Avila e Oliveira (2002), não se pode falar em controles internos sem citar o *The Committee of Sponsoring Organizations* (COSO). Em 1992, o COSO produziu um documento intitulado *Internal Control – Integrated Framework*, que define os controles internos como um processo conduzido pelo conselho de administração, pela diretoria, e por outras pessoas da organização, que visa fornecer, com segurança razoável, a consecução de objetivos nas seguintes categorias: confiabilidade das informações; cumprimento de leis, normas e regulamentos (*compliance*); e eficiência de operações nas organizações.

Para o COSO, o controle interno da organização pode ser representado pelos componentes: ambiente de controle, avaliação de riscos, procedimento de controle, informação e comunicação, e monitoramento. Para fins desta comunicação, abordar-se-á o componente informação e comunicação.

Devido à abrangência do documento, organizações internacionais de orientações sobre controles internos revisaram suas publicações no sentido de incorporar os conceitos trazidos pelo COSO, a exemplo da *United States General Accounting Office*, do Banco Mundial e da *International Organization of Supreme Audit Institutions* (INTOSAI). (TCU, 2009).

No Brasil, o Controle Externo é exercido pelo Poder Legislativo, com o auxílio do Tribunal de Contas da União (TCU), que é o órgão técnico, e é a Instituição máxima de fiscalização do Setor Público brasileiro. O TCU representa o Brasil na INTOSAI. A partir de 2010, o TCU passou a cobrar, no relatório de gestão dos Órgãos da Administração Pública Federal brasileira, informações sobre os controles internos, com base no COSO.

As Universidades Federais fazem parte da Administração Pública do Brasil, e devem seguir aos princípios constitucionais, bem como atender normas definidas pelos órgãos de controle. Elas possuem em sua estrutura Bibliotecas, Arquivos e Museus, e formam os profissionais que desempenham atividades e práticas relacionadas à organização da informação e do conhecimento registrado e socializado.

Nesse sentido, em 2011, em nossa pesquisa de mestrado, foi investigado o seguinte problema: a estrutura dos controles internos das Universidades Federais da Região Nordeste é eficiente sob a perspectiva do modelo COSO? Foi idealizado como objetivo geral a avaliação da eficiência da estrutura dos controles internos dessas universidades, e como objetivos específicos: a mensuração do ambiente de controle; da avaliação de risco; dos procedimentos de controle; da informação e comunicação e do monitoramento. Analisando, para isso, os relatórios de gestão do exercício de 2010 dessas universidades.

Portanto, este trabalho tem como objetivo apresentar a análise dos resultados do componente informação e comunicação da estrutura dos controles internos das Universidades Federais da Região Nordeste do Brasil, sob a perspectiva do modelo COSO.

METODOLOGIA

Ao longo da revisão da literatura buscamos os entendimentos necessários para responder ao problema, tendo como tema: ‘A eficiência da estrutura dos controles internos das Universidades Federais da Região Nordeste sob a perspectiva do modelo do COSO’.

O primeiro entendimento é que ser eficiente na Administração Pública significa alcançar o melhor resultado possível de uma operação, considerando o contexto e os recursos disponíveis, analisando o rendimento do serviço prestado em relação ao seu custo e comparando-o a um valor de referência, atendendo, assim, às demandas da sociedade. O segundo é que o controle da Administração Pública visa aperfeiçoar as crescentes funções públicas do Estado diante de recursos cada vez mais limitados, valendo-se para tal de instrumentos como o controle externo, o controle interno, as técnicas de controle, a avaliação de desempenho e a prestação de contas. O terceiro é que o referencial do COSO permite uma forma objetiva de avaliar a estrutura dos controles internos da Administração Pública, valendo-se para tal dos componentes: ambiente de controle, avaliação de risco, procedimento de controle, informação e comunicação, e monitoramento. Por último, entendeu-se que as Universidades Federais, notadamente as da Região Nordeste, representadas pela sua grande importância para o conjunto da sociedade, devem demonstrar competência, integridade, ética e transparência ao público na gestão de recursos que pertencem a toda sociedade.

Esses entendimentos se constituíram no delineamento da pesquisa como verdadeiros elementos apreendidos da revisão da literatura, presentes dentro do referencial do COSO, que permitiram chegar às conclusões.

A pesquisa também utilizou as informações, sobre a educação superior no Brasil, disponibilizadas no censo realizado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), em 2009. O INEP é uma Instituição autárquica federal, vinculada ao Ministério da Educação (MEC). Sua missão é promover estudos, pesquisas e avaliações sobre o Sistema Educacional Brasileiro, com o objetivo, dentre outros, de produzir informações claras e confiáveis aos gestores, pesquisadores, educadores e público em geral.

O referido Instituto realiza o censo da educação superior do Brasil, com o objetivo de oferecer informações, à comunidade acadêmica e à sociedade em geral, detalhadas sobre a situação do setor educacional brasileiro. O censo da educação superior reúne informações regionais, dentre outras, sobre as Universidades Públicas Federais, seus cursos presenciais de graduação, cursos e vagas oferecidos, matrículas realizadas, dentre outras. O censo da educação superior é realizado por meio de um questionário eletrônico, no qual as Instituições de Ensino Superior (IFES) respondem sobre sua estrutura e seus cursos. O INEP, então, faz a análise dos dados do censo, para verificar a consistência das informações. Após essa fase de conferência, em colaboração com os pesquisadores institucionais, o censo é finalizado. Daí, os dados são divulgados e publicados e passam a fazer parte das estatísticas oficiais. Segundo essas informações de 2009, o Brasil conta com 55 Universidades Públicas Federais, 3.868 cursos presenciais de graduação, oferecendo 186.984 vagas, além de realizar 696.693 matrículas nesses cursos. A Tabela 1 sintetiza esses dados por Região do Brasil.

Tabela 1. Dados sobre as Universidades Federais da Região Nordeste do Brasil

Regiões	Nº Universidades	Cursos	Vagas	Matrículas
Norte	8	627	19.387	86.827
Nordeste	14	1.003	59.618	221.034
Sudeste	19	1.071	54.746	202.763
Sul	9	706	32.397	117.131
Centro-oeste	5	461	20.836	68.938
Total	55	3.868	186.984	696.693

Fonte: INEP, 2009.

A maior quantidade de universidades se encontra na Região Sudeste do Brasil, seguida pela Região Nordeste. A Região Sudeste lidera quando se trata do número de cursos de graduação presenciais oferecidos pelas Universidades Federais do Brasil, sendo 1.071 cursos contra 1.003 da Região Nordeste – que ocupa a segunda colocação. Entretanto, quando se trata vagas oferecidas e matrículas realizadas nesses cursos, a Região Nordeste lidera com cerca de cinco mil vagas e dezoito mil matrículas a mais em relação à Região Sudeste. Assim, as Universidades Federais da Região Nordeste alimentam a sociedade com um maior número de vagas oferecidas e de matrículas realizadas, além disso, é sabido que a Região Nordeste é uma das regiões mais carentes do país. Portanto, essas informações foram determinantes na escolha das 14 Universidades Federais da Região Nordeste como amostra da presente pesquisa.

Para realização deste trabalho, foram utilizados os dados e informações presentes nos questionários dos relatórios de gestão relativos ao exercício de 2010 das Universidades Federais da região Nordeste do Brasil. O questionário - instrumento de coleta de dados – é constituído por uma escala Likert e consta na Portaria nº 277/2010 do Tribunal de Contas da União (TCU). O referido questionário trata sobre o funcionamento da estrutura dos controles internos da Administração Pública Federal brasileira, e é baseado no referencial do COSO. Tal questionário traz ainda afirmativas, as quais o respondente (neste caso, cada universidade federal) tem de marcar as opções ‘1 Totalmente inválida’, ‘2 Parcialmente inválida’, ‘3 Neutra’, ‘4 Parcialmente válida’ e ‘5 Totalmente válida’.

A estruturação dos elementos para a compreensão das conexões entre as afirmativas e o componente de análise, bem como a eficiência deste com a estrutura dos controles internos, constitui-se no recorte do modelo de análise para os dados do componente Informação e Comunicação, conforme Quadro 1.

Quadro 1 - Esquema das conexões entre as assertivas, o componente de análise e a eficiência

Eficiência	Componente de análise	Afirmativas
Controles Internos	Informação e Comunicação	1. A informação relevante é devidamente identificada, documentada, armazenada e comunicada tempestivamente às pessoas adequadas.
		2. As informações consideradas relevantes são dotadas de qualidade suficiente para permitir ao gestor tomar as decisões apropriadas.
		3. A informação disponível é apropriada, tempestiva, atual, precisa e acessível.
		4. A Informação divulgada internamente atende às expectativas dos diversos grupos e indivíduos, contribuindo para a execução das responsabilidades de forma eficaz.
		5. A comunicação das informações perpassa todos os níveis hierárquicos, em todas as direções, por todos os seus componentes e por toda a sua estrutura.

Fonte: Adaptado TCU, 2010.

Para o tratamento dos dados, foi utilizada a planilha Excel. E foi adotado como critério de aferição de eficiência um percentual igual ou superior a 60% (índices da coluna “Totalmente válida”), abaixo desse percentual a estrutura dos controles internos das Universidades pesquisadas foi considerada pouco eficiente. A análise dos resultados na presente comunicação é, portanto, relativa ao componente Informação e Comunicação da estrutura de controles internos das Universidades Federais da Região Nordeste do Brasil.

RESULTADOS

A Tabela 2 apresenta os dados referentes às afirmativas do componente informação e comunicação da estrutura dos controles internos das Universidades Federais da Região Nordeste do Brasil.

Tabela 2. Distribuição percentual do componente informação e comunicação

Afirmativa	Totalmente inválida (1)	Parcialmente inválida (2)	Neutra (3)	Parcialmente válida (4)	Totalmente válida (5)	Total	
						Resp.	Não resp.
1. A informação relevante é devidamente identificada, documentada, armazenada e comunicada tempestivamente às pessoas adequadas.	0,0%	0,0%	14,2%	42,9%	42,9%	14	0
2. As informações consideradas relevantes são dotadas de qualidade suficiente para permitir ao gestor tomar as decisões apropriadas.	0,0%	0,0%	14,3%	50,0%	35,7%	14	0
3. A informação disponível é apropriada, tempestiva, atual, precisa e acessível.	0,0%	28,6%	7,1%	42,9%	21,4%	14	0
4. A Informação divulgada internamente atende às expectativas dos diversos grupos e indivíduos, contribuindo para a execução das responsabilidades de forma eficaz.	0,0%	21,4%	28,6%	35,7%	14,3%	14	0

5. A comunicação das informações perpassa todos os níveis hierárquicos, em todas as direções, por todos os seus componentes e por toda a sua estrutura.	0,0%	35,7%	28,6%	14,3%	21,4%	14	0
Total de participantes	14	14	14	14	14	14	

Fonte: Elaborado pelo autor.

A Tabela 3 apresenta os dados consolidados do componente informação e comunicação da estrutura dos controles internos das Universidades Federais da Região Nordeste do Brasil.

Tabela 3. Distribuição percentual do componente informação e comunicação

Componentes	Totalmente inválida (1)	Parcialmente Inválida (2)	Neutra (3)	Parcialmente Válida (4)	Totalmente válida (5)	Total de respostas	
						Resp.	Não resp.
Informação e Comunicação	0,0%	17,1%	18,6%	37,1%	27,1%	70	0
Total de participantes	14	14	14	14	14	14	

Fonte: Elaborado pelo autor.

Quanto ao componente informação e comunicação é necessário tecer algumas observações sobre seus aspectos, pois trazem resultados interessantes. A primeira delas trata da identificação, documentação, armazenamento e comunicação tempestiva, às pessoas adequadas, das informações relevantes para cada uma das universidades pesquisadas.

A informação relevante que é devidamente identificada, documentada, armazenada e comunicada tempestivamente às pessoas adequadas existe em 42,9% das Universidades Federais da Região Nordeste. Esse tratamento da informação foi validado parcialmente por 42,9% delas, enquanto 14,2% não se posicionaram. É um foco de fragilidade que, em apenas 42,9% das universidades pesquisadas, a informação relevante seja devidamente identificada, documentada, armazenada e comunicada tempestivamente às pessoas adequadas. Alarmante é que esse processo se dá de forma parcial em quase metade delas (42,9%), o que parece significar que ou as informações são pouco relevantes para as universidades, ou estas não estão devidamente identificadas, documentadas, armazenadas e comunicadas tempestivamente às pessoas adequadas. Ainda preocupante é que parece não existir a identificação, documentação, armazenamento e comunicação tempestiva às pessoas adequadas das informações relevantes em 14,2% delas.

Para Davenport e Prusak (2003), o conhecimento, nas organizações, costumar estar embutido não só em documentos, como também em processos, rotinas, procedimentos, práticas e normas organizacionais. Para que os funcionários realizem suas atividades de forma eficiente é necessário que conheçam suas atribuições e responsabilidades. E, para isso, é importante ter informação relevante. A informação sobre as transações e eventos, por exemplo, para ser relevante e confiável necessita que seu registro, classificação e organização sejam feitos de forma adequada e imediata. Logo, a informação precisa ser identificada, documentada e armazenada para que possa ser comunicada tempestivamente às pessoas adequadas. Tudo isso permite proporcionar uma comunicação dentro da organização com a qualidade adequada, possibilitando, assim, uma melhora no processo de construção do conhecimento.

E os dados revelam que isso não vem sendo atingido eficientemente nas Universidades Federais da Região Nordeste do Brasil, pois se mostrou presente em apenas 42,9% delas. Talvez algum alento esteja

no fato de que em algumas universidades pesquisadas dão algum tratamento as suas informações relevantes, mesmo que de forma parcial (42,9%). Nesse caso, se somarmos os dois índices totaliza 85,8%, esse percentual representa aquelas universidades que de alguma forma dão relevância as suas informações. Porém, essa situação necessita de ações de melhoria.

O percentual de 85,8% das Universidades pesquisadas que dão alguma relevância a sua informação talvez ocorra, pelo tratamento da informação, nas suas fases de identificação e armazenagem. Ou seja, pode haver fragilidade na comunicação e na construção do conhecimento coletivo, e essa fragilidade pode ser atribuída a uma possível cultura de não disseminação das informações relevantes dentro das universidades pesquisadas ou mesmo dificuldades nos meios de comunicação.

As observações seguintes são sobre a qualidade que as informações relevantes devem ter para permitir ao gestor tomar as decisões apropriadas; e, que a informação disponível seja apropriada, tempestiva, atual, precisa e acessível.

As informações consideradas relevantes que são dotadas de qualidade suficiente para permitir ao gestor tomar as decisões apropriadas estão presentes em 35,7% das Universidades Federais da região Nordeste. Essa qualidade das informações para a tomada de decisões foi validada parcialmente por 50% delas, enquanto 14,3% se manifestaram de forma neutra. O fato de apenas 35,7% das Universidades pesquisadas terem as informações relevantes dotadas de qualidade suficiente para permitir ao seu gestor tomar as decisões apropriadas merece cuidados. Pois em metade delas isso ocorre de forma parcial, o que indica que ou as informações relevantes são dotadas de pouca qualidade, ou elas não permitem ao gestor tomar decisões apropriadas. Certo é que em 14,3% delas não existem informações dotadas de qualidade suficiente para permitir ao seu gestor tomar decisões apropriadas.

A informação disponível é apropriada, é tempestiva, é atual, é precisa e acessível em apenas 21,4% das Universidades Federais da região Nordeste. Para 42,9%, isso ocorre de forma parcial, enquanto 7,1% se posicionaram de forma neutra. Além disso, 28,6% delas invalidaram parcialmente a existência de tais qualidades para as informações disponíveis. Da mesma forma que nas questões referentes à informação dentro das universidades pesquisadas, o fato de em apenas 21,4% delas existir informações apropriadas, tempestivas, atualizadas, precisas e acessíveis inspira cuidados. Pois em mais da metade (71,5%, somando os índices de 42,9% e 28,6%) delas isso ocorre de forma parcial, o que indica que ou as informações são pouco acessíveis, ou elas estão desatualizadas. Outro fato é que o índice de universidades onde parece que não há informação tempestiva, atualizada, precisa e acessível alcança 7,1%.

Para INTOSAI (2004), os controles internos estão relacionados à eficiência das operações, cumprimentos de procedimentos e normas, além da obrigação de prestar contas, que podem ser evidenciada pela disponibilização de informações através de relatórios financeiros e não financeiros. Uma informação relevante e dotada de qualidade suficiente, atrelado a um adequado processo de organização da informação e do conhecimento registrado, permite a administração pública tomar decisões coerentes. Deste modo, criar mecanismo para melhorar esse fundamento reflete diretamente na qualidade das informações e, conseqüentemente, nas decisões tomadas em benefício do alcance dos objetivos. Assim, a formalização e documentação das informações de qualidade permitem ao gestor tomar decisões apropriadas, além de contribuir para a promoção da obrigação de *accountability*.

Os dados apontam que isso vem ocorrendo de uma forma pouco eficiente nas Universidades Federais da Região Nordeste do Brasil, pois apenas 35,7% delas consideram que as informações relevantes são

dotadas de qualidade suficientes que permitam ao gestor tomar decisões apropriadas. O mesmo ocorre com a tempestividade, atualização, precisão e acessibilidade da informação disponível, ou seja, essa eficiência no tratamento da informação só está totalmente presente em 21,4% dos respondentes. Para Choo (2003), uma organização pode melhorar ainda mais esse processo tirando vantagem do conhecimento especializado de profissionais da informação. Assim, a informação que é divulgada internamente passa a ter melhores condições de atender as expectativas dos diversos grupos e indivíduos.

A quarta observação diz respeito a se a informação divulgada internamente atende às expectativas de todos, e contribui para a execução das responsabilidades de forma eficaz; e, a quinta e última se a comunicação das informações perpassa todos os níveis hierárquicos, em todas as direções, por todos os componentes e por toda estrutura das universidades pesquisadas.

A divulgação interna de informações atende às expectativas dos diversos grupos e indivíduos, contribuindo para a execução das responsabilidades de forma eficaz, em apenas 14,3% das Universidades Federais da Região Nordeste. Para 35,7%, isso ocorre de forma parcial, enquanto 28,6% se posicionaram de forma neutra. Além disso, 21,4% das universidades invalidaram parcialmente a existência de divulgação interna de informações que atende às expectativas dos diversos grupos e indivíduos.

O fato de apenas 14,3% das universidades pesquisadas declararem possuir divulgação interna de informações que atendam às expectativas dos diversos grupos e indivíduos e contribuam para a execução das responsabilidades de forma eficaz inspira cuidados. Pois em mais da metade (57,1%, somando os índices de 35,7% e 21,4%) delas isso ocorre de forma parcial, o que indica que ou a divulgação interna de informações atende pouco às expectativas dos diversos grupos e indivíduos, ou estas contribuem pouco para a execução das responsabilidades de forma eficaz. Outro fato é que o índice de universidades onde parece não existir divulgação interna, ou contribuição, alcança 28,6%.

Se a comunicação das informações perpassa todos os níveis hierárquicos das Universidades Federais da região Nordeste, em todas as direções, por todos os seus componentes e por toda a sua estrutura foi validada totalmente por apenas 21,4% delas. Essa amplitude da comunicação foi validada parcialmente por 14,3% das universidades, enquanto 28,6% se manifestaram de forma neutra. Para 35,7%, a amplitude da comunicação foi invalida parcialmente.

Chama a atenção que apenas 21,4% das universidades pesquisadas possuam comunicação de informações que perpassem todos os níveis hierárquicos, em todas as direções, por todos os seus componentes e por toda a sua estrutura. Porque em metade delas (somando os índices de 14,3% e 35,7%), a existência dessa amplitude da comunicação se dá de forma pouco eficiente, o que indica que ou a comunicação das informações não perpassa todos os níveis hierárquicos, ou não perpassa em todas as direções, ou não atingem todos os componentes nem toda a estrutura. Acrescente-se a isto que em 28,6% delas parece não existir nenhum destes elementos.

Para Schmidt e Santos (2009), a informação e o conhecimento registrado têm estreita relação com a comunicação e devem atender às expectativas de vários grupos e indivíduos da organização, permitindo-lhes executar suas responsabilidades de forma eficaz. E parece que isso não vem ocorrendo eficientemente nas Universidades Federais da Região Nordeste do Brasil, pois somente para 14,3% delas, a informação divulgada internamente atende às expectativas dos diversos grupos de indivíduos das universidades pesquisadas e contribui para a execução das suas responsabilidades.

A comunicação é um elemento essencial para os controles internos, é o fluxo de informações e de conhecimento dentro da organização e, segundo Schmidt e Santos (2009), esse fluxo deve fluir em todas as direções da entidade, em todos os níveis hierárquicos, dos superiores aos inferiores e vice-versa. A informação, a comunicação e o conhecimento devem, assim, auxiliar na eficiência das operações das instituições, perpassando por toda a organização, fluir para todos os lados, por todos os componentes, isto é, pela estrutura inteira. E isso parece que também não vem ocorrendo eficientemente nas universidades pesquisadas, pois somente para 21,4% delas a comunicação das informações perpassa todos os níveis hierárquicos, em todas as direções, por todos os componentes e por toda estrutura.

Finalmente, tudo o que até aqui foi discorrido explica a pouca eficiência do componente informação e comunicação, que alcançou um índice de apenas 27,1% das Universidades Federais da Região Nordeste do Brasil, conforme a Tabela 3. Isso pode ser decorrente da pouca eficiência dos aspectos referentes à organização da informação e do conhecimento registrado, bem como do processo não muito satisfatório da comunicação dentro das universidades pesquisadas.

CONCLUSÃO

O presente trabalho revelou que as informações das Universidades Federais da Região Nordeste do Brasil são afetadas pela pouca eficiência na sua identificação, documentação, armazenamento e transmissão tempestiva às pessoas adequadas. Provocando uma qualidade pouco satisfatória para permitir ao seu gestor tomar as decisões adequadas, que significa dizer que as informações não são muito apropriadas, tempestivas, atualizadas, precisas e acessíveis.

Isso tudo gera um cenário em que poucas Universidades Federais da Região Nordeste possuam divulgação interna de informações que atendam às expectativas dos diversos grupos e indivíduos e contribuam para a execução das responsabilidades de forma eficiente e que, poucas delas possuam comunicação de informações que perpassem todos os níveis hierárquicos, em todas as direções, por todos os seus componentes e por toda a sua estrutura.

Tais resultados se por um lado indicam fragilidades, por outro, indicam também oportunidades de melhoria nas práticas relacionadas à organização da informação e do conhecimento registrado e socializado e, conseqüentemente, de melhoria na eficiência dos controles internos das instituições pesquisadas.

Finalmente, há para o presente trabalho alguns limites. A presente comunicação se restringiu à análise, apenas, do componente Informação e Comunicação dos controles internos das Universidades pesquisadas, baseada nos questionários por elas respondidos em 2010. Há, portanto, necessidade de novas investigações que aprofundem, assegurem e ampliem os conhecimentos da informação e comunicação dos controles internos das Universidades Federais do Nordeste, bem como de toda Administração Pública do Brasil.

REFERÊNCIAS BIBLIOGRÁFICAS

Choo, C. W. (2003). *A organização do conhecimento: como as organizações usam a informação para criar significado, construir conhecimento e tomar decisões*. São Paulo, Brasil: SENAC.

Constituição da República Federativa do Brasil de 1988.

Controladoria Geral da União. (2011). *Relatório de gestão das Universidades Federais – exercício 2010*. Brasília, Brasil: CGU.

Davenport, T. H., Prusak, L. (2003). *Conhecimento Empresarial; como as organizações gerenciam o seu capital intelectual*. Rio de Janeiro, Brasil: Elsevier.

D'Avila, M. Z., Oliveira, M. A. M. de. (2002). *Conceitos e técnicas de controles internos de organizações*. São Paulo, Brasil: Nobel.

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2011). *Censo da Educação Superior 2009*. Brasília, Brasil: INEP.

International Organization of Supreme Audit Institutions. (2004). *Guía para las normas de control interno del sector público*. Bruxelas, Bélgica: INTOSAI.

Lopes, E. C.; Valentim, M. L. P.; Almeida Júnior, O. F. de. (2011). Disseminação da informação em empresas de capital aberto e os processos de mediação da informação. *Em Questão*, 17, 77-89. Porto Alegre, Brasil.

Saracevic, T. Ciência da informação: origem, evolução e relações. *Perspectivas em Ciência da Informação*, 1(1), 41-62. Belo Horizonte, Brasil.

Schmidt, P., Santos, J. L. (2009). *Fundamentos de controladoria*. São Paulo, Brasil: Atlas.

Tribunal de Contas da União. (2009). *Critérios gerais de controles internos na administração pública: um estudo das normas disciplinadoras em diversos países*. Brasília, Brasil: TCU.

Tribunal de Contas da União. (2010). *Portaria 277, de 07 de dezembro de 2010*. Brasília, Brasil: TCU.



A FUNÇÃO DO REGISTRO DE AUTORIDADE DE ASSUNTO NA CONSTRUÇÃO E USO DE LINGUAGENS DE INDEXAÇÃO PARA CATÁLOGOS ONLINE

Mariângela Spotti Lopes Fujita¹, Luciana Beatriz Piovezan², Natany de Souza dos Santos³

¹Professora titular; Universidade Estadual Paulista-Unesp, 0000-0002-8239-7114, fujita@marilia.unesp.br

²Doutoranda em Ciência da Informação; Universidade Estadual Paulista – Unesp, nanipiovezan@gmail.com,

³Graduada em Biblioteconomia; Universidade Estadual Paulista – Unesp, souza.nthy@yahoo.com.br

RESUMO Considerando a importância do controle de vocabulário para as atividades de organização e representação do conhecimento em catálogos online, materializado no uso de linguagens de indexação no processo de indexação de assuntos, discute-se a construção de linguagens de indexação contemplando a formalização dos termos autorizados por meio do registro de autoridade de assunto, segundo o formato MARC autoridades. Com a proposta de verificar a composição do registro de autoridade de assunto, o objetivo da pesquisa é investigar quais funções estes registros realizam na construção e uso de linguagens de indexação em catálogos online. Conclui-se que a construção da linguagem de indexação com o uso do recurso do registro de autoridade de assunto é adequada e coerente na medida em que pretende assegurar a qualidade do controle de vocabulário e a mediação do processo de pesquisa em catálogos online pretendidas pelos sistemas de organização do conhecimento.

PALAVRAS-CHAVE *Linguagem de indexação, controle de vocabulário, registro de autoridade de assunto.*

ABSTRACT Vocabulary control in knowledge organization and representation for online catalogs using indexing languages in the subject indexing process is a fundamental requirement. On account of this, this paper presents a discussion on the construction of indexing languages through the formalization of the terms authorized by subject authority records. In order to verify the composition of a subject authority record, a proposal is made to investigate its functions in the construction and use of indexing languages for online catalogs. It is concluded that the construction of an indexing language using subject authority records is adequate and coherent in the sense that this procedure ensures the quality of vocabulary control and the mediation of the research process in the online catalog as required by knowledge organization systems.

KEY WORDS *Indexing language, vocabulary control, subject authority record*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

O uso de uma linguagem de indexação adequada e atualizada para a representação dos conceitos significativos de um documento no processo de indexação é essencial para a garantia da recuperação da informação em um sistema. A construção de uma linguagem deste porte é um processo complexo que requer que se respeitem os princípios teóricos estabelecidos na literatura de Organização do Conhecimento (Campos, 2001, Dodebei, 2002, Currás, 2005, Moreira & Moura, 2006, Pereira & Bufrem, 2005) e que seja norteado pelas diretivas estabelecidas nas normas internacionais de construção (International Organization for Standardization, 2011).

Entretanto, com os desenvolvimentos tecnológicos ocorridos nas últimas décadas, este processo de construção de linguagens de indexação tem sido facilitado pelo desenvolvimento de softwares que já podem ser interligados às bases de dados bibliográficos e aos catálogos de acesso público. Esse recurso representa não apenas uma vantagem ao profissional de informação, que constrói a linguagem e a utiliza nos processos de tratamento temático, mas também aos usuários que podem ter à sua disposição uma ferramenta integrada que lhe permita efetivamente realizar suas pesquisas com a mediação da linguagem de indexação, realizando buscas de modo mais intuitivo e, assim, obtendo resultados mais condizentes com seus interesses.

Porém, para que estes desenvolvimentos ocorram, um elemento se coloca como chave no processo de construção de linguagens de indexação, o registro de autoridade de assunto. O registro de autoridade de assunto é o responsável pelo controle de vocabulário automático no sistema, ao incluir para cada termo a descrição do seu significado e das relações com outros termos, sejam hierárquicos ou associativos e, com isso realizar adequações e correções automaticamente referentes a cada termo quando de seu uso no campo de assunto do registro bibliográfico no formato MARC21.

Dessa forma, o registro de autoridade de assunto é elemento intrínseco das linguagens de indexação, portanto neste trabalho trataremos dos registros de autoridade de assunto como os registros da linguagem de indexação, ou seja, os registros referentes aos termos que compõem a linguagem de indexação e suas relações.

Nesse sentido, este estudo pretende abordar a composição do registro de autoridade de assunto no formato MARC Autoridades, com objetivo de verificar sua função na construção de linguagens de indexação em catálogos online e sua importância na visualização da linguagem de indexação pelo usuário que consulta o catálogo. Para tanto, iniciaremos apresentando a concepção de linguagem de indexação como instrumento de organização do conhecimento que explicita as relações entre conceitos constituindo-se em uma ferramenta para a representação e a recuperação da informação.

LINGUAGENS DE INDEXAÇÃO PARA CATÁLOGOS ONLINE

As linguagens de indexação são instrumentos de controle de vocabulário que objetivam a representação e a recuperação de informação, utilizados na indexação, para explicitar o conteúdo de documentos e representá-los em sistemas de recuperação da informação, e na expressão das necessidades de pesquisa de usuários, para formulação de estratégias de busca.

A construção das linguagens de indexação tem por fundamento os princípios teóricos e metodológicos provenientes da literatura sobre Linguística, Terminologia, Teoria do Conceito, Teoria da Classificação e Recuperação da Informação.

O processo de construção das linguagens de indexação também é guiado pelas diretrizes estabelecidas pelas normas internacionais de construção, cuja evolução levou às atuais ANSI/NISO Z39.19 *Guidelines for the construction, format, and management of monolingual controlled vocabularies* (American National Standards Institute/ National Information Standards Organization, 2005) no contexto norte-americano e a ISO 25964 *Information and documentation: part 1: thesauri and interoperability with other vocabularies* (2011) e *Information and documentation – thesauri and interoperability with other vocabularies - part 2: Interoperability with other vocabularies* (2013).

A construção e uso de linguagens de indexação tem por objetivos o controle de vocabulário, no que tange à eliminação da ambiguidade, controle de sinonímia presentes na linguagem natural e a normalização das variações de gênero e número; oferecer auxílio ao indexador na seleção de termos que melhor representem o assunto dos documentos; a mediação entre o conhecimento armazenado em sistemas e os usuários e facilitar a seleção de termos para formulação de estratégias de busca.

Para atingir a estes objetivos, além da seleção e normalização dos conceitos, são estabelecidas relações de tipo lógico-hierárquicas, tais como as relações de gênero/espécie e todo/parte e relações semânticas, tais como as relações de equivalência e associativas.

A linguagem de indexação que melhor desenvolveu-se com relação à sua estrutura e funções e que se considera mais adequada para a representação e recuperação da informação dotada de precisão e relevância nos catálogos online é o tesouro (Cervantes & Fujita, 2012). O tesouro é compreendido no âmbito da norma ISO 25964 como um “vocabulário controlado e estruturado no qual conceitos são representados por termos, organizados para que as relações entre conceitos sejam feitas explícitas e termos preferidos estejam acompanhados por entradas para sinônimos ou quase-sinônimos” (International Organization for Standardization, 2011, p. 12, tradução nossa). Considera-se o tesouro como uma linguagem de indexação cuja prerrogativa é a sua composição que destaca as relações existentes entre conceitos e termos.

Esta concepção de tesouro está baseada na abordagem teórico-metodológica denominada de “abordagem conceitual” que enfatiza a relação entre a teoria do conceito de Dahlberg (1978) e a teoria da classificação, especialmente a herança de Ranganathan, para a construção de tesouros (Campos & Gomes, 2006).

Neste processo, a abordagem conceitual de construção do tesouro conceitual adota como unidade de trabalho o conceito e não o assunto - como nos cabeçalhos de assunto, ou o termo - como no tesouro documental. Para Dahlberg (1978) o conceito é entendido como “unidades de conhecimento, identificadas por meio de enunciados verdadeiros acerca de determinado objeto e representadas por uma forma verbal”. A abordagem conceitual pressupõe, portanto, o conceito como elemento central da linguagem de indexação, de onde se estabelece a partir do objeto/referente o melhor correspondente para sua designação, na forma do termo, e também as relações entre conceitos e termos. Ao abordar esta questão com relação à norma ISO 25964 (2011), Dextre Clarke e Zeng (2012) afirmam que por conceito e designação serem inseparáveis, o único modo de um tesouro apresentar sua listagem alfabética de conceitos é com o uso de termos, o que torna inevitável o uso de termos em um tesouro, ou seja, um tesouro irá conter conceitos e termos que os designam.

REGISTRO DE AUTORIDADE DE ASSUNTO: ANTECEDENTES E PERSPECTIVAS

Para compreender a função do registro de autoridade de assunto na construção de linguagens de indexação, é preciso recuperar o seu contexto de desenvolvimento no âmbito da catalogação descritiva e de assunto.

O processo de criação de registros de autoridades, pautados nos princípios do controle de autoridade, remonta aos primórdios da padronização do trabalho bibliotecário, em que eram registrados em fichas construídas manualmente. De acordo com Campos (2003) o controle de autoridade tem como objetivos a validação do trabalho do catalogador, percebido pela qualidade do catálogo, por meio da criação e manutenção de registros de autoridade e a garantia do acesso à informação consistente e confiável.

Esse processo de controle de autoridade encontrou solo fértil para seu desenvolvimento com o avanço do uso das tecnologias de informação nos ambientes de bibliotecas. Na visão de Campos (2003), o registro de autoridade tornou-se elemento obrigatório nos catálogos online, pois se conecta aos registros bibliográficos e é a partir das autoridades que se dá o processo de consulta à base de dados do catálogo, assim como, da disposição dos resultados pertinentes.

A preocupação com a normalização e a orientação à criação de registros de autoridade padronizados tem estado presente nas discussões da *International Federation of Library Associations and Institutions* (IFLA) com a constituição de grupos de trabalho voltados a estes temas, o que teve como resultados a publicação das *Guidelines for Authority and References Entries* (GARE) em 1984, atualizada em 2001 como *Guidelines for Authority Records and References* (GARR), e refletida na composição dos formatos de metadados de autoridade, como o *United International Machine Readable Cataloging* (UNIMARC) (Llanes Padrón, Bastos & Fujita, 2014).

Nesse contexto, hoje os pontos de acesso por assuntos em geral são construídos a partir da criação de um registro de autoridade de assunto por meio de um arquivo de autoridade em formato *Machine Readable Cataloging* - MARC21.

Dentre os avanços ocorridos no campo da construção de registro de autoridade de assunto deve-se dar destaque à publicação em 2010 do modelo conceitual *Functional Requirements for Subject Authority Data* (Requisitos Funcionais para Dados de Autoridade Assunto – FRSAD), que vem a fazer parte do que se tem denominado “Família FRBR”, já composta pelo próprio *Functional Requirements for Bibliographic Records* (Requisitos Funcionais para Registros Bibliográficos - FRBR) e cuja primeira expansão, no que tange às formas autorizadas de nome, havia tomado forma no modelo conceitual *Functional Requirements for Authority Data* (Requisitos Funcionais para Dados de Autoridade – FRAD).

Nesse cenário, o modelo FRBR define três grupos de entidades, sendo que o primeiro grupo trata daquilo que é descrito no registro bibliográfico, ou seja, os próprios documentos, e é abordado no FRBR; o segundo grupo aborda as entidades responsáveis pelo conteúdo do documento e é abordado no modelo FRAD; o terceiro grupo, composto pelas entidades *thema* e *nomen*, é abordado no modelo FRSAD, enfocando as tarefas de representação temática da informação (International Federation of Library Associations and Institutions, 2010).

O modelo conceitual FRSAD pode ser considerado uma tendência na representação dos assuntos de documentos em catálogos, tendo como objetivos “prover uma compreensão clara e compartilhada das informações que os dados de autoridade assunto devem fornecer para corresponder às necessidades dos usuários” e o uso e compartilhamento de dados de autoridade de assunto (Melo & Brascher, 2014, p. 109).

O REGISTRO DE AUTORIDADE DE ASSUNTO: COMPOSIÇÃO E APLICAÇÃO

Segundo MARC21 (2005), o registro de autoridade consiste no estabelecimento de formas padronizadas de nomes (pessoais, corporativos, eventos, jurisdição, título uniforme, combinações nome/título) e assuntos (termos tópicos, nomes geográficos, nomes com subdivisão de assunto, termos e subdivisão de assunto), sendo ambos usados como pontos de acesso, ao padronizarem as entradas principais e secundárias no registro bibliográfico com o termo nome para entrada principal e secundária e o termo assunto no catálogo online construído por meio de software de gestão e manutenção de catálogos, como por exemplo, o Aleph, Pergamum, Sophia e outros.

Os pontos de acesso por assunto “[...] são constituídos por termos (palavra ou grupo de palavras) que representam o documento. Por eles serem padronizados permitem que a representação da informação seja feita de maneira consistente” (Martinho, 2010, p. 57). Ou seja, por meio de um registro de autoridade de assunto pode-se estabelecer a forma padronizada dos pontos de acesso por assunto selecionados na indexação para representação no registro bibliográfico.

O registro de autoridade de assunto é composto por indicações de relações e notas de escopo, e tem quatro funções principais, fornecer o termo de assunto autorizado ou referir ao termo autorizado, indicar se um termo pode ou não ser subdividido geograficamente, fornecer uma nota de escopo, quando houver necessidade de explicar a cobertura a que o termo se aplica, e indicar relações entre termos (Fujita & Santos, 2016). Em seguida, apresenta-se exemplo de registro de autoridade de assunto de termo tópico da ferramenta *Classification Web* da *Library of Congress Subject Headings* (LCSH) em formato de ficha terminológica simples e registro MARC.

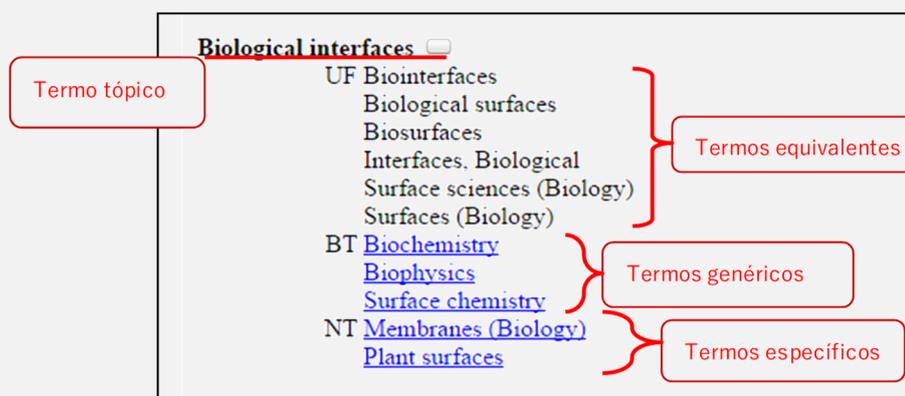


Figura 1 - Registro de autoridade em ficha terminológica simples

The screenshot shows a MARC record for 'Biological interfaces'. The record is displayed in a table-like format with fields for ID, Entered, Replaced, and various control numbers. The main body of the record contains MARC fields 010 through 675. Annotations in red boxes and arrows identify specific fields: 'Termo tópico' points to field 150, 'Termos equivalentes' points to fields 450, and 'Termos genéricos' points to fields 550.

Field	Value	Annotation
010	\$a sh 86004344	
035	\$a (DLC)sh 86004344	
035	\$a (DLC)148871	
040	\$a DLC \$c DLC \$d DLC	
150	\$a Biological interfaces	Termo tópico
450	\$a Biointerfaces	Termos equivalentes
450	\$a Biological surfaces	
450	\$a Biosurfaces	
450	\$a Interfaces, Biological	
450	\$a Surface sciences (Biology)	Termos genéricos
450	\$a Surfaces (Biology)	
550	\$w g \$a Biochemistry	
550	\$w g \$a Biophysics	
550	\$w g \$a Surface chemistry	
670	\$a Work cat. : Institute of Basic Medical Sciences. Interaction of cells with natural and foreign surfaces, 1986.	
670	\$a McGraw-Hill dict. life sci. \$b (Interface)	
670	\$a Biol. horizons in surface science, 1973.	
670	\$a Jones. Biological interfaces, 1975.	
670	\$a LC data base, 6/18/86 \$b (Biological interfaces. Biosurfaces. Biological surfaces)	
675	\$a Web. 3; \$a Henderson. Dict. sci. terms	

Record: 128997
Added: Thu Jul 17 00:00:00 1986
Modified: Fri Nov 13 17:36:37 1987

Figura 2 - Registro de autoridade em registro MARC

A função de um registro de autoridade em um sistema de informação é estabelecer a forma autorizada do assunto principal chamado de cabeçalho (ou de termo tópico ou assunto tópico, compõe o campo MARC 150 do registro de autoridades) de um documento, e também as formas não autorizadas, ou seja, que não são estabelecidas como assunto principal, mas que podem ser considerados como termos equivalentes (chamados de remissivas e são dispostos no campo MARC 450 do registro de autoridades) e fazer a conexão entre a busca do usuário e a recuperação da informação, para que o usuário não tenha que pensar em uma sucessão de possíveis termos relacionados ao documento buscado. O usuário não precisa buscar necessariamente o termo principal, pois o sistema apura tanto o principal quanto as remissivas; se um usuário pesquisa um assunto que é uma remissiva, o sistema vai rastrear o termo tópico no qual a remissiva está registrada, sendo interligado diretamente com o registro bibliográfico, pois o campo 150 do registro de autoridade é o equivalente ao campo MARC 650 (Assunto – termo tópico) do registro bibliográfico, que contém o termo de indexação.

Outro tipo de relação entre termos que é indicado no registro de autoridade de assunto é o de termo genérico/termo específico, em que é indicada a posição do termo na hierarquia de assuntos. Um termo genérico é indicado no campo MARC 550 do registro de autoridades, dispondo-se o código *g* no subcampo *w*. Como esta relação é recíproca, o sistema deve ser apto a incluir a relação inversa automaticamente. Este tipo de relação entre termos é útil ao catalogador, na decisão sobre qual termo de assunto indicar como ponto de acesso a um documento, bem como ao usuário, que poderá navegar pela hierarquia de termos para decidir qual assunto é mais relevante ao seu interesse.

Também se indica no registro de autoridade a relação entre termos relacionados, ou termos associados, que conecta dois termos que possuem alguma conexão diferente das relações anteriores (de equivalência e de hierarquia). Esta relação também é indicada no campo MARC 550 subcampo *a* do registro de autoridades.

DISCUSSÃO SOBRE O REGISTRO DE AUTORIDADE NA CONSTRUÇÃO E USO DE LINGUAGEM DE INDEXAÇÃO

O controle de vocabulário é um ponto de sensível importância para a prática da representação da informação no que se refere à compatibilização das ocorrências linguísticas que se encontram em um sistema de informação, manifestada pelos discursos de autores, indexadores e usuários. As linguagens de indexação sejam tesouros, taxonomias, ontologias – são considerados Sistemas de Organização do Conhecimento com o propósito definido de desempenhar funções de controle de vocabulário, embasando os processos de organização e representação do conhecimento (Melo & Brascher, 2014).

A construção e atualização de linguagens de indexação realizadas por meio de softwares específicos com possibilidade de integração aos softwares de gestão de catálogos online permitirão a intermediação com os registros de autoridade que se articularão aos registros bibliográficos para realizar um contínuo trabalho de controle, correção e atualização linguística. Para isso, a construção de registros de autoridade de assunto tem importância pela representação sintática e semântica de cada termo na linguagem de indexação e de sua interligação aos registros bibliográficos, pois como afirma Martins (2013) será por meio desta ferramenta que os indexadores irão alimentar e organizar continuamente todos os registros bibliográficos que compõem as bases de dados dos catálogos online.

Tratamos do uso concomitante de softwares específicos para a construção e atualização de linguagens de indexação com o software de gestão do catálogo tendo em vista que nem todos os softwares de gestão de acervos bibliográficos possuem as funcionalidades requeridas para atender às necessidades de gestão de tesouros, que demandam, por exemplo, o estabelecimento das relações padrão entre termos e conceito, bem como a criação de novos tipos de relações; a gestão do fluxo dos termos (desde sua condição como termo candidato); a importação e exportação completa ou parcialmente da linguagem; a visualização da linguagem em sua estrutura hierárquica, entre outras. Exemplo desse caso é o software Aleph utilizado nos maiores sistemas de bibliotecas universitárias do Brasil e na plataforma Periódicos CAPES.

A integração das linguagens de indexação com os sistemas de recuperação da informação por meio do estabelecimento dos registros de autoridade de assunto também colabora com a realização das pesquisas pelos usuários. Isto, em razão de a linguagem de indexação possuir seu vocabulário e relações sintáticas, semânticas e lógico-hierárquicas representadas por meio de registros de autoridade em um arquivo de autoridades, com visibilidade amigável ao usuário para que este conheça e consulte essa linguagem de indexação ao fazer suas pesquisas no catálogo online.

No entanto, a linguagem de indexação será acionada para realização de consultas, mesmo que o usuário não opte por consultá-la previamente, tendo em vista estar contida no próprio sistema de gerenciamento do catálogo, o que garante a mediação da pesquisa do usuário pela linguagem. De acordo com Martins (2013, p. 108) os metadados incluídos nos arquivos de autoridade “reproduzem todos os dados relativos aos cabeçalhos e subcabeçalhos, relacionando-os com os registros bibliográficos, o que do ponto de vista do acesso temático, permite dirigir o usuário dos seus próprios termos para os existentes no sistema, ou sugerir termos alternativos de pesquisa”.

Compreende-se que para instituições que já se utilizam dos registros bibliográficos em formatos de metadados padronizados, tais como o MARC21, a construção e manutenção da linguagem de indexação com o uso deste recurso se vê como adequada e coerente.

Para a qualidade do controle de vocabulário há que se mencionar a vantagem de que, depois de criados os registros de autoridade de assuntos, o sistema de gestão da base de dados irá proceder à interligação com o catálogo bibliográfico, importando e utilizando o termo autorizado. Havendo a necessidade de atualização ou correção do termo, a manipulação do registro de autoridade de assunto implicará na correção simultânea de quaisquer registros bibliográficos em que este termo tenha sido empregado (Martins, 2013). Função que ocorrerá caso a linguagem seja construída *a priori* e os registros bibliográficos indexados com esta linguagem. Caso a linguagem seja adotada *a posteriori* é necessária a reindexação do acervo para que esta possa garantir a qualidade da indexação. Tais operações dependem do trabalho contínuo de organização e representação da linguagem de indexação em registros de autoridade.

CONSIDERAÇÕES FINAIS

Considerando-se o exposto, compreendemos que a construção e uso de linguagens de indexação e sua interligação com as bases de dados bibliográficas e com os catálogos online muito se beneficia com o uso de formas padronizadas de apresentação de metadados e que, portanto, deve ser feita privilegiando a composição dos registros de autoridade de assunto.

Quanto aos desenvolvimentos dos formatos de manipulação dos dados de autoridade, no sentido dos modelos conceituais referidos previamente, compreende-se que os principais elementos das linguagens de indexação são os conceitos, os termos e as relações que se estabelecem entre eles, que também se encontram presentes no modelo FRSAD, evidenciando que sua proposta pode se adequar às diferentes linguagens de indexação.

REFERÊNCIAS BIBLIOGRÁFICAS

American National Standards Institute/ National Information Standards Organization Z39.19-2005. (2005) Guidelines for the construction, format, and management of monolingual controlled vocabularies. Bethesda, Maryland: NISO Press.

Campos, F. M. (2003). *Controle de autoridade: novos contextos e soluções*. Lisboa: Biblioteca Nacional. Obtido de: <http://purl.pt/331/1/docs/comunicacao/12manha/controlodeautoridade.doc>

Campos, M. L. de A. (2001). *Linguagem documentária: teorias que fundamentam sua elaboração*. Niterói: UFF.

Campos, M. L. de A. & GOMES, H. E. (2006). Metodologia de elaboração do tesouro conceitual: a categorização como princípio norteador. *Perspectivas em Ciência da Informação*, 11(3), 348-359.

Cervantes, B.M.N. & Fujita, M. S. L. (2012). Uma perspectiva diacrônica de parâmetros e modelos de construção de tesouros. // Cervantes, B. M. N. (Org.). *Horizontes da organização da informação e do conhecimento*. Londrina: EDUEL. v. 1, p. 1-18.

Currás, E. (2005). *Ontologias, taxonomia y tesouros: manual de construcción y uso*. 3.ed.atual.ampl. Gijón: Ediciones TREA. 337 p.

Dahlberg, I. (1978). Teoria do conceito. *Ciência da informação*, 7(2).

Dextre Clarke, S. G., Zeng, M. L. (2012). From ISO 2788 to ISO 25964: the Evolution of thesaurus standard towards interoperability and data modeling. // *Information Standards Quarterly*, 24(1), 20-26.

DIRECTRICES para registros de autoridad y referencias. (2001). 2 ed. rev. por el Grupo de Trabajo de la IFLA para la Revisión de las GARE. Obtido de: <<http://www.ifla.org/files/assets/cataloguing/garr/garr-es.pdf>>.

Dobedei, V. L. D. (2002). *Tesouro: linguagem de representação da memória documentária*. Niterói: Intertexto; Rio de Janeiro: Interciência. 120p.

Fujita, M. S. L. & Santos, L. B. P. dos. (2016). A estrutura lógico-hierárquica de linguagens de indexação utilizadas por bibliotecas universitárias. *Scire: representación y organización del conocimiento*, 22(2), 37-46.

International Organization for Standardization. (2011). ISO/FDIS 25964-1: Information and documentation – Thesauri and interoperability with other vocabularies – Part 1: Thesauri for information retrieval. Geneva: International Organization for Standardization. 2011.

International Organization for Standardization. (2013). ISO/FDIS 25964-2: Information and documentation – thesauri and interoperability with other vocabularies - part 2: Interoperability with other vocabularies. Geneva: International Organization for Standardization. 2013.

Llanes Padrón, D., Fujita, M. S. L., & Bastos, F. M. (2014). Os registros de autoridade em sistemas de informação: uma perspectiva biblioteconômica e arquivística. *Enancib*, 15.

MARC 21: formato condensado para dados de autoridade. (2005). Tradução e adaptação de Margarida M. Ferreira. Marília: FUNDEPE.

Martinho, N. O. (2010). A dimensão teórica e metodológica da catalogação de assunto. Dissertação (Mestrado em Ciência da Informação). Marília: Faculdade de Filosofia e Ciências, Universidade Estadual Paulista “Júlio de Mesquita Filho”, 2010.

Martins, M. C. B. (2013). *Indexação e controlo da terminologia em bibliotecas do ensino superior politécnico em Portugal: o sistema no Instituto Politécnico de Portalegre*. Salamanca: Universidad de Salamanca. Obtido de: <http://hdl.handle.net/10400.26/7171>

Melo, M. A. F., & Bräscher, M. (2014). Requisitos Funcionais para Dados de Autoridade Assunto (FRSAD): entidades, atributos e relacionamentos. *RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação*, 12(2), 102-119.

Moreira, M. P.; Moura, M. A. (2006). Construindo tesouros a partir de tesouros existentes: a experiência do TCI – Tesouro em Ciência da Informação. *DataGramZero: Revista em Ciência da Informação*, 7(4), ago.

Pereira, E. C.; Bufrem, L. S. (2005). Princípios de organização e representação de conceitos em linguagens documentárias. *Encontros Bibli: Revista eletrônica de Biblioteconomia e Ciência da Informação*, Florianópolis, 20.

REQUISITOS Funcionales para Datos de Autoridad de Materia (FRSAD). Grupo de Trabajo de IFLA sobre los Requisitos Funcionales para Registros de Autoridad de Materia (FRSAR). Obtido de: <http://www.ifla.org/files/assets/cataloguing/frsad/frsad-final-report-es.pdf>



RECORDS IN CONTEXTS Y LA PUBLICACIÓN DE CONJUNTOS DE DATOS ARCHIVÍSTICOS INTEROPERABLES

Juan-Antonio Pastor-Sánchez¹, Dunia Llanes Padrón²

¹Universidad de Murcia, 0000-0002-1677-1059, pastor@um.es

²Universidad de La Habana, 0000-0001-8639-4706, duniialp@yahoo.es

RESUMEN Introducción: Este trabajo analiza el modelo conceptual Records in Contexts (RiC-CM) propuesto por el Consejo Internacional de Archivos (ICA) para la descripción archivística, así como su implementación en una ontología OWL y su aplicación para la publicación de datos en el ámbito Linked Data. **Metodología:** Se ha estudiado la estructura y elementos de RiC-CM en relación con normas de descripción ya existentes para la elaboración de la ontología. También se han considerado, las condiciones necesarias para la elaboración y mantenimiento de conjuntos de datos interoperables que contemplen la aplicación de buenas prácticas para la publicación de datos estructurados en la web. **Resultados:** Se muestra un análisis del modelo conceptual RiC-CM junto con una propuesta de ontología OWL para su representación. También se detalla una lista de buenas prácticas para la publicación de conjuntos de datos archivísticos. **Conclusiones:** El trabajo revela el carácter integrador del modelo conceptual de Records in Contexts, así como las posibilidades que ofrece una mayor interoperabilidad de los datos de descripciones archivísticas.

PALABRAS CLAVE Descripción Archivística, Records in Contexts, web semántica, conjuntos de datos, Linked Data

ABSTRACT Introduction: This work analyzes the conceptual model Records in Contexts (RiC-CM) proposed by the International Council of Archives (ICA) for the archival description, as well as its implementation in an OWL ontology and its application for the publication of data in the field Linked Data. **Methodology:** The structure and elements of RiC-CM have been studied in relation to existing norms of description for the elaboration of the ontology. The necessary conditions have also been considered for the development and maintenance of interoperable data sets that include the application of good practices for the publication of structured data on the web. **Results:** An analysis of the RiC-CM conceptual model is presented together with an OWL ontology proposal for its representation. A list of good practices for the publication of archival data sets is also detailed. **Conclusions:** The work reveals the integrative nature of the conceptual model of Records in Contexts, as well as the possibilities offered by greater interoperability of archival descriptions data.

KEYWORDS Archival Description, Records in Contexts, semantic web, datasets, Linked Data

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUCCIÓN

El proceso de representación archivística ha cambiado en lo que respecta a sus conceptualizaciones y prácticas debido, principalmente, a las tecnologías basadas en internet y a las nuevas necesidades de acceso a la información por parte de los usuarios. Las nuevas formas de representación de la información abarcan a bibliotecas, archivos, museos y galerías de arte y es la interoperabilidad semántica quien permite relacionar datos estos diferentes ámbitos, ayudando a eliminar las líneas de separación que existen entre dichas instituciones. La elaboración de modelos conceptuales para la descripción archivística es un paso importante para lograr la interoperabilidad semántica de todos los recursos de información.

Después de la intensa actividad normativa desarrollada por el *International Council Archives* (ICA) y algunas comunidades nacionales en la década de los 90 del pasado siglo, comenzaron a surgir los primeros modelos conceptuales para la Archivística. Los nuevos preceptos no sólo incluyen las relaciones entre las sub-entidades del documento (descripción multinivel) sino que analizan los vínculos de estos con su contexto, analizando su producción, gestión y conservación. Estos modelos son conocidos como pluridimensionales pues incluyen diversas entidades, sus relaciones y definen propiedades.

Los primeros modelos nacionales para la descripción de objetos archivísticos fueron publicados en Australia, Nueva Zelanda, España y Finlandia¹ y tuvieron gran impacto tanto a nivel nacional como en el contexto internacional. El ICA, motivado por el auge de la normalización y por la aplicación de los estándares, en el 2012, formó un grupo de expertos para desarrollar un modelo conceptual internacional en materia de archivos. En 2016 se publica el primer borrador de “*Records in Contexts: A Conceptual Model for Archival Description*” (RiC-CM). Este modelo completa, amplía e interrelaciona las normas ISAD(G), ISAAR(CPF), ISDF e ISDIAH y se adapta a las nuevas posibilidades que brindan las tecnologías de la información y las comunicaciones para integrar datos y servicios en la web.

Los modelos conceptuales archivísticos, en general, representan e interrelacionan las entidades involucradas en el proceso descriptivo, facilitando la interoperabilidad entre registros de información de archivos, bibliotecas, museos o cualquier otra institución cultural. En este sentido, la interoperabilidad semántica permite que los sistemas de información intercambien datos directamente (Sheth, 1999, p. 4; Hernández-Pérez et al., 2009, p. 68) facilitando la reutilización de la información y aumentando la capacidad para generar y transformar datos, contenidos, conocimiento, productos y servicios, armonizando datos y procesos (Tolk y Muguira, 2003, p. 2).

¹ Australian Government Recordkeeping Metadata Standard. (Australia); Technical Specifications for the Electronic Recordkeeping Metadata Standard. (Nueva Zelanda); Modelo Conceptual de Descripción Archivística y Requisitos de Datos Básicos de las Descripciones de Documentos de Archivo, Agentes y Funciones. (España); Finnish Conceptual Model for Archival Description (Finlandia).

La web ofrece diversas tecnologías para acceder, publicar y compartir información siguiendo los principios de la interoperabilidad semántica. RDF es un modelo de datos que describe las propiedades y relaciones de los recursos de información y junto con RDF Schema y OWL permiten definir ontologías con un alto grado de expresividad semántica. El uso de RDF para publicar datos archivísticos permitiría una interoperabilidad mayor que la que ofrecen formatos como EAD o EAC. Los modelos conceptuales, como Records in Contexts, pueden implementarse con RDFS y OWL para lograr la interoperabilidad semántica en la web de los conjuntos de datos archivísticos. Además, las tecnologías de la web semántica y los servicios basados en Linked Data ofrecen una serie de ventajas evidentes para los Archivos (Baker et al., 2011; Sulé et al., 2016, p. 12; Hidalgo-Delgado et al., 2016, p. 2; McKenna, 2013, p. 352).

Este trabajo analiza el Modelo Conceptual Records in Contexts (RiC-CM) de Descripción Archivística elaborado por el International Council of Archives (ICA) al tiempo que se propone su adaptación en forma de ontología OWL y la publicación de conjuntos de datos archivísticos adecuados a las buenas prácticas de publicación de datos en la web. Para ello, se realiza un estudio de RiC-CM de los propósitos, entidades y relaciones de este. También se presentan los criterios de modelado de la ontología OWL propuesta y la aplicación de las buenas prácticas para la publicación de datos en la web recomendada por el W3C.

METODOLOGÍA

En un primer momento se procedió a la contextualización del trabajo en relación a la actividad previa realizada con respecto a la normativa de descripción. Igualmente se revisaron las tecnologías más relevantes de la web semántica y su vinculación con la interoperabilidad de conjuntos de datos. Se ha realizado un análisis del borrador del modelo conceptual propuesto por el ICA, abarcando las entidades, las propiedades y las relaciones de RiC-CM.

A partir de dicho análisis se ha procedido al modelado de la ontología OWL. En una primera versión se realizó una implementación directa del modelo conceptual. Una segunda versión de la ontología se centró en la optimización de propiedades y relaciones. Para ello se unificaron algunas propiedades de datos y de objeto de la primera versión y se definieron doblemente algunas propiedades de datos para dar soporte a la representación de datos a partir de la vinculación de recursos de términos de listas y vocabularios controlados.

Finalmente se procedió a un análisis de las buenas prácticas de datos en la web y se identificaron aquellas cuya aplicación resultan más pertinentes para la publicación de conjuntos de datos archivísticos.

RECORDS IN CONTEXTS: EL NUEVO MODELO CONCEPTUAL PARA LOS ARCHIVOS

En RiC-CM se establecen los propósitos esenciales de la descripción archivística en los contextos actuales de posmodernidad. El ICA propone tres roles fundamentales: Gestión Documental (*Management of Records*); Preservación de Documentos (*Preservation of Records*) y Reutilización de datos/información (*Ongoing Use and Reuse of Records*), (ICA, 2016, p. 14-15). La descripción se adapta a los nuevos tiempos, donde la representación de la información archivística a partir de modelos

de datos es determinante para la gestión, la preservación digital y la reutilización e interoperabilidad de datos en la web semántica. Representar las funciones, actividades y transacciones que dan origen a la documentación, también, facilita el uso continuo de los registros y conserva la memoria colectiva de las naciones, aspectos determinantes para la gestión. El ICA, al pronunciarse por una descripción encauzada a la gestión, sitúa a esta actividad en un lugar indispensable dentro de la gestión integrada de procesos en las organizaciones.

Según el RiC-CM, al describir entidades archivísticas, a través de metadatos, se posibilita el mantenimiento de la autenticidad e integridad de los documentos. En el modelo se expone que “documentar el contexto describiéndolo es esencial para la conservación de los registros” (ICA, 2016). Por último, el texto hace referencia a la importancia de la descripción para la reutilización de los datos archivísticos en entornos web. Además, la reutilización de información está asociada al acceso, la calidad y la usabilidad de datos públicos abiertos (*Open Data Governmet*) permiten la creación de nuevos datos, contenidos, conocimientos, productos o servicios.

El ICA, partiendo de sus cuatro estándares y de otros modelos de la Ciencias de la Información ha propuesto un conjunto de entidades que representan la realidad de la descripción archivística: documento, componente documental, conjunto documental, agente, ocupación, posición, función, función abstracta, actividad, norma, forma documental, fechas, lugar, concepto/asunto. El RiC-CM contempla las entidades definidas en ISAD(G) (Documento), en ISAAR(CPF) e ISDIAH (Agente) y en ISDF (funciones - actividades), unificando todos los componentes del conjunto normativo. Asimismo, se puede apreciar que RiC-CM tiene entidades coincidentes con los *Functional Requirements for Bibliographic Records* y en el *Conceptual Reference Model (CRM) del International Council of Museums (ICOM)* y en los modelos de Australia, Nueva Zelanda, España y Finlandia; las entidades documento, agente, función, norma, lugar, período y concepto son denominadores comunes en los modelos conceptuales desarrollados en las CI y constituyen entidades principales.

Este modelo presta gran importancia al contexto de producción de los documentos y así lo deja explícito tanto en sus propósitos como con las entidades función, función abstracta, ocupación, actividad, posición o mandato. Con ello se define el contexto de creación de los records, garantizando la veracidad, integridad y autenticidad de los documentos archivísticos. Además, define entidades como fecha, lugar, concepto o asunto esenciales para la creación de los puntos de acceso normalizados determinantes en el proceso de búsqueda y recuperación de información.

Las 69 propiedades definidas en el modelo son otro aspecto determinante en *Records in Contexts*, para la descripción de las 14 entidades archivísticas, reduciendo de esta forma la cantidad de atributos, con respecto al conjunto normativo. Por lo tanto se pretende presentar descripciones completas pero a su vez sucintas. El esquema de elementos de RiC-CM está basado en los trabajos normativos internacionales, nacionales y regionales desarrollados durante más de 20 años, además de las experiencias prácticas a través del uso de determinados software.

La ejemplificación de las relaciones entre las entidades es uno de los temas más trabajados y conseguidos en este nuevo estándar conceptual. La determinación de una amplia variedad de relaciones garantiza la interrelación de los contenidos de los documentos con los diversos ambientes en los que fueron producidos. En el RiC-CM se definen cerca de 800 relaciones (pueden aumentar en el texto definitivo) para definir las asociaciones entre entidades para incrementar la capacidad de búsqueda, recuperación y uso de la información. La figura siguiente muestra un ejemplo de descripción archivística conforme las relaciones y las entidades definidas en el RIC-CM.

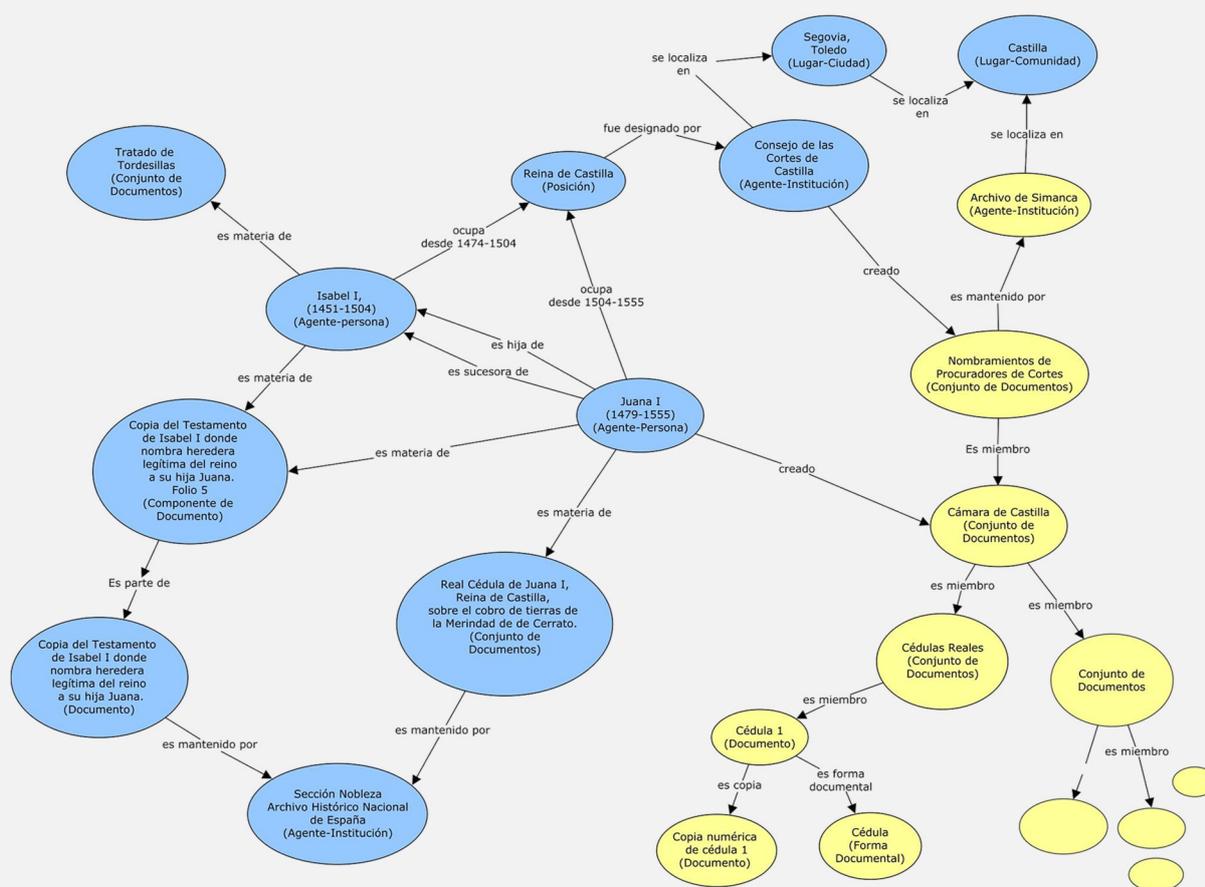


Figura 1. Ejemplo de descripción archivística a partir del RIC-CM.

Fuente: Llanes-Padrón & Moro-Cabero, 2017, p. 532

La descripción archivística, a partir del RIC-CM, facilita la consulta de recursos diversos entrelazados a partir de un enfoque único y como resultado permite descubrir y compartir más información en un modelado de red. El desarrollo de una ontología de este modelo, basada en RDF y OWL es el punto concluyente para posicionar a los recursos archivísticos en la dinámica de la publicación directa en la web, los datos abiertos entrelazados, la reutilización de datos y la interoperabilidad con sistemas de información de instituciones diversas.

INTEROPERABILIDAD, REUTILIZACIÓN Y TECNOLOGÍAS DE LA WEB SEMÁNTICA

En entornos web, las descripciones de recursos de información se realizan mediante modelos descriptivos y sus respectivos vocabularios cuya implementación suele realizarse a través de esquemas de metadatos y ontologías. Los metadatos son descripciones estructuradas y codificadas para describir características y propiedades de objetos y recursos (Gilliland-Swetland, 2003). Las ontologías proporcionan un modelo lógico-conceptual común en un área de conocimiento, definiendo, a diferentes niveles de formalización, el significado de las clases de objetos, las relaciones entre ellos y las propiedades que los caracterizan (Gruber, 1995).

Los metadatos y las ontologías requieren de tecnologías concretas para su aplicación práctica. Bajo la denominación de web semántica se agrupan un conjunto de especificaciones, tecnologías, desarrollos y aplicaciones para acceder, publicar y compartir información siguiendo los principios de la interoperabilidad semántica que permiten representar el significado de los datos. Se sigue haciendo uso del protocolo HTTP (*HyperText Transfer Protocol*) para establecer las conexiones entre clientes y servidores y las referencias IRI (*Internationalized Resource Identifier*) permiten identificar de forma inequívoca cualquier recurso. Cuando una IRI puede utilizarse para localizar y acceder a un recurso se dice que está desreferenciada como URL (*Uniform Resource Locator*).

RDF (*Resource Description Framework*) es un modelo de datos que aporta un marco semántico para describir recursos mediante sentencias que forman tripletas (*triplets*) compuestas por un sujeto, un predicado y un objeto. Tanto los recursos como las propiedades se identifica de manera unívoca mediante referencias IRI definidas dentro de un espacio de nombres XML propio (que pueden abreviarse mediante prefijos). Las descripciones se realizan mediante la asignación de valores a determinados atributos mediante propiedades de datos, o con el establecimiento de relaciones entre dos recursos utilizando propiedades de objeto (Cyganiak et al., 2014). Las sentencias RDF se almacenan utilizando algún formato legible por máquina como RDF/XML, Turtle, JSON-LD o RDFa entre otros (Gandon y Schreiber, 2014; Prud'hommeaux et al., 2014; Sporny et al., 2014; Herman et al., 2015). RDF se amplía con RDF Schema y con OWL (*Ontology Web Language*) para definir jerarquías de clases y propiedades, restricciones de cardinalidad, equivalencia entre clases, relaciones inversas, etc (Brickley, Guha y McBride, 2014; Hitzler et al., 2012).

PROPUESTA DE ONTOLOGÍA OWL BASADA EN EL MODELO CONCEPTUAL DE RECORDS IN CONTEXTS

La definición de ontologías OWL a partir de modelos conceptuales han sido propuestas ampliamente para FRBR (Ciccarese y Peroni, 2011; Shotton y Peroni, 2016), RDA (Hillmann et al., 2010) y CIDOC-CRM (Reinhardt, 2009). El EGAD está desarrollando una ontología OWL denominada RiC-O para implementar el modelo conceptual. La propuesta recogida en este trabajo puede contribuir a un debate que haga de RiC-O una ontología más eficiente y se encuentra disponible para su descarga y referencia en el espacio de nombres <http://purl.org/ica/ric>².

La ontología OWL modela las entidades del modelo conceptual mediante 14 clases principales sobre las que se definen las correspondientes propiedades. El borrador del RiC-CM especializa la entidad “RiC-E4 Agent” al definir las relaciones (ICA, 2016, p. 30-90). De esta forma se han definido tres subclases específicas para la entidad “RiC-E4 Agent” (agentes delegados, grupos y personas). Teniendo en cuenta estas consideraciones la Tabla 1 muestra una propuesta de taxonomía de clases para la ontología.

² La documentación de la ontología se encuentra disponible en <http://skos.um.es/TR/ric>

Tabla 1. Modelado de clases OWL a partir de las entidades de RiC-CM. Fuente: elaboración propia.

Entidad RiC	Denominación	Clase OWL
RiC-E1	Record	ric:RiC-E1-Record
RiC-E2	Record Component	ric:RiC-E2-Record Component
RiC-E3	Record Set	ric:RiC-E3-Record Set
RiC-E4	Agent	ric:RiC-E4-Agent
	Delegate Agent (subtipo de Agent)	ric:RiC-E4-Agent-DelegateAgent
	Person (subtipo de Agent)	ric:RiC-E4-Agent-Person
	Group (subtipo de Agent)	ric:RiC-E4-Agent-Group
RiC-E5	Occupation	ric:RiC-E5-Occupation
RiC-E6	Position	ric:RiC-E6-Position
RiC-E7	Function	ric:RiC-E7-Function
RiC-E8	Function (Abstract)	ric:RiC-E8-FunctionAbstract
RiC-E9	Activity	ric:RiC-E9-Activity
RiC-E10	Mandate	ric:RiC-E10-Mandate
RiC-E11	Documentary Form	ric:RiC-E11-DocumentaryForm
RiC-E12	Date	ric:RiC-E12-Date
RiC-E13	Place	ric:RiC-E13-Place
RiC-E14	Concept/Thing	ric:RiC-E14-ConceptThing

Fuente: Elaboración propia.

La definición de tipos más específicos para las subclases que representan personas, grupos y agentes-delegados se realiza mediante la propiedad “RiC-P32 Type” (ric:RiC-P32-type en la ontología propuesta).

La mayor parte de las propiedades de RiC-CM son específicas para ciertas entidades y algunas de ellas tienen una semántica similar e incluso idéntica. En estos casos, se han modelado como una única propiedad cuyo dominio es la unión de las clases de las entidades correspondientes. Este es el caso de las propiedades “Type”³ cuya semántica no es exactamente la misma puesto que su dominio es distinto. Sin embargo, estas propiedades tienen un objetivo común: la categorización de un recurso dentro de una clasificación o lista de tipos específica de cada entidad. En consecuencia, la ontología propuesta considera que la semántica de dichas propiedades es la misma. De esta forma se define una única propiedad ric:RiC-P-type cuyo dominio es la unión de las oportunas clases (ver Tabla 2).

3 Los números de estas propiedades en el modelo conceptual son: RiC-P23, RiC-P32, RiC-P42, RiC-P45, RiC-P48, RiC-P52, RiC-P55, RiC-P58, RiC-P61, RiC-P63 y RiC-P66.

Tabla 2. Definición del dominio de la propiedad ric:RiC-P-type

1	ric:RiC-P-type rdf:type owl:ObjectProperty ;
2	rdfs:domain [rdf:type owl:Class ;
3	owl:unionOf (ric:RiC-E10-Mandate
4	ric:RiC-E11-DocumentaryForm
5	ric:RiC-E12-Date
6	ric:RiC-E13-Place
7	ric:RiC-E14-ConceptThing
8	ric:RiC-E3-RecordSet
9	ric:RiC-E4-Agent
10	ric:RiC-E5-Occupation
11	ric:RiC-E6-Position
12	ric:RiC-E7-Function
13	ric:RiC-E9-Activity)
14] .

Fuente: *Elaboración propia.*

El rango de las propiedades de datos se representan mediante literales RDF (rdfs:Literal). Sin embargo, RiC-CM establece el rango de algunas propiedades a partir de los términos de un vocabulario controlado. Por ello, y considerando que OWL no considera mutuamente excluyentes a owl:DatatypeProperty y owl:ObjectProperty, se han definido este tipo de propiedades como miembros de ambas clases representando las 69 propiedades de RiC-CM mediante 38 propiedades de datos y 15 propiedades de objeto.

También se reduce el número de relaciones de RiC-CM puesto que muchas de ellas tienen una semántica prácticamente idéntica, pudiendo definirse en OWL como una única propiedad de objeto. Esto puede verse en “is member of” que se utiliza con las entidades que representan documentos, conjuntos de documentos y agentes para la que el modelo conceptual define tres relaciones (Tabla 3).

Tabla 3. Relaciones “is member of” de RiC-CM

Número de relación	Nombre	Dominio	Rango
RiC-R18	is member of	RiC-E1 Record	RiC-E3 Record Set
RiC-R143	is member of	RiC-E3 Record Set	RiC-E3 Record Set
RiC-R266	is member of	RiC-E4 Agent	RiC-E4 Agent (of type group)

Fuente: *elaboración propia a partir de ICA, 2016.*

Esta relación, únicamente permite vincular documentos con conjuntos de documentos, conjuntos de documentos con otros conjuntos de documentos y agentes con otros agentes del tipo grupo. En este caso, la ontología propuesta define una única propiedad de objeto ric:RiC-R-isMemberOf cuyo dominio sería la unión de las clases ric:RiC-E1-Record, ric:RiC-E3-RecordSet y ric:RiC-E4-Agent.

Tabla 4. Modelado de las relaciones “is member of”.

1	ric:RiC-R-isMemberOf rdf:type owl:ObjectProperty ;
2	rdfs:domain [
3	a owl:Class ;
4	owl:unionOf (
5	ric:RiC-E1-Record
6	ric:RiC-E3-RecordSet
7	ric:RiC-E4-Agent
8)
9] ;
10	rdfs:range [
11	a owl:Class ;
12	owl:unionOf (
13	ric:RiC-E3-RecordSet
14	ric:RiC-E4-Agent-Group
15)
16] ;
17	rdfs:comment "Relations RiC-R18, RiC-R143, RiC-R266"@en ;
18	rdfs:label "is member of"@en .

Fuente: Elaboración propia.

No obstante, puesto que el rango de las relaciones puede variar en función de las clases sobre las que se definen la unificación de este tipo de relaciones con OWL requiere el establecimiento de restricciones para limitar el uso de una propiedad a los objetos de las clases correspondientes (Tabla 5). En consecuencia, las 792 relaciones del borrador del modelo conceptual se definen con 166 propiedades de objeto en la ontología propuesta.

Tabla 5. Definición de restricción de la propiedad ric:RiC-R-is Member Of sobre la clase RiC-E1-Record.

1	ric:RiC-E1-Record rdfs:subClassOf [rdf:type owl:Restriction ;
2	owl:onProperty ric:RiC-R-isMemberOf ;
3	owl:allValuesFrom ric:RiC-E3-RecordSet
4]

Fuente: Elaboración propia.

Por su parte, las propiedades “RiC-P68 Date” y “RiC-P69 Place” permiten describir las relaciones entre entidades. Por ello se han modelado como propiedades de anotación para su uso mediante reificaciones RDF.

PUBLICACIÓN DE CONJUNTOS DE DATOS ARCHIVÍSTICOS

Además de una ontología basada en RiC-CM también se precisa definir una serie de características que deben reunir los conjuntos de datos archivísticos. Un conjunto de datos debe ser una unidad autodescriptiva con datos estructurados disponibles en varios formatos para su descarga completa o acceso selectivo. Los metadatos que describen diversos aspectos de un conjunto de datos deben proporcionar información sobre su contenido y estructura, así como sobre sus diferentes versiones y distribuciones, para facilitar la confianza en los datos y su reutilización.

Los principios Linked Open Data (LOD) tienen como objetivo la interconexión de conjuntos de datos, utilizando IRIs desreferenciables (convertibles a URLs) durante la identificación y acceso a la

descripción de recursos mediante el protocolo HTTP (Berners-Lee, 2006; Haslhofer y Schandl, 2010). Sin embargo los escenarios de acceso y consumo de datos RDF son muy variables: en algunos casos se precise descargarlo en su totalidad y en otros utilizar SPARQL para un acceso selectivo (Harris, Seaborne y Prud'hommeaux, 2013) o utilizar una API especializada.

Esta diversidad de casos de uso exige un análisis detenido del proceso de publicación de un conjunto de datos. El W3C ha creado un grupo para trabajar en la identificación de diversos escenarios y retos a partir de los cuales se han identificado y descrito 35 buenas prácticas para la publicación de datos en la web (Farias, Burle y Calegari, 2016) que pueden resumirse en los siguientes puntos:

- Uso de IRIs persistentes para identificar tanto a las distintas versiones de un dataset como a los datos que contiene.
- Reutilización de vocabularios de datos para incrementar la interoperabilidad semántica y escogiendo niveles adecuados de modelado y formalización.
- Uso preferente de vocabularios y formatos estándar para que los datos sean legibles tanto por personas como por máquina, haciendo más flexible su reutilización.
- Enriquecimiento de los datos para aportar valor añadido a través del uso de otros conjuntos, bases de datos u otros recursos de información.
- Proporcionar acceso en tiempo real a datos actualizados, no solamente descargando todo el conjunto de datos, sino también mediante servicios web y APIs bien documentadas.
- Garantizar la privacidad y la seguridad de los datos sensibles, identificando qué partes de los conjuntos de datos no están disponibles públicamente e indicando el motivo.
- Informar a los usuarios y consumidores de los datos acerca del modo para contactar con los editores y publicar los resultados de dicha retroalimentación como datos abiertos.
- Incluir una licencia de uso del conjuntos de datos para que los usuarios estén informados sobre las limitaciones y posibilidades de reutilización de los datos.
- Incluir información sobre la procedencia del contenido del conjuntos de datos para que pueda evaluarse con la finalidad de medir su fiabilidad y confianza.
- Gestionar adecuadamente las distintas versiones de un conjunto de datos, evitando modificar APIs o servicios de acceso a los datos que obliguen a introducir cambios en los clientes que los utilicen.

La aplicación de estas buenas prácticas aporta una serie de beneficios en la publicación de datos en la web y una mejor comprensión de su estructura, naturaleza, significado, y mecanismos de acceso (Lee, Farias y Archer, 2015). Todo esto, deberá ser tenido en cuenta a nivel técnico, para la implementación o adopción de plataformas de publicación y consulta de conjuntos de datos archivísticos.

CONCLUSIONES

RiC-CM expone la importancia de la descripción para la reutilización de los datos archivísticos en la web y sienta las bases para que los archivos se beneficien de las ventajas de las tecnologías de la web semántica en lo concerniente a interoperabilidad, reutilización de datos y servicios basados en Linked Open Data. Además, la mayor parte de las entidades se encuentran en otros modelos conceptuales en el ámbito de las Ciencias de la Información lo que acentúa la idea de crear un modelo semántico integral

que facilite la interoperabilidad de los contenidos de todas las instituciones de información en el contexto web apoyados en desarrollos ontológicos.

El desarrollo de una ontología de este modelo, basada en el Esquema de Datos para la Web Semántica (RDF) es el punto concluyente para posicionar a los recursos archivísticos en la dinámica de la publicación directa en la web, los datos abiertos entrelazados, la reutilización de datos y la interoperabilidad con sistemas de información de instituciones diversas. Los conjuntos de datos archivísticos interoperables y de confianza son una aportación muy valiosa en el universo Linked Open Data y a su vez pueden enriquecerse enormemente mediante enlaces con vocabularios controlados SKOS, entidades de DBpedia, registros bibliográficos, etc. También será preciso alinear cualquier ontología basada en RiC-CM con otros vocabularios de metadatos y ontologías ya existentes como FOAF, SKOS o Dublin Core.

Cuando el EGAD termine sus trabajos y el ICA publique las versiones definitivas del modelo conceptual y de la ontología será el momento en el que las herramientas de gestión y descripción archivística deberán adaptarse a esta propuesta. El futuro dirá la repercusión de Records in Contexts en la publicación de datos archivísticos en la web. La publicación final de RiC trae consigo un nuevo desafío para la Ciencia de la Información que transita rumbo a la creación de un único modelo conceptual de descripción de recursos de información. Cada vez se avanza más en la ruptura de las líneas teóricas y prácticas que separan los procedimientos de trabajo en archivos, bibliotecas, museos y otras instituciones culturales. Sin lugar a dudas, los modelos conceptuales, las ontologías y los formatos son el camino para lograr la compatibilidad e interoperabilidad total de la información patrimonial y cultural en el contexto de la web semántica.

REFERENCIAS BIBLIOGRÁFICAS

Baker, T., Bermès, E., Coyle, K., Dunsire, G., Isaac, A., Murray, P., ... Zeng, M. (2011). Library Linked Data Incubator Group Final Report. Recuperado de <https://www.w3.org/2005/Incubator/ld/XGR-ld-20111025/>

Berners-Lee, T. (2006). Linked Data. Recuperado de <http://www.w3.org/DesignIssues/LinkedData.html>

Brickley, D., & Guha, R. V. (2014). RDF Schema 1.1. Recuperado de <https://www.w3.org/TR/2014/REC-rdf-schema-20140225/>

Brickley, D., Guha, R.V., & McBride, B. (2014). RDF Schema 1.1. Recuperado de <https://www.w3.org/TR/2014/REC-rdf-schema-20140225/>

Ciccarese, P., & Peroni, S. (2011). Essential FRBR in OWL2 DL. Recuperado de <http://purl.org/spar/frbr>

CNEDA. (2017). NEDA-MC. Modelo conceptual de descripción archivística: Entidades, relaciones y atributos. Recuperado de <https://www.mecd.gob.es/dms/mecd/cultura-mecd/areas-cultura/archivos/mc/cneda/documentacion/NEDA-MC-2017/NEDA-MC-2017.pdf>

Cyganiak, R., Wood, D., & Lanthaler, M. (2014). RDF 1.1 Concepts and Abstract Syntax.

Farias Lóscio, B., Burle, C., & Calegari, N. (2016). Data on the Web Best Practices. Recuperado de <https://www.w3.org/TR/2017/REC-dwbp-20170131/>

Gandon, F., & Schreiber, G. (2014). RDF 1.1 XML Syntax. Recuperado de <http://www.w3.org/TR/2014/REC-rdf-syntax-grammar-20140225/>

Gilliland-Swetland, A. J. (2003). Metadata - Where Are We Going?. En G.E. Gorman (ed.) *International yearbook of library and information management 2003-2004: Metadata applications and management* (pp. 17-33). London: Facet Publishing.

Gruber, T. R. (1995). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal of Human-Computer Studies* 43(5-6), 907-928. Recuperado de <http://doi.org/10.1006/ijhc.1995.1081>

Gueguen, G., Fonseca, V. M. M. da, Pitti, D. V., & Sibille-de Grimouard, C. (2013). Hacia un Modelo Conceptual Internacional de Descripción Archivística: Un informe preliminar del Grupo de Expertos en Descripción Archivística del Consejo Internacional de Archivos. *Métodos de información*, 4(7), 155-172. <https://doi.org/10.5557/IIMEI4-N7-155172>

Haslhofer, B., & Schandl, B. (2010). Interweaving OAI-PMH data sources with the linked data cloud. *International Journal of Metadata, Semantics and Ontologies archive*, 5(1), 17-31. Recuperado de <http://dx.doi.org/10.1504/IJMSO.2010.032648>

Harris, S., Seaborne, A., & Prud'hommeaux, E. (2013). SPARQL 1.1 Query Language. Recuperado de <http://www.w3.org/TR/2013/REC-sparql11-query-20130321/>

Herman, I., Adida, B., Sporny, M., & Birbeck, M. (2015). RDFa 1.1 Primer: Rich Structured Data Markup for Web Documents (Third Edition). Recuperado de <http://www.w3.org/TR/2015/NOTE-rdfa-primer-20150317/>

Hernández-Pérez, T., Rodríguez-Mateos, D., Martín-Galán, B., & García-Moreno, M. A. (2009). El uso de metadatos en la administración electrónica española: los retos de la interoperabilidad. *Revista española de Documentación Científica*, 32(4), 67-91.

Hidalgo-Delgado, Y., Senso, J. A., Leiva-Mederos, A., & Hípola, P. (2016). Gestión de fondos de archivos con datos enlazados y consultas federadas. *Revista española de Documentación Científica*, 39(3), 145.

Hillmann, D., Coyle, K., Phipps, J., & Dunsire, G. (2010). RDA Vocabularies: Process, Outcome, Use. *D-Lib Magazine*, 16(1/2). <https://doi.org/10.1045/january2010-hillmann>

Hitzler, P., Krötzsh, M., Parsia, B., Patel-Schneider, P. F., & Rudolph, S. (2012). OWL 2 Web Ontology Language Primer (Second Edition). Recuperado de <https://www.w3.org/TR/2012/REC-owl2-primer-20121211/>

ICA. (2000). ISAD(G): General International Standard Archival Description. Ottawa: International Council on Archives. Recuperado de http://www.ica.org/sites/default/files/CBPS_2000_Guidelines_ISAD%28G%29_Second-edition_EN.pdf

- ICA. (2016). Records in Contexts: A conceptual model for archival description (Consultation Draft v0.1). Recuperado de <http://www.ica.org/sites/default/files/RiC-CM-0.1.pdf>
- Lee, D., Farias Lóscio, B., & Archer, P. (2015). Data on the Web Best Practices Use Cases & Requirements. Recuperado de <https://www.w3.org/TR/2015/NOTE-dwbp-ucr-20150224/>
- Llanes-Padrón, D., Moro-Cabero, M. (2017). Records in Contexts: un nuevo modelo para la representación de la información archivística en el entorno de la web semántica. *El profesional de la información*, 26(3), 525-533.
- McKenna, G. (2013). Linked Heritage Experience in Linking Heritage Information. *JLIS.it*, 4(1), 325. <https://doi.org/10.4403/jlis.it-6304>
- Prud'hommeaux, E., Carothers, G., Beckett, D., & Berners-Lee, T. (2014). RDF 1.1 Turtle: Terse RDF Triple Language. Recuperado de <http://www.w3.org/TR/2014/REC-turtle-20140225/>
- Reinhardt, S. (2009). CIDOC CRM in OWL 2. Recuperado de <http://bloody-byte.net/rdf/cidoc-crm/>
- Schneider, M. (2012). OWL 2 Web Ontology Language RDF-Based Semantics (Second Edition). Recuperado de <https://www.w3.org/TR/2012/REC-owl2-rdf-based-semantics-20121211/>
- Sheth, A. P. (1999). Changing focus on interoperability in information systems: from system, syntax, structure to semantics. En M. Goodchild, M. J. Egenhofer, R. Fegeas, & C. Kottman (Eds.), *Interoperating geographic information systems* (pp. 5-29). Norwell, MA, USA: Kluwer Academic Publishers. Recuperado de <http://knoesis.org/sites/default/files/S98-changing.pdf>
- Shotton, D., & Peroni, S. (2016), FaBiO, the FRBR-aligned Bibliographic Ontology. Recuperado de <http://purl.org/spar/fabio/>
- Sporny, M., Kellogg, G., Lanthaler, M., & Lindström, N. (2014). JSON-LD 1.0: A JSON-based Serialization for Linked Data. Recuperado de <http://www.w3.org/TR/2014/REC-json-ld-20140116/>
- Sulé, A., Centelles, M., Franganillo, J., & Gascón, J. (2016). Aplicación del modelo de datos RDF en las colecciones digitales de bibliotecas, archivos y museos de España. *Revista española de Documentación Científica*, 39(1), 121. <https://doi.org/10.3989/redc.2016.1.1268>
- Tolk, A., & Muguira, J. A. (2003). The Levels of Conceptual Interoperability Model. En *Fall Simulation Interoperability Workshop*, Orlando, Florida, September 2003. Recuperado de <https://pdfs.semanticscholar.org/f655/af160f630b9be8dbab986f6a96953aa3e986.pdf>



UTILIZACIÓN DE CATEGORÍAS DE WIKIPEDIA EN PROCESOS DE ORGANIZACIÓN DE INFORMACIÓN: HACIA UNA REVISIÓN CUALITATIVA.

Jesús Tramullas¹, Ana I. Sánchez-Casabón², Piedad Garrido-Picazo³

¹Depto. Ciencias de la Documentación, Univ. de Zaragoza, 0000-0002-5374-9993, tramullas@unizar.es

²Depto. Ciencias de la Documentación, Univ. de Zaragoza, 0000-0002-0908-1615, asanchez@unizar.es

³Depto. Informática e Ingeniería de Sistemas, Univ. de Zaragoza, 0000-0002-1750-7225, piedad@unizar.es

RESUMEN Este trabajo revisa el estudio y la utilización del sistema de categorías de Wikipedia en la investigación científica, adoptando para ello una metodología de revisión sistemática de literatura, pero atendiendo a la revisión cualitativa del contenido de un conjunto de trabajos escogidos. Se identifican varios tipos de trabajos, atendiendo al estudio intrínseco del sistema de categorías, o a su uso como herramienta para el análisis de otros corpus documentales diferentes a Wikipedia. Se concluye que el sistema de categorías ofrece un esquema de clasificación válidos para abordar estudios sobre la organización del conocimiento en múltiples contextos.

PALABRAS CLAVE *categorías, Wikipedia, organización del conocimiento, ontologías, clasificación.*

ABSTRACT This paper reviews the study and use of the Wikipedia category system in the scientific research, adopting a systematic literature review approach, but considering the qualitative review of the content of a set of selected papers. Several types of work are identified, depending on the intrinsic study of the category system, or its use as a tool for the analysis of other documentary corpus other than Wikipedia. We conclude that the system of categories offers a valid classification scheme for different approaches and studies on the organization of knowledge in multiple contexts.

KEYWORDS *categories, Wikipedia, knowledge organization, ontologies, classification*

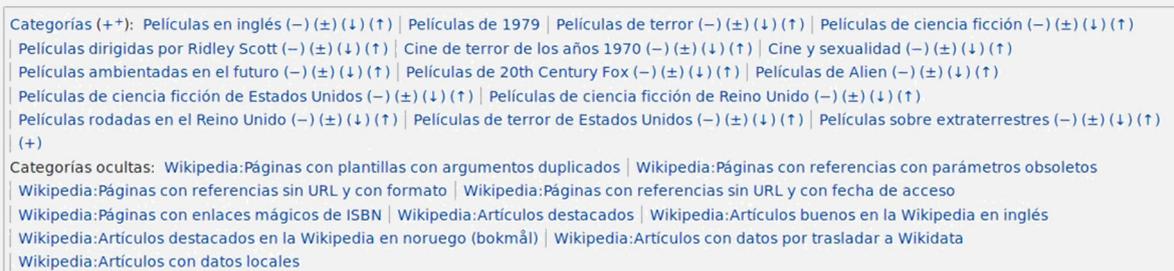
COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUCCIÓN

Wikipedia es el principal recurso de información enciclopédica disponible a escala mundial. Ofrece más de 32 millones de artículos, y es consultada diariamente por millones de usuarios. Puede considerarse en la actualidad como la mayor base de conocimiento, que es organizado y etiquetado conforme a unas instrucciones y parámetros básicos. Si se atiende a datos estructurados según criterios semánticos, entonces es el proyecto Wikidata (Vrandecic, 2013) la mayor base de conocimiento existente. La revisión de la estructura de los artículos en Wikipedia, así como las herramientas de organización y exploración de la enciclopedia, permiten identificar que uno de sus elementos fundamentales, inherentes al contenido, son las categorías (figura 1). El conjunto de las categorías, que

está compuesto por un vasto conjunto de términos, y que se encuentra en constante evolución, es utilizado por los editores para enmarcar los contenidos dentro de una estructura de organización del conocimiento (Capocci, Rao & Caldarelli, 2008). Las categorías fueron introducidas en Wikipedia en 2003 y las páginas de categorías y subcategorías en 2004. Ambos elementos son definidos, mantenidos y actualizados por la comunidad de editores de manera colaborativa (Thornton y MacDonald, 2012). Se trata de un sistema que combina una organización jerárquica con relaciones entre diferentes categorías, de forma que se crean polijerarquías y asociaciones. A todos los efectos, el sistema de categorías es un sistema de organización del conocimiento, y permiten agrupar los artículos en conjuntos y subconjuntos conceptuales o temáticos.

El estudio científico de la dinámicas de Wikipedia ha dado como resultado la publicación de trabajos sobre los procesos de edición colaborativos, las pautas de comportamiento de las comunidades de usuarios, el fenómeno del vandalismo, etc. (Mesgari et alii, 2015). El desarrollo de los estudios sobre el universo de Wikipedia ha permitido elaborar y disponer de varias revisiones sistemáticas de literatura, que atienden a diferentes enfoques (Tramullas, 2015). Dentro del conjunto de trabajos sobre Wikipedia, que muestran diferentes orientaciones, objetivos, métodos y resultados, también es posible identificar en la bibliografía especializada un buen número de estudios que están usando el corpus textual de Wikipedia, y en particular las categorías, como componente para diferentes estudios relacionados con la organización y la recuperación de información, el etiquetado social, la clasificación de documentos, o el etiquetado semántico y las ontologías.



Categorías (+): Películas en inglés (–) (±) (↓) (↑) | Películas de 1979 | Películas de terror (–) (±) (↓) (↑) | Películas de ciencia ficción (–) (±) (↓) (↑)
Películas dirigidas por Ridley Scott (–) (±) (↓) (↑) | Cine de terror de los años 1970 (–) (±) (↓) (↑) | Cine y sexualidad (–) (±) (↓) (↑)
Películas ambientadas en el futuro (–) (±) (↓) (↑) | Películas de 20th Century Fox (–) (±) (↓) (↑) | Películas de Alien (–) (±) (↓) (↑)
Películas de ciencia ficción de Estados Unidos (–) (±) (↓) (↑) | Películas de ciencia ficción de Reino Unido (–) (±) (↓) (↑)
Películas rodadas en el Reino Unido (–) (±) (↓) (↑) | Películas de terror de Estados Unidos (–) (±) (↓) (↑) | Películas sobre extraterrestres (–) (±) (↓) (↑)
(+)
Categorías ocultas: Wikipedia:Plantillas con argumentos duplicados | Wikipedia:Plantillas con referencias con parámetros obsoletos
Wikipedia:Plantillas con referencias sin URL y con formato | Wikipedia:Plantillas con referencias sin URL y con fecha de acceso
Wikipedia:Plantillas con enlaces mágicos de ISBN | Wikipedia:Artículos destacados | Wikipedia:Artículos buenos en la Wikipedia en inglés
Wikipedia:Artículos destacados en la Wikipedia en noruego (bokmål) | Wikipedia:Artículos con datos por trasladar a Wikidata
Wikipedia:Artículos con datos locales

Figura 1. Categorías en el artículo de Wikipedia “Alien: el octavo pasajero” (modo Edición).

En consonancia con todos estos aspectos, este trabajo tiene como objetivo principal identificar los usos y las aplicaciones que los investigadores están haciendo del sistema de categorías de Wikipedia, y cómo esta utilización se está reflejando en la evolución del corpus de literatura científica disponible. En segundo lugar, pretende revisar la forma en que un sistema de organización del conocimiento, elaborado colaborativamente, se está usando como recurso de investigación en diferentes aproximaciones al tratamiento y organización de la información, lo que refrendaría su validez como sistema de clasificación. Finalmente, debe indicarse que este trabajo no entra a valorar la estructura, la evolución ni la calidad del sistema de categorías de Wikipedia.

METODOLOGÍA

La metodología utilizada para la realización del trabajo ha sido la revisión sistemática de bibliografía. En este caso se ha optado por un estudio cualitativo, seleccionando trabajos específicos a revisar, antes que por un estudio cuantitativo descriptivo o inferencial de tipo bibliométrico. La recopilación de los datos bibliográficos se ha llevado a cabo a través de la consulta de las referencias sobre el tema objeto

de estudio disponibles en *Web of Science* y *Scopus*. El método de trabajo ha sido adoptado del propuesto por Okoli y Schabram (2009) para el estudio de los trabajos de investigación sobre Wikipedia, y que ya ha sido aplicado previamente por los autores de este trabajo (Tramullas, Garrido-Picazo y Sánchez-Casabón, 2016).

Durante la primera fase se seleccionaron de fuentes y la consulta a realizar. La consultas sobre *Scopus* y *Web of Science* se llevaron a cabo entre febrero y marzo de 2017, usando la expresión “Wikipedia AND categories”, y limitando la búsqueda a trabajos publicados entre 2002 y 2016. Se obtuvieron 666 resultados en Scopus y 311 en Web of Science. En ambos casos, los primeros trabajos identificados se publicaron en 2006. No se consultaron, por los objetivos y límites establecidos para este trabajo, ni las bibliotecas digitales de ACM ni IEEE, por la repetición de contenidos, ni *Google Scholar*, por la imposibilidad de limitar las búsquedas a posiciones específicas de los documentos o de sus referencias bibliográficas.

En una segunda fase, una vez obtenidos los datos en bruto de las referencias, se ha procedido a su procesamiento. En primer lugar se han fusionado ambos conjuntos, para proceder a la identificación y eliminación de los documentos duplicados. Esta tarea ha reducido el número de trabajos a 669. Posteriormente, se ha procedido a realizar una selección cualitativa de los trabajos, atendiendo a la identificación temática del contenido de los diferentes estudios, mediante la revisión de los títulos, resúmenes y palabras clave asignados a los mismos. El criterio de selección establecido fue el uso o estudio de categorías como elemento importante del trabajo que se revisase en cada caso. Cada trabajo fue revisado por los tres autores de manera independiente. En caso de discrepancia se revisaba el trabajo para decidir su inclusión o rechazo por mayoría. Este tipo de selección no permite valorar la calidad, pero este criterio no ha sido tenido en cuenta, ya que este trabajo no tiene como objetivos ni elaborar una bibliografía seleccionada ni evaluar la calidad de los trabajos o su impacto. El filtrado realizado ha permitido eliminar del conjunto aquellos artículos o comunicaciones cuyo uso o referencia a las categorías de Wikipedia no estaba directamente relacionado con los objetivos planteados en este trabajo. Finalmente, el número de trabajos seleccionados ha ascendido a 518.

Los estudios y trabajos seleccionados han sido revisados para identificar en los mismos la utilización que se ha hecho del sistema de categorías de Wikipedia. Se han identificado el contexto de aplicación, el uso de las categorías, el método utilizado y los resultados obtenidos, con la finalidad de delinear las líneas de investigación que han usado las categorías de Wikipedia como parte integrante de las mismas. Finalmente, se ha procedido a la descripción de los resultados obtenidos y a la elaboración de la síntesis y las propuestas de desarrollo.

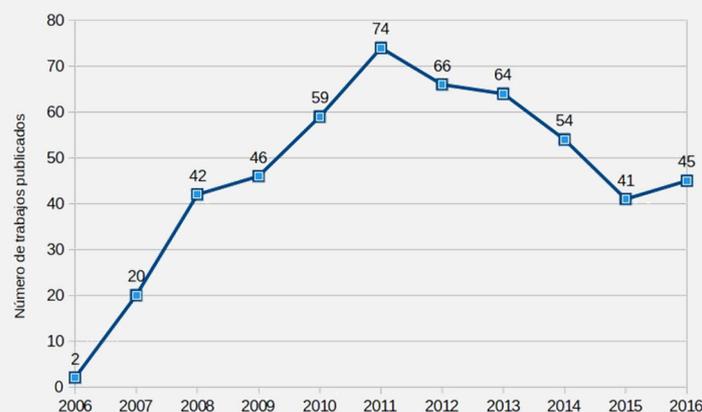


Figura 2. Evolución del número de publicaciones por año (*elaboración propia*).

Todos los datos obtenidos durante el proceso han sido liberados como *Open Data*, y las referencias bibliográficas obtenidas y revisadas se han publicado en grupos específicos de acceso público en *Mendeley*, *Zotero* y *CiteUlike* (tabla 1).

Tabla 1. Datos en acceso abierto

Mendeley	https://www.mendeley.com/community/research-on-wikipedia-categories/documents/
Zotero	https://www.zotero.org/groups/1543457/research_on_wikipedia_categories
CiteUlike	http://www.citeulike.org/groupdefault/20700

RESULTADOS

Los resultados obtenidos de la revisión cualitativa que se ha llevado a cabo demuestran la gran variedad de enfoque, usos y aplicaciones que los investigadores hacen del sistema de categorías de Wikipedia. Esta riqueza, a su vez, implica un límite para el análisis cualitativo que se propone: la combinación de técnicas, enfoques y aplicaciones existente en los trabajos de investigación imposibilita poder establecer divisiones precisas entre tipos de trabajos. Si bien pueden encontrar trabajos que puede adscribirse claramente a un tema (por ejemplo, generación de ontologías), se encuentran acto seguido trabajos que generan ontologías y las combinan con otras técnicas para poder aplicarlas a procesos de recuperación de información, que, a su vez, puedan darse en un contexto genérico, o estar circunscritos a dominios específicos. En consecuencia, un tratamiento cualitativo como el que aborda este trabajo debe limitarse a delinear temas y líneas de investigación claramente identificados. Para poder investigar las relaciones entre los temas y las posibles divisiones o tipos entre ellos sería necesario aplicar técnicas de procesamiento de información basadas en análisis numéricos del corpus bibliográfico, como recientemente han desarrollado Smiraglia y Cai (2017).

Atendiendo a la limitación indicada, la revisión cualitativa permite, en primer lugar, establecer una primera división. En primer lugar, deben señalarse aquellos estudios que analizan el sistema de categorías dentro del propio contexto de Wikipedia (cubriendo aspectos como la organización y el ciclo de vida de los contenidos, su estructura, la utilización por la comunidad de usuarios, o la propia evolución y mejora del sistema de categorías...). En segundo, aquellos trabajos que utilizan las categorías de Wikipedia en el marco de estudios sobre diferentes aspectos del procesamiento y tratamiento de la información, generalmente sobre corpus documentales ajenos a Wikipedia (formados por colecciones de documentos de diferentes tipos y/o temáticas, páginas web, mensajes en redes sociales...). Alrededor del 90% de los estudios revisados corresponden a este segundo grupo. Hay que destacar la presencia de trabajos que usan corpus creados ex-profeso que, a su vez han sido generados o extraídos de la propia Wikipedia. Todos los trabajos revisados podrían encuadrarse en alguna de las cuatro categorías de investigación sobre Wikipedia enumeradas por Nielsen (2017), aunque este investigador sitúa el tema tratado en este trabajo dentro de la categoría genérica de “usos de Wikipedia”.

Dentro de este numeroso grupo de trabajos se pueden establecer varias grandes subdivisiones, aunque, como se ha indicado previamente, hay trabajos que pueden atribuirse a dos o más de ellas. Los grupos que se proponen corresponden a:

- 1) Recuperación de información: se trata de aquellos trabajos en los cuales las categorías se han usado en diferentes procesos y/o técnicas de recuperación de información, tanto

en lo que concierne a la formulación de expresiones de búsqueda, como su refinamiento y mejora, o al filtrado de resultados. Deben destacarse su uso en procesos de evaluación de resultados. También se han usado en procesos de recomendación.

- 2) Procesamiento de entidades: especial interés, también por su número, despiertan los trabajos que buscan identificar entidades (*named entities*) en los documentos. La aplicación de estos estudios es muy amplia, ya que sirven para identificar relaciones semánticas entre términos, resolver problemas de desambiguación, o integrar clasificaciones y taxonomías. Pueden aparecer relacionados con investigaciones sobre procesamiento de lenguaje natural e incluso la elaboración de diccionarios en diferentes lenguas o multilingües.
- 3) Indización y clasificación de corpus documentales: las categorías, o subconjuntos específicos de ellas, han sido usadas como herramienta para proceder a la indización y clasificación de conjuntos de documentos, normalmente dentro de contextos o dominios específicos. Un tipo especial de estos trabajos lo forman aquellos que usan las categorías para etiquetar documentos textuales, en el marco de procesos de indización automática. Otro subconjunto a destacar es el formado por estudios que abordan el etiquetado automático de video o fotografía. Quizá cabe incluir aquí los trabajos que elaboran corpus especializados de forma automática, usando a su vez las categorías de Wikipedia.
- 4) Creación y uso de taxonomías: se trata de uno de los usos más clásicos. Las categorías de Wikipedia son extraídas de su contexto para formar esquemas nuevos, que puedan ser aplicados en dominios específicos. Pueden ser completadas con el uso combinado de otras taxonomías. En relación con esta aproximación se han identificado algunos trabajos que proponen la creación de clasificaciones de estructura clásica, como clasificaciones jerárquicas o tesauros. En numerosas ocasiones las taxonomías creadas se integran en procesos de indización y clasificación de corpus, como se recoge en el punto 3 de esta enumeración.
- 5) Creación y uso de ontologías: el segundo de los usos clásicos. Alrededor de un 15% de los trabajos revisados se ocupan de la creación y uso de taxonomías y ontologías desde el sistema de categorías de Wikipedia. Al igual que las taxonomías indicadas en el punto 4, se usan en procesos de clasificación de documentos, pero también en la ingeniería ontológica y en el desarrollo de relaciones semánticas entre entidades.
- 6) Tratamiento semántico: en este grupo se englobarían diferentes aproximaciones que se caracterizan por estar basadas en los principios y técnicas del web semántico. Se incluirían aquí técnicas como la creación de grafos de categorías, la creación de árboles y esquemas semánticos, la extracción de tripletas, la identificación de relaciones significativas entre términos y su aprovechamiento semántico, etc. Las ontologías, aunque parte integrante del web semántico, se han incluido en un grupo aparte por su importancia.
- 7) Otros usos: dentro de este grupo pueden englobarse aplicaciones muy específicas o poco representadas numéricamente en los trabajos revisados. Ejemplo de ellos pueden ser definición de perfiles de usuario, patrones de edición colaborativa, o la identificación de eventos y personas.

Con el objetivo de identificar con mayor precisión los temas que han sido objeto de investigación se ha elaborado un listado de términos y expresiones significativas, tomadas de los títulos y resúmenes

revisados, y cuyo contenido se recoge a continuación, en correspondencia con los siete grandes grupos delineados en los párrafos anteriores:

- Recuperación de información: Automatic question generation, Automatic subject induction, Discovering answers, Entity retrieval, Exploratory search, Organizing search results, Query classification, Query expansion for entity, ranking, Query phrase expansion, Semantic question answering, Supervised question classification, Tagging, Wikipedia categories for ad hoc search.
- Procesamiento de entidades: Automatic keyword extraction, Computing word relatedness, Context and keyword extraction, Entity disambiguation, Entity retrieval, Entity semantics, Keyphrase extraction, Matching named entities, Named entity extraction, Named entity linking, Query expansion for entity ranking, Semantic tags, Wikipedia entity expansion, Word sense disambiguation.
- Indización y clasificación de corpus documentales: Automatic document classification, Automatic document tagging, Categories for document labelling, Conceptual indexing, Corpus building in machine learning, Document clustering, Document context similarity, Document indexing, Document topics, Multimodal document classification, Relevant features for text classification, Tagging, Text categorization.
- Creación y uso de taxonomías: Analysis of cluster structure, Automatic taxonomy extraction, Comparing taxonomies, Derivation of “is a” taxonomy, Method for refining a taxonomy, Taxonomy and clustering, Taxonomy-based information content, Twixonomy, Web taxonomies, Wikipedia category graph.
- Creación y uso de ontologías: Automated construction of domain ontology taxonomies, Automatic ontology generation, CyC ontology, Deriving domain taxonomies, Domain ontological structure, Evolution of ontologies, Extracting ontologies, Mining concepts, Ontological models, Ontology density, Ontology evaluation, Rich ontology extraction, Semi-automatic ontology creation, Topic ontology.
- Tratamiento semántico: Category annotation recommendation, Category graph, Concept graph, Disambiguation of keyword search results on highly heterogeneous structured data, Domain semantic networks, Entity linking, Knowledge trees, Ontology density, Ranking related entities, Semantic knowledge base, Semantic knowledge extraction, Semantic relatedness, Semantic relationships extraction, Semantic resource extraction, Semantic tagging, Semantically related category hierarchies, Triples extraction.
- Otros usos: Automatic blog classification, Automatic mapping of Wikipedia categories, Automatic thesaurus generation, Concept hierarchies, Dbpedia, Dynamic facet hierarchy construction, Generation of dictionaries, Multilingual domain specific resources, Semantic interest profiles, Semantic knowledge base, Semantic recommender, Semantic tags, Terminology, Wikipedia categories clustering, WordNet, YAGO-NAGA.

Si bien la variedad de términos y expresiones utilizados permite afirmar y reforzar los resultados indicados en la revisión cualitativa, sin embargo pone de manifiesto un problema subyacente a las revisiones sistemáticas, como es la disparidad de criterio de los autores en la redacción de títulos, resúmenes y en la selección de palabras clave. Incluso en algunos casos puede detectarse el uso como sinónimos de términos o expresiones que no lo son. En lo referido a la clasificación e identificación del

contenido de los trabajos con una mayor precisión, pone de manifiesto que los investigadores recurren a enfoques mixtos y combinan métodos y técnicas, lo que dificulta una aproximación tradicional, y requiere métodos de procesamiento automático de la información para obtener mejores resultados.

CONCLUSIONES

Debe admitirse que Wikipedia está teniendo una importante influencia en la forma en la que los usuarios se aproximan a la resolución de sus problemas de información, y con notable influencia positiva (Fallis, 2008). La investigación científica no resulta tampoco ajena a la importancia de este recurso de información (Tomaszewski y MacDonald, 2016). A ello no resulta ajena la investigación en organización y recuperación de información, que encuentra en Wikipedia un banco de pruebas de gran valor (Mehdi et alii, 2017).

La primera conclusión que se extrae del estudio realizado es que Wikipedia es un objeto de investigación de interés para diferentes enfoques de la investigación sobre sistemas y técnicas de representación, organización y recuperación de información, tanto en lo que concierne a sus aspectos internos, como al uso externo de sus datos en otras áreas y enfoques de investigación. En segundo lugar, el sistema de categorías de Wikipedia sirve para analizar la evolución socio-temporal de los procesos de organización del conocimiento en entornos colaborativos. En tercer lugar, hay que destacar su uso como herramienta de apoyo y validación en diferentes tipos de aproximaciones al estudio y análisis de corpus documentales.

Al tratarse los artículos de Wikipedia y su sistema de categorías de un corpus documental en continua evolución, cabe señalar que los resultados obtenidos en los diferentes estudios pueden variar, a medio o largo plazo, en virtud del desarrollo de factores externos e internos a la propia enciclopedia. Sirvan como referencia los sistemas de etiquetado social, objeto de gran interés durante la pasada década, que han ido desapareciendo progresivamente de los intereses de la investigación sobre organización del conocimiento.

Finalmente, cabe destacar el potencial que ofrece el sistema de categorías de Wikipedia, en cuanto se trata de un esquema de clasificación universal desarrollado de forma colaborativa, que se asemeja a un tesoro, y que se contrapone a los esquemas de clasificación especializados elaborados en contextos cerrados. Ello ofrece un amplio campo tanto para la validación o la comparación entre esquemas de clasificación existentes, como para creación de otros nuevos desde una perspectiva que permita combinar ambas aproximaciones a la organización del conocimiento. Varios trabajos recientes inciden en esta cuestión (Salah, Gao, Suchecki y Scharnhorst, 2012; Kiyota et alii, 2009), así como en las ventajas y avances del sistema de Wikipedia frente a clasificaciones tradicionales (Jiménez-Pelayo, 2009).

REFERENCIAS BIBLIOGRÁFICAS

- Capocci, A., Rao, F., & Caldarelli, G. (2008). Taxonomy and clustering in collaborative systems: The case of the on-line encyclopedia Wikipedia. *EPL (Europhysics Letters)*, 81(2), 28006. <https://doi.org/10.1209/0295-5075/81/28006>
- Fallis, D. (2008). Toward a Epistemology of Wikipedia. *Journal of the American Society for Information Science and Technology*, 59(10), 1662-74. DOI 10.1002/asi.20870
- Jiménez-Pelayo, J. (2009). Wikipedia como vocabulario controlado: ¿está superado el control de autoridades tradicional?. *El Profesional de la Información*, 18(2), 188-201. DOI 10.3145/epi.2009.mar.09.
- Kiyota, Y., Nakagawa, H., Sakai, S., Mori, T., & Masuda, H. (2009). Exploitation of the Wikipedia category system for enhancing the value of LCSH. En *2009 ACM/IEEE Joint Conference on Digital Libraries, JCDL '09* (p. 411). <https://doi.org/10.1145/1555400.1555488>
- Mehdi, M., Okoli, C., Mesgari, M., Nielsen, F. Å., & Lanamäki, A. (2017). Excavating the mother lode of human-generated text: A systematic review of research that uses the wikipedia corpus. *Information Processing & Management*, 53(2), 505–529. <https://doi.org/10.1016/j.ipm.2016.07.003>
- Mesgari, M., Okoli, C., Mehdi, M., Nielsen, F.A., & Lanamäki, A. (2015). “The sum of all human knowledge”: A systematic review of scholarly research on the content of Wikipedia. *Journal of the American Society for Information Science and Technology*, 66(2), 219-245.
- Nielsen, F.A. (2017). *Wikipedia research and tools: Review and comments*. Working paper. http://www2.compute.dtu.dk/pubdb/views/edoc_download.php/6012/pdf/imm6012.pdf
- Okoli, C., & Schabram, K. (2009). Protocol for a Systematic Literature Review of Research on the Wikipedia. En *MEDES '09 Proceedings of the International Conference on Management of Emergent Digital EcoSystems*, Art. 74. <https://doi.org/10.1145/1643823.1643912>
- Salah, A. A., Gao, C., Suchecki, K., & Scharnhorst, A. (2012). Need to Categorize: A Comparative Look at the Categories of Universal Decimal Classification System and Wikipedia. *Leonardo*, 45(1), 84–85. <https://doi.org/10.2307/41421810>
- Smiraglia, R.P., & Cai, X. (2017). Tracking the Evolution of Clustering, Machine Learning, Automatic Indexing and Automatic Classification in Knowledge Organization. *Knowledge Organization*, 44(3), 215-233.
- Thornton, K., & McDonald, D.W. (2012). Tagging Wikipedia: collaboratively creating a category system. En *Proceedings of the 17th ACM International Conference on Supporting Group Work*, 219-228. doi: 10.1145/2389176.2389210.
- Tomaszewski, R., & MacDonald, K.I. (2016). A Study of Citations to Wikipedia in Scholarly Publications. *Science & Technology Libraries*, 35(3), 246–261. <https://doi.org/10.1080/0194262X.2016.1206052>
- Tramullas, J. (2015). Wikipedia como objeto de investigación. *Anuario ThinkEPI*, 9, 223–226. <https://doi.org/10.3145/thinkepi.2015.50>

Tramullas, J., Garrido-Picazo, P., & Sánchez-Casabón, A.I. (2016). Research on Wikipedia Vandalism: a brief literature review. En *CERI '16 Proceedings of the 4th Spanish Conference on Information Retrieval*, Art. 15. DOI 10.1145/2934732.2934748

Vrandecic, D. (2013). The rise of wikidata. *IEEE Intelligent Systems*, 28(4): 90–95. DOI 10.1109/MIS.2013.119



CONSTRUÇÃO DE TESAURO NO DOMÍNIO DOS METADADOS PARA DESCRIÇÃO DE RECURSOS BIBLIOGRÁFICOS

Jefferson Leite Oliveira Ferreira¹, Heliomar Cavati Sobrinho², Virgínia Bentes Pinto³

¹Mestrando Programa de Pós-Graduação em Ciência da Informação/Universidade Federal do Ceará
Jefferson.leite@ufc.br

²Programa de Pós-Graduação em Ciência da Informação/Universidade Federal do Ceará
heliomarcavati@yahoo.com.br

³Programa de Pós-Graduação em Ciência da Informação/Universidade Federal do Ceará
vbentes@ufc.br

RESUMO Descreve as etapas envolvidas na construção de um tesauro no domínio dos metadados aplicados à descrição de recursos bibliográficos. Discute as linguagens documentárias como instrumentos que agregam aportes teóricos-epistemológicos da lógica e da semântica descritiva, a partir do referencial teórico e metodológico de Cavati(2014) e Cervantes(2009). Adota o modelo metodológico integrado como diretriz para a organização sistemática do domínio notional e posterior desenvolvimento do percurso empírico da construção do vocabulário com a sistematização de suas etapas e discussão dos resultados.

PALAVRAS-CHAVE *Metadados descritivos, domínio bibliográfico, representação da informação, tesauro.*

ABSTRACT Describes the steps involved on building of metadata domain thesaurus applied to bibliographic resources description. Discusses documentary languages as tools that aggregates epistemological-theoretic backgrounds from logic and semantic descriptive, starting from Cavati (2014) and Cervantes (2009). Adopts integrated methodological model as guideline to systematic organization of notional domain and later development of empirical roadmap aiming the building of vocabulary with systematization of its steps and results discussions.

KEYWORDS *Descriptive metadata, bibliographic domain, information representation, thesaurus.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Desenvolver linguagens documentárias implica empreender esforços que contemplam, além das tarefas de análise, seleção e organização de termos, o adequado tratamento da linguagem representativa do domínio. Não se trata de uma tarefa pragmática ou um exercício técnico linear e assistemático; envolve um arcabouço teórico-metodológico capaz de gerar uma representação enriquecida de significado semântico e contextual. Isto implica que no contexto de um vocabulário controlado, a representação promovida pela linguagem como instancia simbólica do real assume a dimensão das conexões possíveis não-lineares e dinâmicas. Para Caixeta e Souza (2013, p. 39) uma outra dimensão da linguagem era

representada pelos gregos pela palavra *logos*, que significava a palavra racional do conhecimento real (daí a lógica). No âmbito da estrutura lógica das linguagens documentárias enquanto instrumentos de representação do conhecimento, a representação pode ser tratada como “doutrina segundo a qual nosso conhecimento nunca é absoluto, mas sempre flutuante, como se estivesse em um *continuum* de incerteza e indeterminação”(COSTA; MOURA, 2013. p. 13). Isto posto, considera-se necessário reconhecer o caráter multidisciplinar do exercício de construção das LD’s, concordando com Cervantes (2009, p. 27) ao asseverar que o conceito de representação se relaciona com a Psicologia, Linguística, Semiótica, Ciência da Informação, especialmente, Documentação e outras.

Em particular e no que concerne às linguagens documentárias, a tarefa de tradução da linguagem natural para a linguagem do sistema ou controlada deve considerar a mínima perda ou ruído na pós-coordenação. Trata-se de signos que, embora estejam agregados estruturalmente, não devem prescindir de uma semântica capaz de propiciar relações significativas para o alcance da precisão terminológica.

No caso específico da representação da informação é possível estabelecer uma relação semiótica entre o conteúdo temático dos documentos e os termos e expressões adotados para representá-los nos sistemas de recuperação de informação. (COSTA; MOURA, 2013. p. 20)

Para a elaboração de um tesouro, os termos são estruturados, seguindo um triplo sistema de relações que servem de delimitador do conteúdo semântico de cada termo e os vínculos que os relacionam (equivalência, hierarquização e associação semântica) (CAVATI, 2014, p. 45).

Este artigo descreve as etapas da construção de um tesouro no domínio dos metadados para descrição de recursos bibliográficos, que segundo a IFLA, estes recursos, por ocasião dos avanços tecnológicos das últimas décadas, passaram a incorporar novas formas de expressão para além do documento textual, o que implicou no desenvolvimento de novos esquemas de metadados para sua representação.

METODOLOGIA

A partir do modelo integrado proposto por Cervantes (2009) e de sua implementação na pesquisa de Cavati (2014), ambos descritos em pesquisa de doutorado, será apresentado a seguir, a sistematização necessária ao desenvolvimento da linguagem proposta, que contempla os seguintes passos:

- Trabalho preliminar: consiste na escolha da área de especialidade, que envolve, também, a análise do usuário em termos de levantamento de suas demandas.
- Método de compilação: nesta etapa são consideradas as abordagens dedutiva e indutiva com ênfase no alcance de maior flexibilidade para a estruturação do tesouro. Inclui a coleta do *corpus* do trabalho terminológico, estabelecimento da árvore de domínio e a expansão da representação do domínio escolhido.
- Registro de termos: Envolve a coleta e classificação dos termos a partir do *corpus* selecionado na etapa precedente. Implica no julgamento do termo mais adequado para representar determinado conceito, segundo método analítico proposto por Rondeau (1984).
- Verificação de termos: esta etapa consiste na verificação da adequabilidade entre termos e conceitos, o que traz benefícios como aumento da precisão nas relações a serem estabelecidas e na classificação definitiva.

- Forma de apresentação do tesouro: apresenta a estrutura final do vocabulário em sua versão definitiva.

As etapas descritas acima não esgotam as diversas propostas para a construção de linguagens documentárias, mas atendem aos propósitos para sua elaboração na medida em que primam pela objetividade e contemplam os requisitos apoiados na literatura especializada para esta finalidade.

Ao justificar a adoção do modelo metodológico integrado para construção de tesouros, Cavati (2014, p. 97) classifica o modelo de Cervantes como exequível para construção de LD's. Acrescentamos a relevância dos requisitos necessários a esta tarefa, que incluem da caracterização do usuário às formas de apresentação do tesouro.

Das fases preliminares do trabalho, foram adotadas a escolha do domínio e da língua do tesouro bem como a delimitação do subdomínio, delineadas no quadro 1 abaixo. A delimitação do subdomínio corresponde a um recorte necessário para contemplar um universo de termos e conceitos e suas relações que, por si só, já tem o condão de representar com elevado rigor de detalhe o escopo conceitual ou ainda as redes nocionais. Cervantes (2009, p. 147).

Para esta pesquisa foi escolhido o domínio bibliográfico como categoria mais geral para análise terminológica, em seguida, circunscrevemos o subdomínio com o repertório de descritores extraídos dos seguintes vocabulários: Biblioteca Nacional, Rede Pergamum e *Library of Congress Subject Heading*.

Os descritores analisados e extraídos correspondem aos metadados empregados na descrição de recursos bibliográficos, que, por sua vez, integram o domínio ora em estudo.

Quadro 1: Definição das etapas preliminares. Adaptado de Cervantes (2009, p. 163)

ETAPAS PRELIMINARES	
Escolha do domínio e da língua do tesouro	<ul style="list-style-type: none"> • Domínio bibliográfico em tesouro de língua portuguesa e inglesa
Delimitação do subdomínio	<ul style="list-style-type: none"> • Metadados empregados na catalogação de recursos bibliográficos e extraídos dos cabeçalhos da Biblioteca Nacional, Rede Pergamum e <i>Library of Congress</i>.

A compilação inicial dos descritores consta no apêndice A, que inclui a terminologia mais geral do subdomínio consultado bem como as fontes utilizadas para tanto. Atenção especial foi dispensada para eventuais homógrafos e sinônimas com vistas à garantia da consistência na representação conceitual e na posterior construção das hierarquias e associações condizentes ao domínio e usualmente adotadas pela comunidade.

Foram compilados um total de 50 termos representativos do domínio bibliográfico e divididos em categorias para análise e delineamento da árvore de domínio, quadro 2 abaixo, de acordo com a hierarquização nas fontes consultadas. O repertório terminológico compilado foi inicialmente arrolado em planilha Excel e separado por fontes consultadas.

Agência	Qntd. de descritores identificados
Rede Pergamum	17
Biblioteca Nacional	15
Library of Congress	18

Na ocorrência de duplicidades de um mesmo descritor, foi dada preferência àqueles constantes no cabeçalho da *Library of Congress*, pela sua relevância como fonte de informação, aplicando-se o mesmo raciocínio entre a Rede Pergamum e a BN. Neste caso, a escolha recaiu para a segunda opção em virtude da cobertura temática do referido repertório terminológico.

Estas etapas envolvem o julgamento dos termos mais adequados, ou candidatos, a inclusão no tesauro e implica a análise dos conceitos atrelados do geral ao específico (top down) ou, ainda, do específico ao geral (bottom up) segundo modelo proposto por Cervantes (2009, p. 202) .

O levantamento dos conceitos adotou como fonte de informação os próprios cabeçalhos de autoridade, uma vez que tais conceitos remetem a entradas autorizadas empregadas na catalogação o que atesta sua validade e, por outro lado, compila, organiza e sistematiza a relação terminológica no domínio bibliográfico com mais exatidão e razoável exaustividade.

Na etapa denominada “abordagem empírica” Cervantes arrola dois métodos que auxiliam a tarefa do indexador, ou comitês, no julgamento da organização terminológica. Trata-se do método dedutivo que desdobra os termos em relações hierárquicas do mais geral ao mais específico e do indutivo que segue o caminho inverso com ênfase particular na inclusão de novos termos.

Ao desenvolver o tesauro de domínio bibliográfico foram adotados, recursivamente, os dois métodos acima descritos com o objetivo de assegurar a coerência interna do arranjo hierárquico, sem prescindir da garantia literária mediante consultas a artigos de periódicos, anais de eventos, coletâneas, resenhas críticas e bases de citações. A título de exemplo, o descritor “*Information Retrieval*” consta no cabeçalho da *Library of Congress* como um termo genérico a partir do qual se vincula outros como *Cross-language information retrieval* e *Information Services*. Tal arranjo é ratificado na literatura em Biblioteconomia na medida em que as fontes supracitadas validam a hierarquia conceitual e conseqüentemente o desdobramento terminológico.

A forma de apresentação do tesauro adotou o modelo hierárquico convencional com a indicação do termo geral seguido dos termos relacionados constantes no apêndice B.

CONCLUSÃO

Definir estratégias de representação da informação por meio de vocabulários terminológicos constitui prática de forte engajamento inter e multidisciplinar. O modelo metodológico integrado para sua elaboração, na proposta de Cervantes e Cavati, descreve uma seqüência operativa que abstrai detalhes específicos da incorporação do vocabulário controlado a dispositivos tecnológicos, na medida em que se concentra nos aspectos indispensáveis à sua elaboração, tais como a participação de comissões especializadas, delimitação de domínio e subdomínio, registro e verificação de termos dentre outros.

A construção do tesauro no domínio dos metadados para descrição bibliográfica traz diversos e potenciais benefícios ao campo da Biblioteconomia e da Ciência da Informação, dentre os quais o suporte a levantamentos historiográficos do domínio bibliográfico e a inserção, ao longo do tempo, das tecnologias de informação na abordagem dos metadados.

Sabe-se que a finalização de um tesauro não implica a conclusão definitiva de uma tarefa. Conceitos são regularmente revistos, categorias e subcategorias de termos são modificados na medida em que

novas construções teóricas e novos paradigmas forçam o abandono de um termo em detrimento a outro e assim por diante. Na condição de um vocabulário, o tesouro para ser consistente na sua representação e útil como fonte para indexação e recuperação da informação, deve estar inserido em políticas contínuas de revisão e edições que se fazem necessárias

REFERÊNCIAS BIBLIOGRÁFICAS

Caixeta, M., & Souza, R. R. (2013). Representação do conhecimento: história, sentimento e percepção. *Inf. Inf.*, 18(3), 45-67.

Cavati, H. (2009). *A representação documentária do domínio da economia: análise de estruturas de representação em linguagens documentárias e documentos específicos de economia*. São Paulo / SP: UNESP.

Cervantes, Brígida Maria Nogueira (2009). *A construção de tesouros com a integração de procedimentos terminográficos*. Marília / SP , Brazil: UNESP.

Costa, Mateus Uerlei Pereira da, & Moura, M. A. (2013). A representação da informação em contextos de comunicação científica: a elaboração de resumos e palavras-chave pelo pesquisador- autor. *Inf. Inf.*, 18(3), 45-67.

International Federation of Library Associations and Institutions. (2016). *Functional Requirements for Bibliographic Records*. IFLA, 142.

International Organization for Standardization. (1986) ISO 2788-1986: Documentation - Guidelines to establishment and development of monolingual thesauri.

Geneva : ISO.

Rondeau, G. (1984) *Introduction à laterminologie*. (2a ed). Québec, Canadá: Gaëtan Morin



O USO DE VOCABULÁRIO DE NOMES POPULARES DE DOENÇAS COMO FERRAMENTA DE REPRESENTAÇÃO DA INFORMAÇÃO EM SAÚDE

Virgínia Bentes Pinto¹, Lidia Eugenia Cavalcante², Hamilton Rodrigues Tabosa³,
Odete Máyra Mesquita Sales⁴

¹Universidade Federal do Ceará, 00003-1283-8292, vbentes@ufc.br

²Universidade Federal do Ceará, 0000-0002-3190-6900, lidia@ufc.br

³Universidade Federal do Ceará, 0000-0002-2757-195X, hamilton.rt@ufc.br

⁴Universidade Federal do Ceará, 0000-0002-9208-3071, mayra.mesquita@ufc.br

RESUMO Apresenta os resultados da pesquisa sobre representação da informação em saúde, tendo como objetivo geral construir um vocabulário controlado de termos populares de doenças, na perspectiva de reduzir os riscos de interferências decorrentes da utilização de terminologias não coincidentes, melhorando o processo de comunicação entre médico e paciente. A metodologia utilizada concerne em pesquisa descritiva aplicada, pautada na fenomenologia, sendo o estudo empírico realizado junto à comunidade dos índios Tremembés, situada no município de Almofala – Ceará. Como instrumento de coleta de dados, utilizou-se a entrevista semi-estruturada, aplicada *in loco*. Como resultado, foram mapeados cento e quarenta termos populares referentes a nomes de doenças identificadas pela população pesquisada. Estruturam-se esses termos em um Mapa Conceitual, construído utilizando o software Cmap Tools desenvolvido e disponibilizado gratuitamente pelo *Institute for Human & Machine Cognition (IHMC) - University of West Florida*, disponível no site <http://cmap.ihmc.us/>. Conclui-se que os processos cognitivos que envolvem a construção de mapas conceituais e sua utilização nas áreas pesquisadas (Ciência da Informação e Ciências da Saúde) potencializam a realização de estudos que envolvem diferentes domínios, dentre eles: terminologia, linguística, antropologia cultural, tecnologias digitais e representação da informação e do conhecimento.

PALAVRAS-CHAVE *Representação da informação em Saúde, Vocabulário controlado, Termos populares de doenças, Mapas conceituais.*

ABSTRACT It presents the research results on health information representation, in which the general objective concerns the construction of a controlled vocabulary of popular terms of diseases. It is focusing on reducing the risks of interferences caused by the use of non-coincident terminologies, improving the communication process between doctor and patient. The approach applying descriptive research is based on the phenomenology. The empirical study was carried out with the community of Tremembés Indians, located in the municipality of Almofala—Ceará. As a data-collection tool, the semi-structured interview was performed locally. As a result, the surveyed population mentioned one hundred and forty popular terms referring to names of diseases that were mapped. These terms are structured in a conceptual map built using the CMAP Tools software developed and made available for free by the Institute for Human & Machine Cognition (IHMC)—University of West Florida, available at <http://cmap.ihmc.us/>. It is concluded that the cognitive processes involving the construction of conceptual maps and their uses can be applied to the research areas including the Information Science and Health Sciences. Also, it makes possible to carry out studies concerning different domains, such as terminology, linguistics, cultural anthropology, technologies and representation of information and knowledge.

KEYWORDS *Health information representation, Controlled vocabulary, Popular disease terms, Conceptual maps.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

As denominações das doenças acompanham a evolução científica e tecnológica das disciplinas que compõem a área da Saúde, notadamente a Medicina. Na formação do léxico dessa área, normalmente, os termos são formados por prefixos e sufixos originários do grego e do latim. Esse conjunto de vocábulos vai se constituir na chamada terminologia de especialidades. Porém, além dessa terminologia, no processo de comunicação entre os profissionais da saúde e os pacientes, entram em cena, por parte desses últimos, outros termos que constituem o vocabulário de cada cidadão, comunidade, estrato social, grupos de faixa etária etc. Assim, os profissionais da Saúde vão incorporando também esse vocabulário do cotidiano cultural que nomeia as doenças, os signos e os sintomas ao seu repertório. Essa inserção se dá de forma natural ou espontânea fazendo que, no processo de comunicação, principalmente quando da anamnese, sejam registrados nos prontuários do paciente, os termos enunciados pelos pacientes.

Acontece que esse processo de comunicação envolvendo termos populares e científicos pode trazer consequências desastrosas, conforme salienta Dahlberg (1978, p. 102): “Quando, porém, se trata de linguagens especializadas as conseqüências [no ue se refere a conceitos mal compreendidos entre os envolvidos no processo comunicacional] podem ser desagradáveis. Neste caso deve-se fazer todo esforço para que os conceitos sejam definidos com toda precisão.” Essa é a realidade verificada no âmbito da Saúde aqui investigado, quando esses atores não falam a mesma língua, como é o caso dos médicos que vieram de outros países por meio do Programa “Mais Médicos”, do Ministério da Saúde/Governo Federal, para atuarem no Brasil, considerando a possibilidade de polissemia ou mesmo de esvaziamento semântico da comunicação médico-paciente e vice-versa.

O referido programa se configura como uma Política de Governo e foi instituído pela Lei Nº 12.871, de 22 de outubro de 2013, trazendo no Capítulo I, entre outros, os seguintes objetivos:

I - diminuir a carência de médicos nas regiões prioritárias para o SUS, a fim de reduzir as desigualdades regionais na área da saúde; II – fortalecer a prestação de serviços de atenção básica em saúde no País; IV - ampliar a inserção do médico em formação nas unidades de atendimento do SUS, desenvolvendo seu conhecimento sobre a realidade da saúde da população brasileira. (BRASIL..., 2013, p. 01).

Os médicos, majoritariamente vindos de Cuba, tiveram dois grandes desafios:

- 1) Conhecer os termos médicos em língua portuguesa;
- 2) Relacionar os termos populares, específicos de cada região brasileira, aos termos técnicos da área da Saúde.

No tocante ao aperfeiçoamento desses profissionais, assunto tratado no Capítulo IV, o parágrafo 3º do Art. 14, determina que o primeiro módulo, além de outras coisas, “contemplará conteúdo relacionado à legislação referente ao sistema de saúde brasileiro, ao funcionamento e às atribuições do SUS [Sistema Único de Saúde], notadamente da Atenção Básica em Saúde, aos protocolos clínicos de atendimentos

definidos pelo Ministério da Saúde, à **língua portuguesa** e ao código de ética médica”. (BRASIL..., 2013, p. 8, grifo nosso).

Observa-se que tal programa visa aproximar esses médicos da população na qual estão atuando, porém, mesmo com todo esse esforço, haverá interferências quando do processo de comunicação entre eles e os pacientes, pois, nesse processo, estão embutidos os aspectos do multilinguismo e multiculturalismo, por se tratar de sujeitos oriundos de outras culturas.

Em que concerne ao conceito de cultura, Laraia (2001) destaca que o homem compreende o mundo por meio das suas vivências, considerando o seu modo de vida como algo natural. Assim, a cultura se configura como um sistema simbólico de significados, correspondentes às práticas e comportamentos antropológicos e sociais, transmitidos de geração a geração, como por exemplo, linguagens, línguas, ritos, maneiras de vestir, gastronomia, credo etc.

Isto posto, esta pesquisa tem como objetivo geral: construir um vocabulário controlado de termos populares de doenças, na perspectiva de reduzir os riscos de interferências decorrentes da linguagem divergente entre médicos e pacientes, melhorando o processo de comunicação.

Os objetivos específicos são:

- 1) Mapear, na literatura, os termos populares utilizados para nomear as doenças;
- 2) Realizar estudo empírico nas comunidades selecionadas para a pesquisa no Estado do Ceará, a fim de identificar quais termos populares são empregados para nomear doenças em geral;
- 3) Correlacionar os termos populares utilizados pelos sujeitos da pesquisa com outros termos em relação sinonímica, por meio de um mapa conceitual.

Nas profissões relativas às disciplinas que compõem as áreas da Saúde e da Ciência da Informação, notadamente a Biblioteconomia, a Museologia e a Arquivologia, o multiculturalismo também está presente, embora não seja amplamente percebido na literatura e absorvido pelos sujeitos em suas práticas profissionais. Essas áreas estão ligadas de modo particular ao contato direto com as pessoas, seja do ponto de vista de um problema de saúde, ou das demandas de informações para resolver questões do cotidiano da vida do cidadão ou de pesquisas, sendo inevitavelmente carregadas de marcas culturais e linguísticas.

Por isso, nas áreas citadas, é importante evidenciar o multiculturalismo na construção de vocabulários controlados, levando em consideração os termos populares das doenças, para que, no processo de comunicação médico-paciente, as necessidades dos indivíduos sejam contempladas, respeitando-se as particularidades de cada cultura. Agindo assim, o processo de comunicação entre os sujeitos, e o atendimento às demandas poderá se efetivar com maior valor agregado e menos interposição, para que as consequências negativas possam ser evitadas.

CONCEPÇÕES TEÓRICAS SOBRE LÉXICO E VOCABULÁRIO CONTROLADO

No âmbito das linguagens escritas e faladas, entram em cena o vocabulário e o léxico. Conforme Vilela (1994, p. 10), o léxico diz respeito à “[...] totalidade das palavras duma língua, ou, como o saber

interiorizado, por parte dos falantes de uma comunidade linguística”. Para Rey-Debove e Morais (1984), o léxico é o domínio menos específico da linguagem, reportando-se ao universo referencial. Os autores argumentam que o conjunto de signos que compõem o léxico confere a ele um estatuto semiótico, criando um sistema de mundo, sendo um conjunto aberto, com infinitas possibilidades de expansão. Isso certamente acompanha a dinamicidade da comunicação humana, sempre em movimento, posto que criamos, transformamos, usamos e abandonamos representações simbólicas verbais e não-verbais que nos possibilitam o referenciamento linguístico do real e do irreal que fazem parte das nossas vidas em sociedade.

Por sua vez, o vocabulário constitui-se de uma parte do léxico que representa uma determinada área de conhecimento (BARBOSA, 1995). Na visão de Aguiar (2008), o vocabulário é um recurso normativo para nomear as atividades e os procedimentos funcionais.

Logo, esses dois elementos estão presentes no processo de comunicação entre os indivíduos que buscam socorro nas organizações de saúde, os médicos e a equipe multiprofissional, que vão atuar nos cuidados no que tange à questão emergencial do atendimento básico.

Além do léxico e do vocabulário, também é ímpar se pensar nos conceitos, quando da construção de linguagens documentárias, que são consideradas como “um conjunto de termos, providos ou não de regras sintáticas, utilizado para representar conteúdos de documentos técnico-científicos, com fins de classificação ou busca retrospectiva de informações”. (GARDIN et. al., 1968 *apud* CINTRA et. al., 1994, p. 25). Para Lancaster (2004), as linguagens controladas se configuram como uma lista de termos autorizados que inclui, em geral, uma forma de estrutura semântica, que visa, principalmente:

- a) Controlar sinônimos, por meio da padronização de uma forma única de entrada, porém, com remissivas para todas as outras. Assim, nesta pesquisa levantaremos os termos populares utilizados para nomear as doenças e faremos as relações com a terminologia de especialidades da área da Saúde;
- b) Estabelecer diferenças para termos homógrafos. Do mesmo modo será estabelecida a desambiguação dos termos ou conceitos mapeados;
- c) Reunir ou ligar termos que tenham relações mais estreitas com seus significados, por meio das relações hierárquicas e associativas.

Assim, os vocabulários controlados são um tipo de linguagem documentária e vêm sendo construídos na perspectiva de contribuir para a redução de interferências no processo de comunicação em sistemas de recuperação de informação.

Quanto aos conceitos, Dahlberg (1978) afirma que ele é constituído de elementos que se articulam numa unidade estruturada. Conforme Vygotsky (2001, p. 349), na formação dos conceitos espontâneos e científicos, há fortes influências recíprocas, haja vista que

O conceito espontâneo, que passou de baixo para cima por uma longa história em seu desenvolvimento, abriu caminho para que o conceito científico continuasse a crescer de cima para baixo, uma vez que criou uma série de estruturas indispensáveis ao surgimento de propriedades inferiores e elementares do conceito. De igual maneira, o conceito científico, que percorreu certo trecho de seu caminho de cima para baixo, abriu caminho para o desenvolvimento dos conceitos espontâneos, preparando de antemão uma série de formações estruturais indispensáveis à apreensão das propriedades superiores do conceito. Os conceitos científicos crescem de cima para baixo através dos espontâneos. Estes abrem caminho para cima através dos científicos.

Na perspectiva desse autor, os conceitos científicos e espontâneos não são excludentes, muito pelo contrário, um alimenta o outro em uma relação natural. Outrossim, conforme defendem Smit e Kobashi (2003), a construção de um vocabulário se efetiva por meio dos conceitos que devem refletir as terminologias das áreas de conhecimento e a linguagem do público-alvo.

Boccatto (2011) fez um estudo das normas que regem, entre outras coisas, a construção de vocabulários e apresenta de forma resumida o que cada uma propõe, conforme apresentado no Quadro 1.

Quadro 1 – Demonstrativo das ocorrências mais significativas disponíveis nas normas ANSI/NISO Z39.19-2005, BS 8723 e ISO 25964

OCORRÊNCIAS	ANSI/NISO Z39:19-2005	BS 8723	ISO 25964
Título	Guidelines for the Construction, Format, and Managment of Monolingual Controlled Vocabularies.	Structured vocabularies for information retrieval - guide.	Information and documentation - Thesauri and interoperability with other vocabularies.
Estrutura	Onze seções e seis apêndices: ampliação e atualização da norma ANSI/NISO Z39:10-2003.	Cinco partes: a parte 2 substitui a norma BS 5723-1987; a parte 4 substitui a norma BS 6723-1987.	Duas partes: a parte 1 corresponde as partes 1, 2 e 5 da BS 8723. A parte 2 corresponde às partes 3 e 4 da BS 8723. Já a ISO 25964 substitui as normas ISO 2788-1986 e ISO 5964 -1985.
Vocabulário controlado	Listas, anéis de sinônimos, taxonomias, redes semânticas e tesauros.	Sistemas de classificação, sistemas de classificação de negócios para gerenciamento de registros, redes semânticas, taxonomias, listas de cabeçalhos de assunto, ontologias.	Sistemas de classificação, sistemas de classificação de negócios para gerenciamento de registros, listas de cabeçalhos de assunto, taxonomias, mapas conceituais, redes semânticas, terminologias/banco de termos, listas de autoridades de nomes e anéis de sinônimos.
Pontos de atualização/ implicação	Interoperabilidade entre vocabulários controlados; aplicação da análise facetada na construção de tesauros; vocabulários controlados multilíngues; definição de protocolos e formatos necessários para o intercâmbio de dados	Interoperabilidade entre vocabulários controlados; aplicação da análise facetada na construção de tesauros; vocabulários controlados multilíngues; definição de protocolos e formatos necessários para o intercâmbio de dados	Interoperabilidade entre vocabulários controlados; aplicação da análise facetada na construção de tesauros; vocabulários controlados multilíngues; definição de protocolos e formatos necessários para o intercâmbio de dados

Fonte: BOCCATO (2011, p. 11)

Gardin, desde 1966, já chamava atenção para as questões de construção de vocabulários e defendia a necessidade desse trabalho para que o acesso à informação fosse efetuado com menos interferências. Assim, ele estruturou um modelo de vocabulários, no qual defende que a elaboração desses instrumentos linguísticos e terminológicos está inserida em uma proposta lexicográfica, por isso, devem-se observar os aspectos quantitativos do léxico.

METODOLOGIA

O estudo em lide é uma pesquisa aplicada, de cunho analítico-descritivo, iniciada por uma revisão de literatura sobre os temas centrais abordados no estudo, e baseada na Fenomenologia como método de investigação, pois conforme Sadala (2014), ele se mostra mais adequado para nortear pesquisas acerca das questões humanas do que o método cartesiano, não excluindo a importância das pesquisas estatísticas, mas agregando novas perspectivas e ampliando o universo do conhecimento.

Creswell (1998) conceitua o método fenomenológico como sendo a descrição das experiências vividas de vários sujeitos sobre um conceito ou fenômeno, com o objetivo de buscar uma estrutura essencial.

Anteriormente à pesquisa empírica, foi realizado um levantamento em fontes terminológicas especializadas em nomes populares de doenças do Nordeste brasileiro. Nesse levantamento, foram mapeados cerca de 220 (duzentos e vinte) termos que, posteriormente, foram comparados com aqueles coletados junto à comunidade pesquisada.

Alicerçados na abordagem fenomenológica, realizou-se também uma pesquisa empírica na comunidade indígena Tremembé, localizada em Almofala, distrito de Itarema-CE, situada a 224 km da capital cearense, onde se concentra considerável população indígena. Tal comunidade foi selecionada devido a sua estrutura sociocultural e às peculiaridades linguísticas regionais.

A coleta de dados foi realizada no mês de setembro de 2016, onde foram entrevistados o pajé, o cacique e as benzedeiras da aldeia. Mediante entrevistas, colhemos os termos populares, utilizados pela referida comunidade, a fim de denominar as doenças com as quais esses indivíduos são eventualmente acometidos.

Durante as entrevistas, os participantes foram indagados sobre os nomes das doenças que, costumeiramente, acometem as pessoas da comunidade. Depois de relatarem aquilo de que lembraram, os pesquisadores questionaram sobre o léxico dos nomes populares das doenças identificados nas fontes terminológicas.

As entrevistas foram gravadas com auxílio de um aparelho celular, anotadas e posteriormente transcritas, de modo a favorecer a tabulação e análise dos dados, para que, em seguida, fosse realizada a etapa de mapeamento dos termos, resultando na construção de um mapa conceitual a fim de estruturar a terminologia proposta e evidenciar os resultados da pesquisa.

RESULTADOS

O estudo ora apresentado, obteve como resultado um vocabulário com 140 (cento e quarenta) termos populares obtidos na pesquisa empírica e cotejados a partir dos 220 (duzentos e vinte) termos identificados na pesquisa bibliográfica e suas possíveis denominações conceituais sinonímicas. Tais achados foram estruturados em um mapa conceitual, elaborado na ferramenta *CmapTools*, desenvolvida e disponibilizada gratuitamente pelo *Institute for Human & Machine Cognition (IHMC) - University of West Florida*, disponível no site <http://cmap.ihmc.us/>.

O mapa conceitual foi construído com a finalidade de auxiliar no processo de comunicação entre médicos e pacientes estabelecendo-se relações hierárquicas e associativas entre os termos e conceitos,

história, a evolução dos conceitos científicos parece ter se sustentado. É importante frisar que, na literatura sobre a construção de vocabulários controlados, a tônica é de que eles devam não somente refletir a terminologia da área, mas também, que sejam voltados para os usuários a quem se destinam, como argumentam Smit e Kobashi (2003). Nas reflexões teóricas e pragmáticas desse tema, portanto, o léxico popular também deve ser observado. Por todos esses aspectos, é que Cabré (2003, p. 12) chama atenção para a elaboração de vocabulários considerando uma normalização de base sociolinguística, a fim de que se preserve a língua e a cultura das comunidades. Além de eles nominarem os objetos, também servem para expressar as diferenças conceituais relacionadas às culturas das comunidades ou grupos sociais desempenhando papel fundamental para a preservação cultural. “o caráter multidimensional da linguagem se reflete nos termos enquanto unidades que formam parte dele e que se atualizam dentro de uma língua determinada.” (CABRÉ, 1999, p. 148).

Afora esses termos, foram mapeados 67 (sessenta e sete) outros, cuja ambiguidade ficou evidente, posto que, se um participante expressasse um termo, os demais iam complementando com outro para nomear a mesma doença. Então, observando esse fato, consideramos mais razoável estruturarmos um mapa, a fim de que fosse viável expressar esses achados de modo contundente. Vejam-se na Figura 2 os termos populares de doenças que possuem outras denominações. Isso demonstra como a cultura dos índios Tremembés de Almofala encontra-se preservada no aspecto linguístico, no contexto desta pesquisa, mesmo que convivam em ambiente citadino. Corroborando com essa ideia, Cunha (1995, p. 8) afirma que “[...] a representação de uma história, referenciada em sua experiência de vida é que orienta sua forma de estar no mundo.” Ou seja, a comunidade guarda em sua cognição uma linguagem marcante de suas origens.

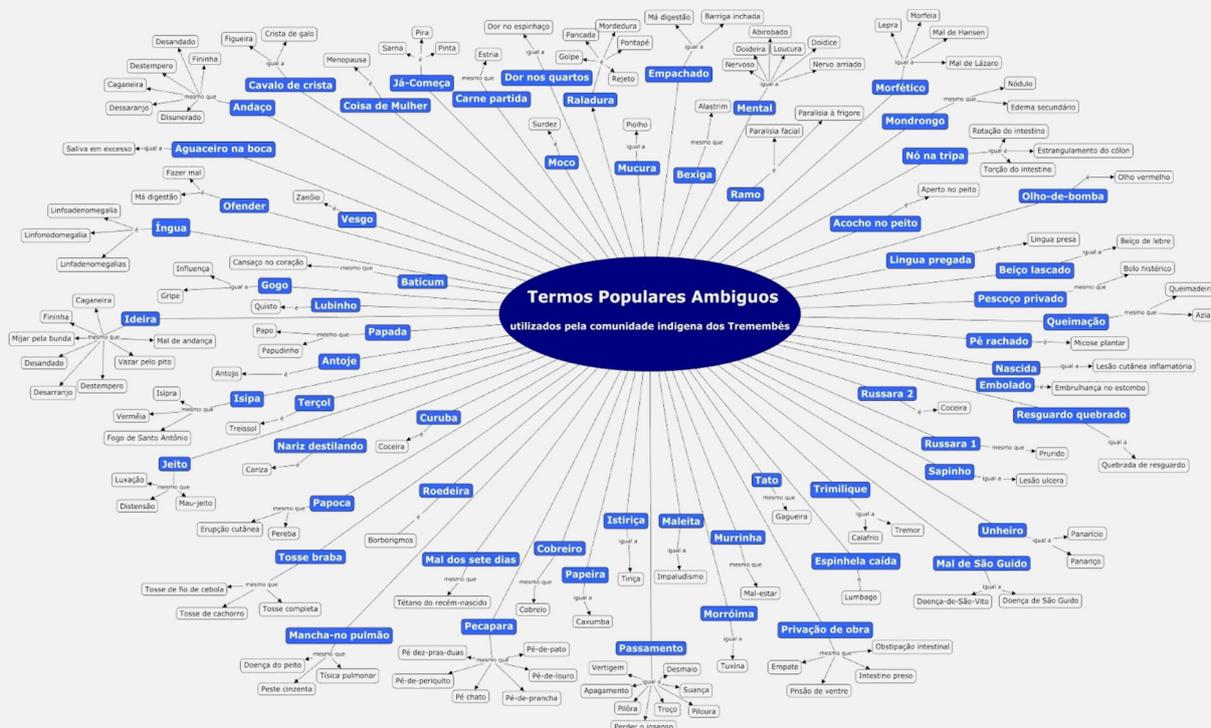


Figura 2 – Termos populares ambíguos que a comunidade pesquisada adota para nomear as doenças

Fonte: Dados da pesquisa in loco

Os resultados desta pesquisa evidenciam que, no momento das ações de cuidados de um paciente, em qualquer organização de saúde, é ímpar que a comunicação entre os envolvidos se efetive com certa

fluidez. Sabe-se que essa comunicação ocorre ao menos por meio de duas linguagens: uma de especialidades e outra popular.

Entendemos que os nomes populares das doenças adotados pelas comunidades faz parte de um repertório de seu léxico e está inserido em uma cultura. Por isso, Martins (1989, p. 111) defende que tal cultura deve ser percebida “[...] como conhecimento acumulado, sistematizado, interpretativo e explicativo [...]”. Isso poderá trazer contribuições ímpares na atenção básica de cuidados no contexto da saúde.

CONCLUSÕES

Os mapas conceituais são ferramentas que trazem efetivas contribuições e favorecem tanto a área da Ciência da Informação, quanto a área de Ciências da Saúde. Vale salientar que, no contexto do “Programa Mais Médicos”, o vocabulário produzido a partir do mapa conceitual tende a possibilitar melhor comunicação e integração entre os atores envolvidos, tendo potencial de tornar a mediação da linguagem entre médico e paciente mais dinâmica, facilitando as conexões relacionadas à anamnese e às ações de cuidado do paciente.

Embora a metodologia de mapa conceitual não tenha sido pensada na perspectiva da construção de vocabulários controlados, ainda assim, evidenciou-se, no decorrer da pesquisa, que essas ferramentas contribuem positivamente na compreensão lexicográfica, uma vez que facilitam as relações evidentes entre os componentes conceituais do léxico proposto. Assim, os processos cognitivos que envolvem o mapa conceitual e sua utilização nas áreas pesquisadas potencializam a realização de estudos que envolvem diferentes domínios, dentre eles: Terminologia, Linguística, Antropologia Cultural, Tecnologias Digitais e Representação da Informação e do Conhecimento.

Infere-se ainda que, na representação temática da informação, sejam levados em consideração os léxicos das comunidades, de modo a contribuir no processo de comunicação e para a preservação da cultura.

Consideramos ainda que, uma ação interessante do Ministério da Saúde seria a divulgação de um vocabulário controlado que levasse em consideração os termos de especialidades relacionados aos termos populares, junto às unidades de saúde, principalmente, no caso daquelas atendidas pelo “Programa Mais Médicos”, cujos integrantes, em sua maioria, são oriundos de outros países. Essa proposta poderia contribuir para facilitar a comunicação entre médico e paciente.

Salientamos ainda que os resultados a que chegamos com este estudo, naturalmente, referem-se à realidade investigada com certas limitações, dada a necessidade de se fazerem recortes e determinadas escolhas metodológicas que assegurassem a exequibilidade da pesquisa. Uma dessas escolhas diz respeito aos sujeitos entrevistados, que podem trazer em seu vocabulário, traços da sua cultura indígena. Os pesquisadores poderiam ter conseguido outros achados se, ao invés de coletar dados em uma tribo indígena, tivessem realizado a pesquisa empírica em favelas situadas no Estado do Ceará, onde certamente a linguagem popular teria suas peculiaridades.

Sabe-se que o Programa Mais Médicos não se limitou ao Estado do Ceará e que cada região brasileira reflete em seus cidadãos aspectos de identidade e cultura locais, de modo que esta pesquisa poderia ser reproduzida em outros espaços geográficos, possibilitando como produto final um vocabulário de termos populares adequado às mais variadas regiões do país.

Por fim, frisamos que, verdadeiramente, os vocabulários controlados em linguagens de especialidade têm o potencial de facilitar a comunicação entre os especialistas e o público em geral em qualquer área do conhecimento.

AGRADECIMENTOS

Os autores agradecem ao CNPq pelo financiamento da pesquisa. Universal- Processo- 456624/214-7

REFERÊNCIAS BIBLIOGRÁFICAS

- Aguiar, F. L. (2008). *O controle de vocabulário como dispositivo metodológico para a organização, tratamento e recuperação da informação arquivística* (Dissertação de mestrado). Pontifícia Universidade Católica de Campinas, Campinas, SP, Brasil.
- Barbosa, M. A. (1995). Contribuição ao estudo de aspectos da tipologia de obras lexicográficas. *Revista Brasileira de Lingüística*, 8(1), 15-30.
- Bocato, V. R. C. (2011). Os sistemas de organização do conhecimento nas perspectivas atuais das normas internacionais de construção. *Incid: R. Ci. Inf. e Doc.*, 2(1), 165-192. Recuperado de <http://www.revistas.usp.br/incid/article/view/42340>.
- Brasil. Presidência da República. (2013). *Lei nº 12.871, de 22 de outubro de 2013*. Institui o Programa Mais Médicos, altera as Leis nº 8.745, de 9 de dezembro de 1993, e nº 6.932, de 7 de julho de 1981, e dá outras providências. Recuperado de <http://migre.me/jR8FE>.
- Cabré, M. T. (2003). Normalización de la terminología y respeto a la diversidad. In: IULATER. *Terminologia y modelos culturais*. Barcelona: IULA.
- CABRÉ, M. T. (1993). **La terminologia**: teoría, metodología, aplicaciones. Barcelona: Editorial Antártida/Empúries,
- Cintra, A. M. M., Tálamo, M. F. G. M., Lara, M. L. G., & Kobashi, N. Y. (1994). *Para entender as linguagens documentárias*. (2a ed.) São Paulo: Polis.
- Creswell, J. (1998). *Qualitative inquiry and research design: choosing among five traditions*. Thousand Oaks, CA: Sage Publications.
- Cunha, M. (1995). *Parque proletário, grotão e outras moradas: saber e história nas favelas da Penha* (Dissertação de Mestrado). Universidade Federal Fluminense, Niterói, RJ, Brasil.
- Dahlberg, I. (1978). Teoria do conceito. *Ci. Inf.*, Rio de Janeiro, 7(2): 101-107.
- Gardin, J. C. (1966). Elements d' un modele pour la description des lexiques documentaires. *Bulletin des Bibliothèques de France*, 5, 171-182.
- Girão, R. (1967). *Vocabulário popular cearense*. Fortaleza: Imprensa Universitária do Ceará.
- Lancaster, F. W. (1993). *Indexação e resumos: teoria e prática*. Brasília: Briquet de Lemos Livros.

Laraia, R. B. (2001). *Cultura: um conceito antropológico*. 14. ed. Rio de Janeiro: Jorge Zahar.

Martins, J. S. (1989). *Caminhada no chão da noite*. São Paulo: HUCITEC.

Ponte, J. C. C. & Olimar Filho, J. (2001). *Dicionário de medicina popular*. Fortaleza: Livro Técnico.

Rey-Debove, J. & MORAIS, C. B. (1984). Léxico e dicionário. *Alfa*, 28(supl.), 45-69.

Sadala, M. L. A. (2017). *A Fenomenologia como método para investigar a experiência vivida uma perspectiva do pensamento de Husserl e de Merleau Ponty*. Recuperado em 16 março, 2017, de <http://zip.net/btprnY>.

Smit, J. W. & Kobashi, N. Y. (2003). *Como elaborar vocabulário controlado para aplicação em arquivos*. São Paulo: Imprensa Oficial.

Vilela, M. (1994). *Estudos de lexicologia do português*. Coimbra: Livraria Almedina.

Vygotski, L. S. (2001). *A construção do pensamento e da linguagem*. São Paulo: Martins Fontes.



FRENTE E ELITE DE PESQUISA DA ISKO-IBÉRICO, A PARTIR DOS TRABALHOS APRESENTADOS NO PERÍODO DE 2005 A 2015

Ely F. Tannuri de Oliveira¹, Leilah Santiago Bufrem², Bruno Henrique Alves³, Natanael V. Sobral⁴

¹Universidade Estadual Paulista, 0000-0003-3365-3000, etannuri@gmail.com

²Universidade Estadual Paulista, 0000-0002-3620-0632, santiagobufrem@gmail.com

³Universidade Estadual Paulista, 0000-0001-8381-2837, brhenriquealves@gmail.com

⁴Universidade Federal da Bahia, 0000-0003-2410-494X, natanvsobral@gmail.com

RESUMO O estudo questiona como se configuram a Elite de Pesquisa e a Frente de Pesquisa por meio dos trabalhos apresentados ao ISKO-Ibérico, no período de 2005 a 2015, em um corpus de 379 pesquisas. Para destacar a Elite de Pesquisa, ranqueia os pesquisadores mais produtivos, aplicando a Lei do Elitismo de Price. Para determinar a Frente de Pesquisa, realiza a análise de citação, tomando os autores mais citados, analisa e compara os dois conjunto de dados, resultando alguns pesquisadores em comum tanto na Frente de Pesquisa como na Elite de Pesquisa. Como principais resultados, verifica que: Guimarães, J.A.C. (15); Fujita, M.S.L. (12); Almeida, C.C. de (8); Bufrem, L.S. (8); e Agustín-Lacruz, M^a del C. são os autores mais produtivos, evidenciando domínio dos pesquisadores brasileiros e espanhóis. Dentre os autores mais citados, destacam-se: Dalhberg, I. (39); Hjørland, B. (35); Campos, M. L. de A. (17); Izquierdo Alonso, M. (16); e Wittgenstein, L. (16), fazendo parte da lista autores advindos de diversas instituições, tais como, Brasil, Estados Unidos, Espanha, Canadá, Dinamarca, Alemanha, França, Reino Unido, Austrália, Áustria, Bélgica, Índia, Portugal, Rússia, Suíça, Tunísia e Uruguai. Por fim, nota-se que os pesquisadores Guimarães, J.A.C.; Lara, M.L.G.; Moreiro González, J.Á.; e Rodríguez Bravo, B. apareceram tanto na Elite de Pesquisa como na Frente de Pesquisa. Isso indica que esses pesquisadores contribuem, efetivamente, para a consagração de elementos teóricos-metodológicos que fundamentam o domínio da Organização do Conhecimento e propõem procedimentos para compreender o domínio em questão.

PALAVRAS-CHAVE *Organização do Conhecimento, Estudos Métricos, Perspectivas de Investigação, Frente de Pesquisa e Elite de Pesquisa.*

ABSTRACT The study questions how the Research Elite and the Research Front are configured through the papers presented to ISKO-Iberico, from 2005 to 2015, in a corpus of 379 researches. To highlight the Research Elite, the study ranks the most productive researchers by applying Price's Elitism Law. To determine the Research Front, it performs the citation analysis, taking the most cited authors, analyzes and compares the two datasets, resulting in some common researchers in both the Research Front and the Research Elite. As main results, it verifies that: Guimarães, J.A.C. (15); Fujita, M.S.L. (12); Almeida, C.C. de (8); Bufrem, L.S. (8); and Agustín-Lacruz, M^a del C. are the most productive authors, evidencing the domain of Brazilian and Spanish researchers. Among the most cited authors, the following stand out: Dalhberg, I. (39); Hjørland, B. (35); Campos, M. L. de A. (17); Izquierdo Alonso, M. (16); and Wittgenstein, L. (16), as part of the list of authors coming from different institutions such as Brazil, the United States, Spain, Canada, Denmark, Germany, France, the United Kingdom, Australia, Austria, Belgium, India, Portugal, Russia, Switzerland, Tunisia and Uruguay. Finally, we note that the researchers Guimarães, J.A.C.; Lara, M.L.G.; Moreiro González, J.Á.; and Rodríguez Bravo, B. appeared in both the Research Elite and the Research Front. This indicates that these researchers effectively contribute to the assignment of theoretical-methodological elements that

underlie the domain of Knowledge Organization and propose procedures to understand the domain in question.

KEYWORDS *Knowledge Organization, Metric Studies, Research Perspectives, Research Front, Research Elite*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A ISKO - *International Society for Knowledge Organization* - é a sociedade científica da área de Organização do Conhecimento (OC) responsável pela construção, consolidação e disseminação do conhecimento científico, considerando os aspectos que envolvem a origem da Teoria do Conhecimento. Ela foi fundada em 22 de julho de 1989, teve, em sua primeira Diretoria, a pesquisadora Ingetraut Dahlberg, fundadora e Presidente de 1989 a 1996 (Fujita, 2008).

A OC pode ser considerada como um conjunto de elementos sistematizados que contribuem para a materialização, socialização e construção do conhecimento científico, no qual, os fundamentos teóricos, ainda, estão sendo desenvolvidos em um processo contínuo de consolidação teórico-metodológica, considerando o recente processo de institucionalização da OC como domínio científico, bem como a investigação epistemológica dos conceitos (Guimarães, Martínez-Ávila & Alves, 2015).

Para Hjørland (2008) a OC envolve uma compreensão da natureza e origem do conhecimento, cognição, linguagem e organização social, pois esses elementos são determinantes para o entendimento da OC em um contexto completo formado por agentes e/ou instituições.

Considerando que a pesquisa em OC, principal foco dos eventos da ISKO, resulta de um contexto dinâmico de produção e reelaboração de conhecimentos sistematizados, enfatiza-se aqui que este contexto está em dinamismo constante em razão da sua aplicação concreta e da comunicação científica. Ao reorganizar-se constantemente na tentativa de sistematizar conceitos e, desse modo, interferir na realidade multifacetada, a OC contribui para a transformação dinâmica dos conceitos anteriores incorporando novas ideias e construindo o conhecimento que pode contribuir diretamente para o desenvolvimento da sociedade, considerando os elementos na contínua construção da ciência que se faz no desenrolar da história.

Segundo Arboit, Oliveira, Bufrem & Gabriel Junior (2015, p. 2),

tal demanda, acentuada pelas crescentes ramificação e inter-relação dos saberes, requer processos mais complexos para o estudo de seus objetos, pois as atividades de pesquisa são determinadas tanto pelo contexto social e histórico quanto pelos marcos teóricos e metodológicos orientadores de conceitos e pressupostos científicos.

O histórico de trabalhos referentes a pesquisas concretas que vêm sendo realizadas no domínio específico da OC é compreendido como resultado de buscas intencionais por respostas a questionamentos sobre a realidade. As características do tema, especialmente as questões ligadas à interdisciplinaridade da OC, torna nebulosa as fronteiras entre a OC e as demais subáreas da Ciência da Informação (CI), Representação do Conhecimento, Perspectivas Epistemológicas em OC, Tendências e Desafios da OC no novo século, entre outros, tem sido objeto de pesquisas no âmbito da ISKO e não

raro discutível sua sistematização, especialmente a partir dos trabalhos apresentados nas Conferências Internacionais da *International Society for Knowledge Organization* (ISKO) como também naquelas setoriais (Zherebchevsky, 2010, Friedman, 2006, Arboit, Grácio, Oliveira & Bufrem, 2012, Arboit, 2014). Entende-se que esta entidade e os conteúdos produzidos e registrados, como representativos da área em função do papel de centralidade que adquire.

Cabe destacar o grande número de pesquisas metodológicas apresentadas, pela relevância que este aspecto da pesquisa apresenta nos trabalhos em OC (Freitas; Nascimento & Bufrem, 2014), percebeu-se um aprofundamento teórico mais intenso em relação ao objeto estudado, como propostas de criação de procedimentos, modelos, metodologias e estruturas para a OC em contextos diversos, o que representa um indício de amadurecimento da pesquisa, graças ao estudo das relações específicas entre variáveis que envolvem os estudos. Em geral os estudos metodológicos remetem à necessidade de conhecimento sobre tecnologias da informação e comunicação.

Sobre as modalidades de pesquisa e opções concretas de dissertações da área de CI, verificou-se que, quanto às técnicas e instrumentos mais utilizados por autores de estudos de caráter exploratório-metodológico, destacam-se os dados colhidos anteriormente ou registros existentes e, em segundo plano, a entrevista, a utilização de modelo e o questionário (Bufrem, 1996).

Observam-se que nos últimos eventos da ISKO, qualquer que seja o âmbito, uma acentuada tendência para contemplar, entre os subtemas, a avaliação da ciência e “perspectivas de investigação em representação e organização do conhecimento, suas atualidades e tendências estudadas”. Essas linhas se remetem às metodologias quantitativas e qualitativas.

O crescimento da ciência em âmbito mundial e, alinhando-se a este, o crescimento da ciência no Brasil, geraram a necessidade de metodologias e procedimentos pertinentes para avaliar o ascendente incremento da produção científica.

Essas metodologias, tanto quantitativas quanto qualitativas, como as demais que utilizam estas duas abordagens de forma simultânea, passam a ser foco de interesse de pesquisadores, especialmente a partir de 1970, quando a hegemonia dos números, em pesquisa, deixa de existir. Começam a ganhar forças, tanto na França quanto nos Estados Unidos, as críticas feitas à pesquisa quantitativa nos campos da Psicologia e Sociologia (Mueller, 2007). Desenvolvem-se as metodologias qualitativas, tais como método da observação, método histórico, pesquisa-ação, pesquisa exploratória, análise do discurso, análise de redes sociais, entre outras.

Os métodos quantitativos em CI constituem o conjunto de conhecimentos relacionados à avaliação da informação e da ciência produzida, fundamentam-se na Sociologia da Ciência e utilizam-se de quantificações, com procedimentos advindos de outras áreas, como a Matemática e Estatística, bem como Computação (Oliveira & Grácio, 2011).

Os Estudos Métricos, de vertente quantitativa, desenvolveram-se a partir da Bibliometria, tendo na Informetria sua maior amplitude. A maioria dos artigos que adota o termo Estudos Métricos é de autoria de pesquisadores da Ibero-América, especialmente de Cuba, Espanha e Brasil e, com menor frequência, artigos advindos de pesquisadores da Índia, Estados Unidos e Hungria. No Brasil, alguns autores apontam que os termos Bibliometria, Metrias da Informação, Estudos Bibliométricos ou Estudos Métricos são considerados sinônimos, ou mesmo avaliação da ciência. Qualquer que seja a terminologia adotada, os pesquisadores incluem nelas as seguintes subáreas: Bibliometria, Cientometria, Webmetria,

Patentometria, Almetria (a mais recente) e Informetria (a mais ampla) (Grácio & Oliveira, 2015). Destaca-se a importância da internet e do advento da Informática como o “grande divisor de águas” na história dos Estudos Métricos da Informação, inicialmente chamados Bibliométricos. Registra-se um renovado e crescente interesse pelos procedimentos e abordagens em Estudos Métricos, que deixam de ter a natureza predominantemente descritiva ou de mera contagem. As abordagens passam a ser mais legítimas e confiáveis, e a tomada de decisões em Planejamento, Ciência e Tecnologia sustenta-se por esses estudos, mais precisos na medida em que trabalham com números contextualizados, segundo cada ambiência.

Com os avanços tecnológicos, alinham-se e ampliam-se os campos da Bibliometria, não só em relação às subáreas que se institucionalizaram como também em relação ao aprofundamento dos estudos em cada uma delas.

Em estudo anterior, desenvolvido por Pavanelli, Oliveira, Bufrem, Prado & Alves (2015), apresentado ao ISKO-Espanha-Portugal - *International Society for Knowledge Organization*-ISKO-Ibérico-em 2015, procedeu-se à avaliação da produção científica dos 320 trabalhos completos dos anais do ISKO-Ibérico, comunicados no período de 2005 a 2013, com o objetivo de identificar evidências concretas que registram as relações entre os principais conceitos, autores e temáticas e seus referentes hegemônicos. No presente estudo, ao acrescentar os 59 trabalhos apresentados na ISKO-Ibérico de 2015, complementando, portanto, o período anterior, procura-se avaliar, por meio da análise de citação das referências dos trabalhos completos, a Frente de Pesquisa da ISKO-Ibérico, no período entre 2005 a 2015, e explicitar sua Elite de Pesquisa, objetivando mostrar as perspectivas de investigação em Organização do Conhecimento e suas tendências, uma vez que o evento trata de uma área do conhecimento em constante dinâmica de revisão e construção dos seus fundamentos teóricos.

Para este estudo, tomou-se a concepção de Frente de Pesquisa de Braga (1973, p. 12) ao defini-la como o conjunto dos artigos mais citados na literatura recente. Ao afirmar que “os artigos surgidos a cada ano estão estreita e multiplamente relacionados a uma seleta pequena parte da literatura recente e relacionados remota e aleatoriamente a uma parte maior da literatura mais antiga”, a autora considera que “apenas uma pequena parte da literatura recente está inter-relacionada pelo grupo de novos artigos”, sendo uma espécie de “camada epidérmica crescente, formadora de uma ativa Frente de Pesquisa (*Research Front*) que distingue a ciência da erudição (*scholarship*)”.

Quanto à Elite de Pesquisa, para defini-la foi utilizado o critério de Price (1976, p. 30), para quem “o número de produtores prolíficos parece equivaler à raiz quadrada do número total de autores”, ou seja, da quantidade de autores responsáveis pelos artigos constituintes do corpus da pesquisa. Espera-se que a Elite de Pesquisa seja responsável por pelo menos por 50% da produção científica do grupo.

Ainda Mostafa e Máximo (2003), analisaram um conjunto de citações, com o objetivo de mostrar a Frente de Pesquisa e visualizar as tendências epistemológicas nos Estudos de Comunicação. Araújo (2006) estudou o conceito de Frente de Pesquisa, que correlaciona os índices absolutos de citação obtidos por cada autor, com a data dos trabalhos publicados por cada autor. Assim, para a contagem da Frente de Pesquisa, só são contabilizados os trabalhos mais recentes desse autor, podendo ser definida a Frente de Pesquisa para os últimos cinco anos, de modo que quanto maior é a percentagem encontrada, mais ativa seria a Frente de Pesquisa. Jarneving (2005) e Grácio (2016) ao estudar questões de acoplamento e análise de cocitação utiliza os estudos de outros autores, que compara artigos bibliograficamente acoplados com artigos cocitados e conclui que a Frente de Pesquisa é retratada de formas distintas, de acordo com o abordagem empregada.

Tanto a Elite, quanto a Frente de Pesquisa, são questões clássicas da Bibliometria. Encontram-se muitos estudos da década de 1970 sobre os temas. Hoje, com crescimento dos estudos sobre produção científica e o desenvolvimento de novas abordagens sobre citações, esses temas se tornam emergentes.

Vale salientar, conforme Glänzel (2003) e Smiraglia (2011), que as citações apontam os paradigmas das comunidades formadas e seus procedimentos metodológicos, identificam os grupos de cientistas e suas publicações e evidenciam os pesquisadores de maior impacto de uma área. Assim, conhecendo-se as citações, conhece-se o domínio da área e a listagem dos autores mais citados constitui o conjunto da Frente de Pesquisadores. Segundo Urbizagástegui-Alvarado (2010), tanto os autores da Elite de Pesquisa quanto àqueles que constituem a Frente de Pesquisa podem ser identificados pela raiz quadrada da população dos autores mais produtivos (Elite de Pesquisa) e mais citados (Frente de Pesquisa).

Assim, esta pesquisa propõe-se a responder às seguintes questões: Qual é a Frente e Elite de Pesquisa da ISKO-Ibérico, a partir dos trabalhos apresentados no período de 2005 a 2015? O que evidencia o delineamento tanto da Frente e da Elite de Pesquisa? De que forma a resposta a estas questões contribuem para o desenvolvimento da Organização do Conhecimento? Propõe-se, no decorrer da pesquisa responder a estas proposições.

A partir das considerações feitas, os objetivos gerais desta pesquisa são: avaliar a Frente de Pesquisa da ISKO-Ibérico, no período em estudo e explicitar sua Elite de Pesquisa. De forma mais específica propõe-se a determinar, por meio de análise de citação das referências, a Frente de Pesquisa e identificar a Elite de Pesquisa a partir dos autores mais produtivos. Por fim, delinear os grandes temas e as grandes linhas em torno das quais se aglutinam os trabalhos da ISKO-Ibérico, no período em destaque.

As respostas a essas perguntas justificam a relevância da pesquisa, na medida em que elucidarão quais as grandes linhas do pensamento dominante em Organização do Conhecimento, além de outras emergentes e quais seus referentes teóricos a partir especialmente da Frente de Pesquisa, além de mostrar os autores mais prolíficos do período estudado (Urbizagástegui-Alvarado, 2009). Apesar desta pesquisa não ter contribuído para o aprofundamento destes conceitos, utiliza-os como aplicações, e contribui para que os pesquisadores atuais repensem seus objetos de estudo, relativos aos conceitos teóricos-metodológicos presentes na Organização do Conhecimento, considerando a sua imersão no contexto político e social do tema em questão (Glänzel, 2003).

METODOLOGIA

Como procedimento de pesquisa, levantaram-se 379 trabalhos completos dos anais das edições do ISKO-Ibérico, sendo: 43 trabalhos de Barcelona, em 2005; 60 de León, em 2007; 87 de Valência, em 2009; 38 de Ferrol, em 2011 e 92 do Porto, em 2013, e 59 em Múrcia, em 2015, considerado um período expressivo para representar as relações entre os principais conceitos, autores e temáticas.

Os 379 trabalhos foram agrupados em um único conjunto e colocados em ordem decrescente de autoria no sentido de se calcular os pesquisadores mais produtivos e desdobradas as coautorias.

Em um primeiro momento, encontraram-se 545 autores participantes da ISKO-Ibérico e extraiu-se a raiz de 545, que resultou aproximadamente 23 pesquisadores que fizeram parte da Elite de Pesquisa, totalizando, no contexto da pesquisa, 30 pesquisadores que publicaram pelo menos quatro trabalhos.

Em um segundo momento, tomaram-se os 379 trabalhos apresentados no período, coletaram-se todas as citações, fazendo-se as devidas triagens.

Para se ter uma estimativa da frequência de autocitações simulou-se por meio da amostragem sistemática, com N=6 (intervalo) os percentuais de autocitações. Foi verificado um alto percentual de autocitações, aproximando-se a 50% de trabalhos, o que se considerou provocaria vieses aos resultados de Frente de Pesquisa. Assim, desconsideraram-se todas as autocitações. Caso permanecessem, consistiria em uma cultura reiterativa do próprio pesquisador, o que comprometeria os resultados do estudo em questão.

Por meio da análise de citação, identificaram-se 2303 pesquisadores citados, em um total de 2608 citações. Construiu-se uma tabela em ordem decrescente de citações, para identificar a Frente de Pesquisa, a partir da raiz de quadrada de 2303, resultando 48 pesquisadores. Foram feitos todos os desdobramentos possíveis dos autores, no caso de coautorias e dos trabalhos que apresentavam “et al.”. A intenção de incluir no corpus todos aqueles com no mínimo 5 citações resultou na soma de 56 pesquisadores.

Analisaram-se a Elite e a Frente de Pesquisa, apresentando-se suas grandes linhas de trabalho. Compararam-se os resultados com destaque para aqueles pesquisadores pertencentes às duas tabelas. Por fim, considerou-se a relevância dos estudos relativos a Elite e Frente de Pesquisa porque elas apresentam os principais rumos e indicações teórico-metodológico, que sustentam a Organização do Conhecimento. Apesar das possibilidades previstas em outros procedimentos para nuclear os autores mais produtivos e mais citados, a busca da Elite e Frente de Pesquisa constitui-se em um procedimento clássico com prolífica utilização.

RESULTADOS

A Tabela 1 apresenta a Elite de Pesquisa constituída pelos 30 pesquisadores e sua afiliação institucional que produziram pelo menos 4 trabalhos.

A relação dos autores mais produtivos é liderada por Guimarães, J. A. C., com 15 trabalhos, seguido de Fujita, M. S. L., com 12, observando-se que a distribuição tem uma variação entre 4 a 15 trabalhos. Há presença de autores advindos de instituições do Brasil, Espanha, Hungria, Portugal e Alemanha, com evidente destaque do Brasil, seguido da Espanha.

Essa presença predominantemente hispano-brasileira sugere que os pesquisadores desses países têm dado os rumos da produção científica nos temas dos eventos da ISKO-Ibérico, no período em estudo.

Vale destacar a observação de Guimarães e Dodebei (2015), na sua "Introdução aos Estudos Avançados em Organização do Conhecimento", que especificamente em relação ao Brasil, a área de Organização (e representação) do Conhecimento vem sendo, há décadas, objeto de pesquisa na área de Ciência da Informação, como demonstra a trajetória do Encontro Nacional de Pesquisa em Ciência da Informação-

ENANCIB (por meio do GT-2 "Organização e Representação do Conhecimento) e dos programas de pós-graduação na área, cujas principais linhas de pesquisa são dedicadas ao campo.

Tabela 1. Elite de pesquisa

Pesquisadores	Número de trabalhos publicados
Guimarães, J. A. C. (UNESP-Brasil)	15
Fujita, M. S. L. (UNESP-Brasil)	12
Almeida, C. C. de (UNESP-Brasil)	8
Austín-Lacruz, M ^a del C. (UNIV. Zaragoza-Espanha)	8
Bufrem, L. S. (UFPE-Brasil)	8
Hajdu Barat, A. (UNIV. of Szeged-Hungria)	7
Lara, M. L. G. de (USP-Brasil)	6
Moreiro González, J. A. (UC3M-Espanha)	6
Simeão, E. L. M. S. (UnB-Brasil)	6
Garrido-Picazo, P. (UNIV. Zaragoza-Espanha)	5
Lima, G. Á. B. de O. (UFMG-Brasil)	5
Martínez-Ávila, D. (UNESP-Brasil)	5
Moraes, J. B. E. de (UNESP-Brasil)	5
Neves, D. A. de B. (UFPB-Brasil)	5
Orrico, E. G. D. (UFRJ-Brasil)	5
Terra, A. L. (Esc. Sup. de Est. Industriais-Portugal)	5
Varela, A. V. (UFBA-Brasil)	5
Barros, T. H. B. (UFPA-Brasil)	4
Fernández-Molina, J. C. (UNIV. de Granada-Espanha)	4
Kobashi, N. Y. (USP-Brasil)	4
Milani, S. O. (UNESP-Brasil)	4
Moura, M. A. (UFMG-Brasil)	4
Ohly, H. P. (Inst. for the Social Sciences-Alemanha)	4
Oliveira, Ely F. Tannuri de (UNESP-Brasil)	4
Paula, C. P. A. de (UFMG-Brasil)	4
Rodríguez-Bravo, B. (UNIV. de León-Espanha)	4
Rodríguez-Yunta, L. (CSIC-Espanha)	4
Sales, R. de (UFF-Brasil)	4
San Segundo, R. (UC3M-Espanha)	4
Tramullas, J. (UNIV. de Zaragoza-Espanha)	4

Referência: Oliveira, Bufrem, Alves e Sobral (2017)

A partir da Frente de Pesquisa, constituída pelo conjunto dos autores mais citados do corpus, apresenta-se a Tabela 2, composta pelos pesquisadores citados pelo menos 5 vezes no conjunto de trabalhos e referências contabilizadas, identificando-se uma variação entre 5 a 39 citações na distribuição.

Destaca-se que Dahlberg, I. foi citada 39 vezes, seguida por Hjørland, B., com 35 citações, e Campos, M. L. de A. primeira autora brasileira do ranking, que obteve 17 citações. Trabalhos anteriores também destacam a posição de Dahlberg, I. e Hjørland, B., tais como de Araújo, C. A. A., Rolim, E. A., Marzano, I. M. G. & Bitencourt, L. G. (2007) e o de Bufrem, Silva e Sobral (2017) que investigaram os autores mais prestigiados por professores e pesquisadores em Ciência da Informação. Quanto a Izquierdo Alonso, pesquisadora de Espanha, cuja participação tem sido assídua nos eventos ISKO, é reconhecida

pela sua posição sócio cognitiva como um marco propício para o ensino, especialmente na indexação e elaboração de resumos. Já como representante dos autores fundantes, Wittgenstein contribuiu decisivamente para os campos da lógica, da filosofia da linguagem, da filosofia da matemática e da filosofia da mente. Sua obra seminal para a área tem sido o *Tractatus Logico-Philosophicus*, na qual discute as condições lógicas para que o pensamento e a linguagem possam representar o mundo. O segundo filósofo do grupo, Serres, defende o desvio do foco da abstração nos verbos ou substantivos, a partir dos quais o conceito filosófico se enuncia, para um tipo de abstração que se verifique no deslocamento, portanto, a partir da relação, do contato (Serres, 1999). Com a mesma quantidade de citações, cinco pesquisadores da área vêm a seguir, García Gutiérrez, A., Gomes, H. E., Guimarães, J. A. C., Lancaster, F. e López-Huertas, M. Como terceiro filósofo na Frente de Pesquisa, Foucault vem contribuindo com a Organização do Conhecimento (OC), com suas análises sobre as instituições sociais, a cultura, a sexualidade e o poder. Segundo ele, as sociedades modernas e contemporâneas são disciplinares e apresentam uma nova organização do poder, que, por sua vez, foi fragmentado em “micropoderes”, estruturas veladas do poder. Para o filósofo, o poder na atualidade engloba os diversos âmbitos da vida social e não somente o poder concentrado no Estado, teoria esclarecida em sua obra “Microfísica do Poder” Percebe-se que a Frente de Pesquisa é constituída em grande parte por pesquisadores da área de Organização do Conhecimento, com presença significativa de filósofos, evidenciando a relação significativa entre os domínios da linguística, da lógica e os modos de Organizar o Conhecimento.

Tabela 2. Frente de pesquisa

Pesquisadores	Número de citações
Dahlberg, I. (Alemanha)	39
Hjørland, B. (Dinamarca)	35
Campos, M. L. de A. (Brasil)	17
Izquierdo Alonso, M. (Espanha)	16
Wittgenstein, L. (Áustria)	16
Serres, M. (França)	13
García Gutiérrez, A. (Espanha)	12
Gomes, H. E. (Brasil)	12
Guimarães, J. A. C. (Brasil)	12
Lancaster, F. (Estados Unidos)	12
López-Huertas, M. (Espanha)	12
Levy, P. (Tunísia)	11
Bellotto, H. L. (Brasil)	10
Foucault, M. (França)	10
Morin, E. (França)	10
Barité, M. (Uruguai)	9
Otlet, P. (Bélgica)	9
Albrechtsen, H. (Dinamarca)	8
Barreto, A. A. (Brasil)	8
Beghtol, C. (Canadá)	8
Lotman, I. M. (Rússia)	8
Mai, J.-E. (Canadá)	8
Ranganathan, S. R. (Índia)	8
Castells, M. (Espanha)	7
Freire, P. (Brasil)	7
Moreiro González, J. A. (Espanha)	7

Nielsen, J. (Dinamarca)	7
Piaget, J. (Suíça)	7
Ribeiro, F. (Brasil)	7
Svenonius, E. (Estados Unidos)	7
Café, L. (Brasil)	6
Decker, S. (Alemanha)	6
Gruber, T. (Estados Unidos)	6
Hartley, J. (Reino Unido)	6
Herrero Solana, V. (Espanha)	6
Hurley, C. (Canadá)	6
Lara, M. L. G. (Brasil)	6
Pombo, O. (Portugal)	6
Putnam, H. (Estados Unidos)	6
Silva, M. C. P. da (Brasil)	6
Álvares, L. (Brasil)	5
Alvite Díez, M. L. (Espanha)	5
Brascher, M. (Brasil)	5
Cook, T. (Canadá)	5
Dervin, B. (Estados Unidos)	5
Dudziak, E. A. (Brasil)	5
Fisher, K. E. (Estados Unidos)	5
Hudon, M. (Canadá)	5
Ingwersen, P. (Dinamarca)	5
Jenkinson, H. (Reino Unido)	5
Kuhlthau, C. C. (Estados Unidos)	5
Marteleteo, R. M. (Brasil)	5
Rodríguez Bravo, B. (Espanha)	5
Schellenberg, T.R. (Estados Unidos)	5
Soergel, D. (Alemanha)	5
Talamo, M. de F. G. M. (Brasil)	5

Referência: Oliveira, Bufrem, Alves e Sobral (2017)

Há presença de autores advindos de instituições do Brasil, Estados Unidos, Espanha, Canadá, Dinamarca, Alemanha, França, Reino Unido, Austrália, Áustria, Bélgica, Índia, Portugal, Rússia, Suíça, Tunísia e Uruguai.

Por fim, foram comparados os resultados, com destaque para aqueles pesquisadores pertencentes às tabelas 1 e 2 e identificou-se que os pesquisadores Guimarães, J. A. C.; Lara, M. L. G.; Moreira González, J. A. e Rodríguez Bravo, B. apareceram tanto na Elite de Pesquisa como na Frente de pesquisa. Isso indica que esses pesquisadores contribuem, efetivamente, para a consignação de elementos teóricos-metodológicos que fundamentam o domínio da Organização do Conhecimento e propõem procedimentos para compreender o domínio em questão.

CONCLUSÕES

Com a intenção de caracterizar, por meio da análise de citação das referências em trabalhos completos apresentados no período de 2005 a 2015 ao ISKO-Ibérico, foi possível identificar 2303 pesquisadores citados, em um total de 2608 referências, cuja Elite de Pesquisa se constitui de um grupo de autores de diversas nacionalidades e instituições. Destaca-se, como primeiro autor da Elite de Pesquisa,

Guimarães, J. A. C., com um total de 15 trabalhos, seguido de Fujita, M. S. L., com 12, observando-se que a distribuição tem uma variação entre 4 a 15 trabalhos. Há presença de autores advindos de instituições do Brasil, Espanha, Hungria, Portugal e Alemanha, com destacada presença do Brasil, seguida da Espanha. Essa presença predominantemente hispano-brasileira sugere que os pesquisadores desses países têm dado os rumos da produção científica nos temas dos eventos da ISKO-Ibérico, no período em estudo e isso reflete uma condição histórica que expressa as relações constantes entre pesquisadores dos dois países.

Sob o ponto de vista metodológico, confirmou-se que a busca da Elite e Frente de Pesquisa é um procedimento clássico com farta utilização em estudos sobre a literatura em Organização do Conhecimento.

REFERÊNCIAS BIBLIOGRÁFICAS

Arboit, A. E., Grácio, M. C. C., Oliveira, E. T. & Bufrem, L. S. (2012). Relationship between authors and main subject categories in the Knowledge Organization domain: a bibliometric approach. In: A. Neelameghan, K. S. Raghavan (Eds.). *Categories, contexts and relations in Knowledge Organization: proceedings of the twelfth international ISKO conference* (pp.44-50). Würzburg: Ergon Verlag.

Arboit, A. E. (2014) *O processo de institucionalização sociocognitiva do domínio de Organização do Conhecimento a partir dos trabalhos científicos dos congressos da ISKO*. Tese de doutorado, Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília.

Arboit, A. E., Oliveira, E. F. T., Bufrem, L. S. & Faustino, R. G. (2015). Noções de modelo científico no domínio de organização do conhecimento: um estudo a partir dos trabalhos apresentados nos congressos internacionais da ISKO. In: *XII Congreso ISKO España y II Congreso ISKO Espapa-Portugal* (pp.1-6). Murcia: Universidad de Murcia.

Araújo, C. A. (2006). Bibliometria: evolução histórica e questões atuais. *Em Questão*, 12 (1) 11-32.

Araújo, C. A. A., Rolim, E. A., Marzano, I. M. G. & Bitencourt, L. G. (2007). A ciência da informação na visão dos professores e pesquisadores brasileiros. *Informação & Sociedade, João Pessoa*, 17 (2), 95-108.

Braga, G. M. (1973). Relações bibliométricas entre a frente de pesquisa (Research Front) e revisões da literatura: estudo aplicado a Ciência da Informação. *Ciência da Informação*, 2 (1), 9-26.

Bufrem, L. S., Silva, F. M. & Sobral, N. V. (2017). Análise das influências intelectuais na produção científica da área de Ciência da Informação: um estudo sobre os bolsistas de produtividade em pesquisa (PQ-CNPq). *Em Questão*, 23 (Edição Especial 5 EBBC), 115-141.

Bufrem, L. S. (1996). *Linhas e tendências metodológicas na produção acadêmica discente do mestrado em ciência da informação do IBICT/UFRJ*. Curitiba. Tese para Concurso de Professor Titular, Universidade Federal do Paraná, Escola de Biblioteconomia.

Freitas, J. L., Nascimento, B. S. & Bufrem, L. S. (2014). A Organização do Conhecimento na dinâmica da pesquisa em artigos da literatura científica da Brapci. *Transinformação*, 26 (3), 295-303.

- Friedman, A. (2006). Concept mapping a measurable sign. In: G. Budin, C. Swertz, K. Mitgutsch (Eds.). *Proceedings of the Ninth International ISKO Conference* (pp.131-139): Würzburg: Ergon Verlag.
- Fujita, M. S. L. (2008). Organização e Representação do Conhecimento no Brasil: análise de aspectos conceituais e da produção científica do ENANCIB no período de 2005 a 2007. *Tendências da Pesquisa Brasileira em Ciência da Informação*. 1 (1), 1-32.
- Glänzel, W. (2003). *Bibliometrics as a research field: a course on theory and application of bibliometric indicators*. Bélgica.
- Grácio, M. C. C. (2016). Acoplamento bibliográfico e análise de cocitação: revisão teórico-conceitual. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, 21(47), 82-99.
- Grácio, M. C. C. & Oliveira, E. F. T. de. (2015). Indicadores de proximidades em Análise de Cocitação de Autores: um estudo comparativo entre Coeficiente de Correlação de Pearson e Cosseno de Salton. *Informação & Sociedade: Estudos*, 21 (2), 105-116.
- Guimarães, J. A. C. & Dodebei, V. L. D. L. de M. (2015). *Organização do conhecimento e diversidade cultural*. Marília: ISKO-Brasil-FUNDEPE.
- Guimarães, J. A. C., Martínez-Ávila, D. & Alves, B. H. (2015). Epistemic communities in knowledge organization: an analysis of research trends in the Knowledge Organization Journal. *Paper presented at the meeting of the International Society for Knowledge Organization - Chapter United Kingdom*, London, UK, 13-14.
- Jarneving, B. (2005). A comparison of two bibliometric methods for mapping of the research front. *Scientometrics*, 65 (2), 245-263.
- Hjørland, B. (2008). What is knowledge organization (KO)? *Knowledge organization*, 35(2/3), 86-101.
- Mostafa, S. P. & Máximo, L. F. (2003). A produção científica da Anped e da Intercom no GT da Educação e Comunicação. *Ciência da Informação*, 32(1), 96-101.
- Mueller, S. P. M. (2007). *Métodos pra a pesquisa em Ciência da Informação*. Brasília: Thesaurus, 2007.
- Oliveira, E. F. T. & Grácio, M. C. C. (2011). Indicadores bibliométricos em Ciência da Informação: análise dos pesquisadores mais produtivos no tema “estudos métricos” na base Scopus. *Perspectivas em Ciência da Informação*, 16 (4) 16-28.
- Pavanelli, M. A., Oliveira, E. F. T. de, Bufrem, L. S., Prado, M. A. R. do & Alves, B. H. (2015). Referentes teóricos basilares na ISKO-Ibérico e suas contribuições para a Organização do Conhecimento (OC). In: J. V. R. Muñoz, I. Gil Leiva, P. M. Díaz Ortuño, F. J. Martínez Méndez (Eds.) - XII Congreso ISKO España y II Congreso ISKO España- Portugal. *Organización del conocimiento para sistemas de información abiertos* (pp.570-578). Murcia: Universidad de Murcia
- Price, D. de S. (1976). *O desenvolvimento da ciência*. Rio de Janeiro: Livros Técnicos e Científicos.
- Serres, M. (1999) *Filosofia mestiça*. Rio de Janeiro: Nova Fronteira.
- Smiraglia, R. P. (2011). Isko 11 'diverse book shielf: an editorial. *Knowledge Organization*, 38(3), 179-189.

Urbizagástegui-Alvarado, R. (2009). A frente de pesquisa na literatura sobre a produtividade dos autores. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, 14 (28), 38-56.

Urbizagástegui-Alvarado, R. (2010). A cientometria como um campo científica. *Informação & Sociedade: Estudos*, 20 (3), 41-62.

Zherebchevsky, S. (2010). Formalism in knowledge organization: thematic analysis of ISKO 10 proceedings. In: C. Gnoli, F. Mazzocchi (Eds.). *Paradigms and conceptual systems in knowledge organization* (pp. 98-105). Würzburg: Ergon Verlag.



PERSPETIVAS DE INVESTIGAÇÃO EM REPRESENTAÇÃO E ORGANIZAÇÃO DO CONHECIMENTO: ATUALIDADE E TENDÊNCIAS

ANÁLISE SOBRE AS ABORDAGENS E APLICAÇÕES DO SKOS NO ÂMBITO DA LITERATURA PERIÓDICA INTERNACIONAL DA CIÊNCIA DA INFORMAÇÃO

José Carlos Francisco dos Santos¹, Walter Moreira²

¹Doutorando Unesp – Marília, 0000-0003-0432-665X, jose.santos@marilia.unesp.br

²Professor Doutor Unesp – Marília, 0000-0001-9454-441X, walter.moreira@marilia.unesp.br

RESUMO Analisa as bases *Information Science and Technology Abstracts (ISTA)*, *Library Information Science Abstracts (LISA)*, *Library, Information Science & Technology Abstracts (LISTA)* e *Scopus* para subsidiar a análise e categorização das abordagens do *Simple Knowledge Organization System (SKOS)* na literatura periódica internacional da Ciência da Informação. A motivação do trabalho foi identificar e analisar os registros dos artigos científicos sobre SKOS, considerando-se a importância deste tema para a interlocução entre os diversos sistemas de organização do conhecimento. A pesquisa utiliza abordagem exploratória e descritiva, com abordagem quantitativa. As atividades foram organizadas em quatro etapas, a saber: coleta, tratamento e análise dos dados e aferição dos resultados de busca. Observou-se que as análises de abordagem e aplicação do SKOS ainda são timidamente abordadas na literatura, com incidência relativamente baixa, o que pode representar um déficit de estudo, principalmente quando se leva em consideração a cobertura internacional da pesquisa. Tem-se também em conta que o permanente estudo de percurso dos artigos científicos, é uma inesgotável temática, tanto na Ciência da Informação quanto na Organização do Conhecimento. Entende-se, portanto, a existência de demanda para novas propostas de pesquisa, no sentido de colaborar com a análise do silêncio dos estudos em SKOS.

PALAVRAS-CHAVE *Simple Knowledge Organization System, SKOS.*

ABSTRACT This paper analyzes the following data sources: *Information Science and Technology Abstracts (ISTA)*, *Library Information Science Abstracts (LISA)*, *Library, Information Science & Technology Abstracts (LISTA)* and *Scopus*. To subsidize the analysis and classify *Simple Knowledge Organization System (SKOS)* approaches, grounded on international periodical publication in *Information Science*. The motivation of this paper was to identify and analyze scientific papers registers about SKOS; considering the importance of the theme for interlocution amid diverse knowledge organization systems. This research utilizes descriptive and exploratory, with quantitative approach. The activities were organized in four stages, namely collection, treatment, data analysis, and measurement of search results. It was observed that the analysis approach and applicability of SKOS are still timidly addressed on the literature, with a relatively low incidence, which may represent a study deficit; especially considering the international coverage of the research. It is also believed that the permanent study of scientific papers pathway is an inexhaustible thematic not only in *Information Science*, but also in *Knowledge Organization*. Consequently, it is understood that there is a demand for new research proposals, in order to collaborate to the analysis of the silence on SKOS studies.

KEYWORDS *Simple Knowledge Organization System, SKOS.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A problemática discutida neste trabalho fundamenta-se em Moreira (2012), em trabalho que questiona o silêncio em relação à consolidação teórica e metodológica da proposta do *Simple Knowledge Organization System* (SKOS), notadamente na Ciência da Informação brasileira. No levantamento bibliográfico realizado nesse estudo (Moreira, 2012), cobrindo a literatura internacional no período 2006 a 2011, foram recuperados 16 trabalhos. Visando a aprofundar a compreensão da questão, estrutura-se a problemática desta pesquisa em torno do comportamento das pesquisas relacionadas ao SKOS em bases de dados que cobrem a literatura especializada da ciência da informação.

Tem-se por objetivo geral identificar e analisar os registros dos artigos científicos sobre SKOS, considerando-se sua abordagem na literatura científica periódica internacional, representada pelas bases ISTA, LISA, LISTA e *Scopus*. Os objetivos específicos são: a) identificar, na literatura os trabalhos relacionados ao SKOS, considerando-se sua presença nos campos de título e/ou resumo e/ou palavra-chave; b) estratificar os resultados em planilhas como suporte para a análise; c) calcular o índice de pontos conforme o local de aparecimento do termo a partir das constatações de localização dos termos; d) analisar e agrupar os resumos em categorias e definidas *a priori*.

A pesquisa caracteriza-se como exploratória e descritiva, em abordagem quantitativa, organizada em quatro etapas executadas para a consecução dos resultados, são elas: coleta e organização do *corpus*, organização dos registros em uma planilha, cálculo dos índices de pontuação e análise dos resumos dos registros dos artigos recuperados das bases. O *corpus* é definido pelos trabalhos relacionados ao SKOS em diversas áreas de aplicação, os quais foram alocados em categorias de análise. Considerando-se o critério da temporalidade, observam-se muitas lacunas em diversos anos. Os trabalhos iniciaram de modo tênue, porém registraram um crescimento relativo, com algumas exceções de declínio. Cabe salientar que os artigos são, na maioria, relacionados à aplicação do SKOS em vocabulários controlados, considerando-se o sentido *lato* deste termo.

Há uma notável semelhança no modo como uma pessoa costuma vivenciar a organização do conhecimento, como uma necessidade, quando é criança e a concepção que passa a desenvolver quando adulta. Em diversos jogos e brincadeiras, podem-se encontrar abordagens que remetem à identificação de padrões de classificação, requerendo a observação e identificação de similaridades e diferenças entre objetos. Aliás, se não forem todos, uma parcela bastante significativa dos sistemas de organização do conhecimento (SOC) conhecidos são, em algum nível, devedores dos conhecimentos filosóficos. Aristóteles, em seu esforço de categorizar o conhecimento, representa o primeiro esforço ocidental de sistematizar a organização do conhecimento. Desse modo, relacionam-se, *e.g.*, a *Dewey Decimal Classification* (DDC) e a *Library of Congress Classification* (LCC) com a concepção filosófica de Bacon acerca da organização do conhecimento.

A definição apresentada por Hodge (2000) para a expressão “sistemas de organização do conhecimento” envolve todos os tipos de esquemas utilizados para a organização da informação e do conhecimento. Inclui esquemas mais tradicionais como os tesouros e os sistemas de classificação e outros, ainda relativamente pouco conhecidos do público leigo, como as ontologias e os arquivos de autoridade utilizados para controlar nomes geográficos e nomes pessoais.

O conceito de SOC tem se tornado comum nas discussões sobre organização e representação do conhecimento e é usado como termo genérico para os conceitos de sistemas de classificação, tesouros,

taxonomias e ontologias, dentre outros. Algumas reflexões muito oportunas sobre os SOC's são propostas por Smiraglia (2014): é possível identificar uma base comum aos tipos de SOC's? Exemplificando, um sistema de classificação pode ser transformado em um tesouro? Um tesouro pode ser transformado em uma ontologia? A resposta a essas questões é de interesse tanto prático como teórico. O interesse teórico advém da necessidade de se encontrar uma resposta precisa acerca das características que diferenciam e as que são complementares entre os diferentes tipos de SOC.

A construção, a manutenção e a avaliação de SOC's, em seus aspectos teóricos e metodológicos, compõem o conjunto das atividades nucleares da ciência da informação. Os elementos componentes de um SOC são, naturalmente, os conceitos, os quais são carregados de teoria, mesmo que algumas vezes seja difícil identificá-las. As teorias fornecem a base para a especificação do sentido que o conceito irá expressar (Hjørland, 2015). Dentre os recursos utilizados para expressão dos conceitos nos SOC's, considerando-se o aspecto de representação computacional, está o SKOS, descrito em *Resource Description Framework (RDF)*.

Pastor-Sánchez & Martínez-Méndez (2010), em tradução para o idioma espanhol do documento “SKOS Simple Knowledge Organization System” da W3C (2009), definem SKOS como um vocabulário em RDF para representação de SOC's, tais como tesouros, taxonomias, sistemas de classificação e listas de cabeçalhos de assunto. A representação do SKOS em RDF é decorrente da necessidade de ser legível por máquinas e proporcionar a interoperabilidade.

Ramalho (2015) apresenta o SKOS como uma aplicação do RDF que traduz a representação de conceitos na Web com possibilidade de ser integrados com outros esquemas conceituais. Esta representação se dá por meio da definição dos conceitos de SKOS em um recurso RDF. Desta forma, o SKOS é fundamentado em vocabulários controlados, na perspectiva de representar os conceitos por meio de rótulos, possibilitando a integração com outros esquemas de representação e padrões de metadados. Catarino, Cervantes & Andrade (2015) complementam que o RDF “é o modelo padrão para intercâmbio de dados na Web recomendado pelo W3C e, portanto, sendo o SKOS escrito em RDF, tornam os esquemas de conceito expressos neste modelo, passíveis de serem lidos por agentes inteligentes conforme as recomendações da W3C” (p. 110).

METODOLOGIA

A pesquisa é caracterizada como exploratória e descritiva com abordagem quantitativa. O *corpus* de análise é composto pelos registros recuperados a partir das representações temáticas por meio do título, resumo e da palavras-chave atribuídas aos artigos científicos eletrônicos indexados nas bases de dados: *Information Science and Technology Abstracts (ISTA)*, *Library Information Science Abstracts (LISA)*, *Library, Information Science & Technology Abstracts (LISTA)* e *Scopus*. Não foi estabelecido nenhum período de cobertura cronológica para as buscas e foram localizados, no total, 232 registros.

A primeira etapa desenvolvida foi determinar as estratégias de busca. Para cumpri-la, respeitaram-se os diferenciais de cada base de dados, na tentativa de padronizar as estratégias. Os dois termos pesquisados foram “SKOS” e “Simple Knowledge Organization System” utilizando o operador “or”. Nas bases ISTA, LISTA e LISA, optou-se por manter a pesquisa em qualquer parte do texto, por dois motivos: a) o retorno de registros foi baixo; b) não existe a opção de seleção por título, resumo e palavras-chave no mesmo critério de busca. Na base Scopus foram filtrados os resultados com a apresentação dos termos

de pesquisa apenas nos campos “título”, “resumo” e “palavras-chave” dos artigos. A aplicação dos filtros foi necessária para garantir a precisão nas buscas. Em relação ao tipo de fonte, a única base em que foi possível selecionar "periódicos acadêmicos e revistas" foi a LISA. Quanto ao tipo de documento, na base LISA foi selecionada a opção "Artigo de periódico", já na ISTA e LISTA "Revistas Acadêmicas", na Scopus "Artigo". O idioma, nas bases LISA, LISTA e ISTA foram selecionados todos, na Scopus não apresentava esta seleção. O mesmo ocorre com o critério "revisados por pares", nas bases LISA, LISTA e ISTA foram selecionados, já na Scopus não existia a opção.

Na segunda etapa, os dados coletados foram registrados e organizados em planilha a partir da exportação dos resultados das bases de dados por meio de arquivos textos e/ou html. Nesta extração dos resultados trabalhou-se com os elementos essenciais: títulos, resumos e palavras-chave, e os elementos complementares: ano, idioma e termos indexados.

Na terceira etapa, realizou-se a análise dos resultados com atribuição de pontuações conforme os campos onde foram encontrados os termos de pesquisa. As pontuações obedecem aos seguintes critérios: termos encontrados no resumo dos artigos 1 ponto; termos encontrados no título 2 pontos e termos encontrados nas palavras-chave dos artigos 3 pontos. Ao fim da contagem dos pontos, foi realizada uma soma geral e verificada a média aritmética, a partir da qual foi definida uma linha de corte para a composição do *corpus* definitivo, utilizado nas demais etapas da pesquisa. Deste modo, compuseram o *corpus* os resumos dos artigos que obtiveram pontuação acima da média que é igual a 2,5, calculada a partir da incidência do termo “SKOS”. Para o termo “*Simple Knowledge Organization System*”, a média foi 1,5, portanto, considerou-se como não significativa sua utilização.

Na quarta etapa, em com vistas à aferição de resultados, foi realizada a categorização e a análise dos resumos componentes do *corpus* (n=71). As categorias de análise foram definidas *a priori*, organizadas conforme o tratamento dado ao SKOS: abordagens relativas a aplicações de SKOS (metadados, ontologias, vocabulários controlados), abordagens conceituais, abordagens comparativas, abordagens relativas à edição de SKOS.

Um alerta que se faz ao leitor quando examina as análises apresentadas é sobre pontos a corrigir em futuros retornos ao estudo: 1) acaba ocorrendo um predomínio na identificação dos textos nas bases de dados pesquisadas; 2) torna-se necessário buscar formas de efetuar maior discussão para a análise dos textos ou de partes deles; 3) será necessário expandir a parte analítica para além da identificação da abordagem que cada texto dá ao tema SKOS.

Os autores reconhecem que analisar a literatura internacional sobre o assunto é uma conduta que implica numa complexidade muito maior. Por ora, e vendo este como um primeiro passo, a abordagem quantitativa se sobrepôs à abordagem qualitativa e não deixa dúvidas quanto a um resgate dos dados para um mais amplo tratamento no campo dos sistemas de organização do conhecimento.

RESULTADOS

A quantidade de registros dos artigos recuperados atendeu às expectativas e possibilitou o cumprimento dos objetivos desta pesquisa. Em termos quantitativos, os resultados foram os seguintes: LISA = 34 registros; ISTA = 32 registros; LISTA = 59 registros; *Scopus* = 107 registros. No total, foram recuperados 232 registros, destes, 90 registros foram recuperados em somente uma base de dados, 14 registros encontrados em duas bases, 18 registros encontrados em três bases e oito registros encontrados

em quatro bases. Ainda somam-se dois registros recuperados em duplicidade na base LISTA, um registro duplicado na base LISTA e que também constou em mais uma base, quatro registros duplicados na base LISTA e que constou em mais duas bases e um registro duplicado na base LISTA, que constou em mais três bases.

Em relação aos idiomas dos artigos recuperados, registra-se grande diversidade: catalão, chinês, croata, francês, alemão, italiano, japonês e português, inglês e espanhol. Estes dois últimos com maior incidência: inglês = 119 registros e o espanhol = 24 registros. O idioma do artigo não caracteriza a abrangência da pesquisa, mas permite observar as ramificações das tendências de pesquisas em SKOS. Outro dado importante é o baixo índice de ausência das palavras-chave nos registros (n = 16), equivalente a 6,95% do total de registros.

O índice foi calculado a partir da pontuação aplicada à localização dos termos “SKOS” e “*Simple Knowledge Organization System*”. Esta localização se deu por meio de busca nos campos: título, resumo e palavras-chave, ao localizar o termo atribuíram-se os respectivos valores: 1, quando a palavra-chave aparecia no resumo, 2, quando aparecia no título e 3, quando ocorria no campo de palavras-chave. A média dos índices para o termo “*Simple Knowledge Organization System*” foi 1,5 e do termo “SKOS” 2,5, portanto, desconsiderou-se o termo “*Simple Knowledge Organization System*”. O valor de corte foi a partir da média dos índices, ou seja, dos 232 registros foram filtrados os índices com valor maior que 2,5. Desconsiderando-se os registros duplicados (tanto em bases diferentes quanto na mesma base), o *corpus* de análise foi composto por 71 artigos. Houve casos de duas ocorrências de índices diferentes para o registro do mesmo artigo, porém, em bases diferentes, o que não interferiu na análise.

Na Tabela 1 apresenta-se a frequência relativa aos índices da pontuação. Observa-se que existe uma quantidade concentrada na frequência 0 e 1, com o termo SKOS 44,4% e com o termo “*Simple Knowledge Organization System*” 93,1%. Com base na Tabela 1 é possível observar que o termo mais utilizado nas pesquisas é “SKOS”, primeiramente no resumo (índice 1), depois nas palavras-chave (índice 3) e finalmente nos resumos e nas palavras-chave (índice 4). O Termo “*Simple Knowledge Organization System*” apresenta-se com mais frequência nos resumos (índice 1).

Tabela 1. Frequência dos resultados do índice

ÍNDICE	FREQUÊNCIA DOS TERMOS	
	SKOS	Simple Knowledge Organization System
0	17	154
1	86	62
2	4	1
3	59	10
4	42	4
5	3	0
6	21	1
Total		232

Fonte: Elaborado pelos Autores

Em relação às palavras-chave atribuídas aos artigos, constante nos registros recuperados em especificamente em inglês, observa-se as ocorrências de variação do termo SKOS: por extenso, por extenso e abreviado, *SKOS extension*, *SKOS-Core*, *SKOSXL*. Demais termos atribuídos e com grande ocorrência são: *Semantic Web* (15), *Linked data* (10), *Controlled vocabularies* (9), *Knowledge Organization Systems* (9), *Ontology* (8) – inclui-se 7 ocorrências de *Ontology* (*ontology building*, *Ontology Design Pattern*, *Ontology Development*, *Ontology mapping*, *Ontology matching*), *Linked Open Data* (7), *Classification* (5), *Metadata* (5). Por outro lado, ocorre 11 vezes o termo *article*, constou em alguns registros somente este termo, o que sugere alguma inconsistência nestes registros.

A análise dos resumos dos artigos foi realizada a partir da leitura e a identificação do foco do artigo em relação ao SKOS. Na Tabela 2 são apresentados os autores e as categorias nas quais eles foram agrupados.

A categoria “aplicação geral”, foi instanciada para agrupar os artigos que apresentavam características de aplicação do SKOS em diferentes áreas. Para efeitos de exemplificação, cita-se Wada (2016), que aborda a representação em SKOS relacionada à síntese em um esquema de classificação de hierarquia baseado na interpretação da *Structure-expression Style Theory*. Outro exemplo é Palmonari (2011), que trabalha com uma linguagem que visa alcançar um bom equilíbrio entre expressividade e facilidade de uso, com emprego de um subconjunto binário que possui sintaxe conceitual, sintaxe concreta e semântica formuladas com base numa extensão do SKOS.

A categoria “aplicação - metadados” foi planejada para agrupar os artigos que visam o SKOS como instrumento para descrição de metadados em recursos da *web*. Por seu turno, a categoria “aplicação - ontologias” agrupa os artigos que trabalham com a representação de ontologias de base em formato SKOS e outras relações de SKOS com as ontologias. A categoria “aplicação – vocabulários controlados” aloca os artigos que estão diretamente voltados para os estudos de vocabulários controlados e suas representações em SKOS. A categoria “comparações” relaciona-se com os artigos que trabalham com comparativos do SKOS com outras tecnologias para representação de sistemas de organização de conhecimento. A categoria “conceitual” associa os artigos cujos objetivos estejam relacionados a discussões teórico-conceituais relativas ao SKOS. A categoria “editor” uniu os artigos relacionados a desenvolvimento de *software* ou aplicação de *software* utilizados exclusivamente para editar e gerenciar representações em SKOS.

Dos artigos que compuseram o *corpus* de análise, houve nove que não se referiam, conforme análise dos campos de título, resumos e palavras-chave, ao SKOS.

A Tabela 2 revela a predominância dos artigos que estão relacionados à categoria de “aplicação - vocabulários controlados”, a abrangência dos vocabulários controlados se concentra mais em tesouros. Por outro lado, as categorias de comparação entre outras tecnologias e a “aplicação – ontologias” versam nas duas categorias uma quantidade equilibrada, porém dispersa no tempo, e não mantém uma discussão periódica. Portanto, cabe observar, que esta expressividade da categoria “aplicação - vocabulários controlados” vai ao encontro com a finalidade do SKOS de ser um vocabulário em RDF para representar SOCs. Em relação à quantidade de autores por artigo, em média é de 2,7, o maior índice são artigos com 2 autores (20), seguido pelos individuais (14), 3 autores (11), 4 autores (7), 5 autores (5), 6 autores (2) e 7 autores (3). A produção de cada autor não é regular, ou seja, não tem uma frequência de estudos com a temática em questão, como pode ser observado, 146 autores possuem somente 1 artigo, 10 autores 2, 1 autor 3 e somente 1 autor (Juan Antonio Pastor Sánchez) que mais publica tem 4 artigos. Este último autor, em seu perfil no Google Acadêmico, dos 77 documentos

listados em sua produção, destes 9 constam no título o termo SKOS. O primeiro trabalho foi em 2009 com o título: *Diseño de un sistema colaborativo para la creación y gestión de tesauros en Internet basado en SKOS* – Universidad de Murcia (Tese Doutorado).

Tabela 2. Categorização dos artigos

AUTORES	CATEGORIAS
Gaona-García, Stoitsis, Sánchez-Alonso & Biniari (2016)	aplicação geral (9)
Kunkel (2015)	
Lacasta, Nogueras-Iso, Lopez-Pellicer, Muro-Medrano & Zarazaga-Soria (2007)	
Masamichi WADA (2016)	
Palmonari (2011)	
Putkey (2011)	
Rodríguez-Enríquez, Alor-Hernández, Mejía-Miranda, Sánchez-Cervantes, Rodríguez-Mazahua & Sánchez-Ramírez (2016)	
Rodríguez-Enríquez, Alor-Hernández, Sánchez-Ramírez & Córtes-Robles (2015)	
Shiri (2013)	
Ford (2013)	aplicação - metadados (6)
Solodovnik (2012)	
Méndez & Greenberg (2012)	
Perožić (2014)	
Vlachidis, Binding, May & Tudhope (2013)	aplicação - ontologias (6)
Wright, Harrison & Watkins (2015)	
Baker, Bechhofer, Isaac, Miles, Schreiber & Summers (2013)	
Cox & Richard (2015)	
Della Rocca, Senatore, & Loia (2017)	
Jianwu, Miao & Tuo (2011)	
Miranda, Orciuoli & Sampson (2016)	
Pastor-Sánchez (2013)	
Albertoni, De Martino, Di Franco, De Santis & Plini (2014)	
Amarger, Chanet, Haemmerlé, Hernandez & Roussey (2014)	
Ameri, Kulvatunyou, Ivezic & Kaikhah (2014)	aplicação – vocabulários controlados (36)
Balkan & Bell (2014)	
Binding & Tudhope (2016)	
Çağdaş & Stubkjær (2015)	
Caracciolo et al. (2012)	
Cohen & Franke (2015)	
Firmino & Baptista (2013)	
Fugazza & Vaccari (2011)	
Gray, Gray, Hall & Ounis (2010)	
Hubain, De Wilde & Van Hooland (2016)	
Ji, Piasecki & Lovell (2015)	
Jia & Wei (2012)	
Keil (2012)	
Kempf & Zapilko (2013)	
Ma Luisa Alvite, Pérez-León, Martínez-González, & Dámaso-Javier Vicente (2010)	
Ma, Carranza, Wu, Van Der Meer & Liu (2011)	
Martínez-González & Alvite Díez (2014)	
Méndez & Greenberg (2012)	
Miles & Pérez-Agüera (2007)	
Nicholson & McCulloch (2006)	
O'DELL (2015)	
Papadakis & Kyprianos (2011)	
Pastor Sánchez (2013)	
Pastor-Sánchez (2013)	
Pastor-Sánchez, Martínez-Méndez & Rodríguez-Muñoz (2012a)	
Pastor-Sánchez, Martínez-Méndez & Rodríguez-Muñoz (2012b)	
Perožić (2014)	
Sánchez-Alonso & García-Barriocanal (2006)	
Singthongchai, Niwattanakul & Chamnongsri (2016)	
Suominen & Mader (2014)	
Tennis & Sutton (2008)	
Vlachidis, Binding, May & Tudhope (2013)	comparações (6)
Wright, Harrison & Watkins (2015)	
Zapilko, Schaible, Mayr & Mathiak (2013)	
Bandyopadhyay & Mukhopadhyay (2015)	
Fedotov et al. (2016)	
Martínez-González & Muñoz Nieto (2012)	
McCulloch & MacGregor (2008)	
Pastor, Martínez & Rodríguez (2009)	
Zeng, Hlava, Busch, Buchel & Žumer (2015)	
Catarino, Nogueira & Andrade (2015)	
Cantara (2006)	editor (2)
Conway et al. (2016)	

Fonte: Elaborado pelos Autores

Na Figura 1 apresenta-se o gráfico pautado nas categorias e temporalidade, observa-se que o volume maior de artigos se concentra de 2012 a 2015 na categoria aplicação – vocabulário controlado, a predominância no gráfico é a falta de publicações.

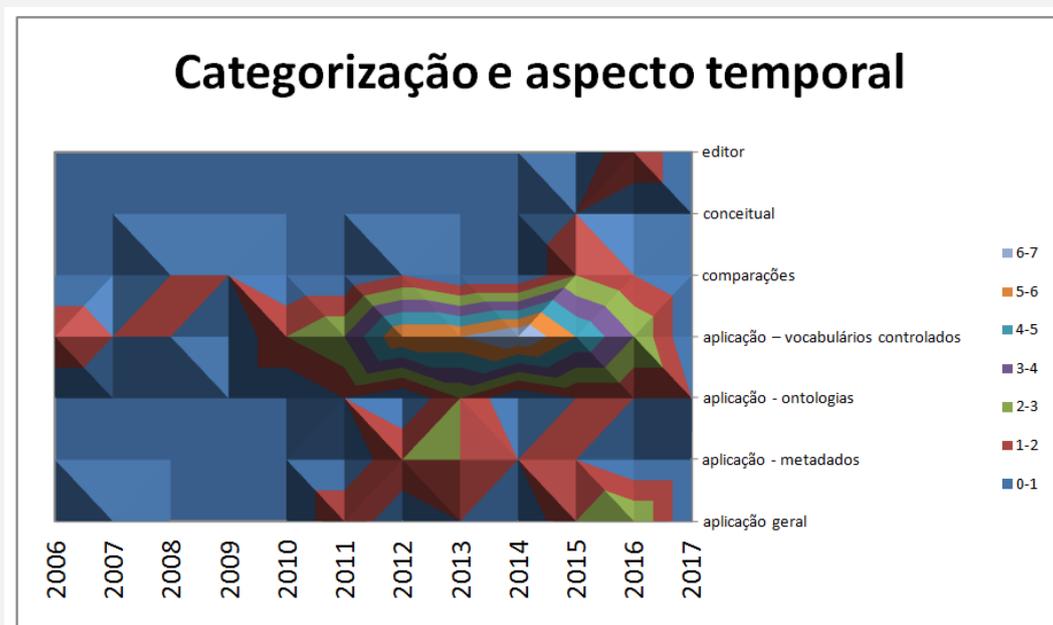


Figura 1. Gráfico de área com a categorização e aspecto temporal

Fonte: Elaborado pelos Autores

No aspecto de distribuição temporal, observa-se um grande número de artigos em 2015 ($n = 14$). Em 2016 ($n = 11$), aconteceu uma queda que foi observada de 2006 (3 artigos) para 2007 (2 artigos) e 2008 (2 artigos) para 2009 (1 artigo), nos demais anos foram crescente o número de publicações. Portanto, o fato em evidência é a forma crescente desde os primeiros artigos, porém em três momentos tiveram quedas. Neste contexto, nota-se também o silêncio das pesquisas relacionadas às ontologias.

CONCLUSÕES

Esperava-se uma quantidade maior de registros, considerando-se a natureza e a abrangência das quatro bases de dados, utilizadas na pesquisa. O volume de registros recuperados manteve-se uma média anual de 5,91 trabalhos. Cabe apontar que os resultados das quatro bases são consideravelmente desiguais, o que é, provavelmente, consequência das variações dos periódicos nelas indexadas.

Observa-se que as bases de dados têm suas particularidades em relação às estratégias de busca que autorizam. Neste estudo, buscou-se obter aproximações nessas estratégias, porém essas diferenças apresentam resultados com outros mecanismos de filtragem, o que pode ter gerado divergência nos resultados, por não serem totalmente iguais às estratégias. Um exemplo de diferença é a não abordagem da opção de filtrar os registros por “revisado por pares” na base *Scopus*. Coincidentemente ou não, esta foi a base com maior número de registros recuperados.

O aparecimento dos termos nos diversos locais dos registros é instigador e provoca a necessidade de análise completa de cada artigo, individualmente, para aprofundar a compreensão sobre o objeto do trabalho. É importante observar também a atribuição de palavras-chave, considerando-se sua forte

relação com a temática do artigo. O aparecimento dos termos de pesquisas nos registros recuperados das bases de dados no campo de palavra-chave evidenciou-se como segundo na análise de ocorrência dos índices. Ou seja, comparativamente ao título é bem menor a inserção dos termos de pesquisas.

As pesquisas em SKOS, tem abrangência suficiente para visualizar os diversos focos de interesse por meio da categorização empregada. A concentração dos estudos se mostra relacionada com o fato de representar, em SKOS, os diversos sistemas de organização do conhecimento, com ênfase maior em tesouros. A contribuição deste estudo deu-se no sentido de buscar a visualização das pesquisas no mesmo sentido que os objetivos do SKOS. A maior identificação das pesquisas desenvolvidas com os aspectos de aplicação tende a reforçar a motivação de aperfeiçoar as formas de ampliar o alcance dessa visualização. O que já se consegue acrescentar é que há pouco destaque para as pesquisas teóricas destinadas a fundamentar ou homologar as práticas.

Embora não seja uma forma exclusiva de mensuração, o levantamento nos moldes em que foi realizado, mostrou que o assunto “SKOS” ainda não é explorado de modo suficiente. Considerando-se a cobertura internacional da literatura periódica desta pesquisa, o relativo silêncio, que disso se depreende, é revelador sobre a necessidade de novos estudos acerca da temática. A imperiosa necessidade de diálogo dos sistemas tradicionais de organização e representação do conhecimento, converte o SKOS em tema de indiscutível relevância para a efetiva compreensão do papel reservado aos SOCs no cenário digital. Esta pesquisa terá prosseguimento com a meta de incursionar pelas relações de complementaridade entre o SKOS e as ontologias a fim de melhor compreendê-las.

REFERÊNCIAS BIBLIOGRÁFICAS

Catarino, M. E., Cervantes, B. M. N., & Andrade, I. A. (2015). A representação temática no contexto da web semântica. *Informação e Sociedade*, 25(3), 105-116.

Hjørland, B. (2015). Theories are knowledge organizing systems (KOS). *Knowledge Organization*, 42(2), 113-128.

Hodge, G. (2000). Knowledge Organization Systems: an overview. In: G. Hodge. *Systems of knowledge organization for digital libraries: beyond traditional authority files* (pp. 3-9). Washington: Council on Library and Information Resources, 2000.

Moreira, W. (2012). Relações conceituais como ponto de inflexão entre linguagens documentais, terminologia e ontologias (pp. 123-127). Scire.

Palmonari, M. (2011). AERIA: Extending SKOS for the practical, yet well-founded, representation and integration of web schemas in the large. *Journal of Emerging Technologies in Web Intelligence*, 3(3), 227-238, doi:10.4304/jetwi.3.3.227-238.

Pastor-Sánchez, J. A., & Martínez-Méndez, F. J. F. (2010). Manual de SKOS (simple knowledge organization system, sistema para la organización del conocimiento simple). *Anales de Documentación*, 13, 285-320.

Ramalho, R. A. S. (2015). Ontologias e Knowledge Organization System (SKOS): aproximações e diferenças. In J. A. C. Guimarães, V. Dodebei (Orgs.). *Organização do conhecimento e diversidade cultural* (pp. 100-107). Marília ISKO-Brasil: FUNDEPE.

Smiraglia, R. P. (2014). *The Elements of Knowledge Organization*. Cham: Springer.

Wada, M. (2016). A SKOS Representation of synthesis in hierarchy classification schemes based on structure-expression style theory. *Library & Information Science*, (76), 1-33.



LOS MODELOS CONCEPTUALES EN LA CIENCIAS DE LA INFORMACIÓN: UN NUEVO RETO PARA LA DESCRIPCIÓN Y EL ACCESO INTEGRADO DE RECURSOS

Dunia Llanes-Padron¹, Manuela Moro-Cabero²

¹Universidad de La Habana, 0000-0001-8639-4706, duniallp@yahoo.es

²Universidad de Salamanca, 0000-0001-5301-1924, morooca@usal.es

RESUMEN En este artículo se analizan algunos modelos de representación desarrollados en las Ciencias de la Información (CI); se pretende mostrar que los modelos conceptuales internacionales y nacionales presentan un diseño novedoso para la descripción y el acceso integrado y entrelazado de la memoria cultural en el contexto de la web semántica. Así mismo, este análisis permite comprender las ventajas que ofrecen modelos como RiC-CM, FRBR, CIDOC-CM entre otros para la integración y compatibilidad de los sistemas de información de archivos, bibliotecas, museos y otras instituciones culturales. Este ensayo tiene como intención exponer una visión general de algunos modelos de la CI y una comparativa entre ellos. Además, pretende subrayar su importancia para el acceso, la distribución y uso abierto de los recursos de información en entornos digitales.

PALABRAS CLAVE *Representación de información, Descripción de Recursos, Modelos Conceptuales de Descripción en la CI, RiC-CM, FRBR, CIDOC-CM.*

ABSTRACT In this article we analyze some models of representation developed in the Information Sciences (CI); It is intended to show that international and national conceptual models present a novel design for the description and integrated access of cultural memory in the context of the semantic web. Also, this analysis allows understanding the advantages offered by models as RiC-CM, FRBR, CIDOC-CM for the integration and compatibility of the information systems of archives, libraries, museums and other cultural institutions. This essay intends to expose an overview of some models of CI and a comparison between them. In addition, it aims to underline its importance for the access, distribution and open access of information resources in digital environments.

KEYWORDS *Representation of Information, Resource Description, Conceptual Models for Description, RiC-CM, FRBR, CIDOC-CM.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUCCIÓN

Los modelos conceptuales desarrollados en el ámbito de la Ciencia de la Información tienen como premisa principal facilitar la interoperabilidad de los datos publicados en la web, con la finalidad de lograr una integración de los contenidos entre las diferentes instituciones de información. Estos modelos

garantizan una interoperabilidad semántica descentralizada y no fundamentada en el intercambio de esquemas de metadatos.

La representación de la información desde la perspectiva del paradigma posmoderno consolida sus bases teóricas a partir de modelos conceptuales, ontologías y aplicaciones tecnológicas. Según un estudio realizado por miembros del Consejo Internacional de Archivos, “un modelo conceptual es una técnica formal para la representación entre los principales conceptos y relaciones en un ámbito dado del conocimiento”. (Gueguen, et al., 2013, p. 9). Los profesionales de bibliotecas, museos y archivos, a partir de modelos creados para ambientes tecnológicos, han elaborado algunas iniciativas de modelado de datos para la representación de objetos de información y su integración en la web.

Las bibliotecas fueron las pioneras, en el contexto de información, en desenvolver un modelo conceptual. En el año 1998, la Federación Internacional de Asociaciones de Bibliotecarios y Bibliotecas (IFLA) publicó los Requisitos Funcionales para Registros Bibliográficos (FRBR, 1998). Estos requisitos presentan una nueva perspectiva del universo descriptivo y trabajan con las relaciones bibliográficas apoyándose en el modelo entidad-relación adoptado de las tecnologías de bases de datos. En los FRBR se definen:

[...] las entidades de los registros bibliográficos, los atributos, los tipos de relaciones,... los vínculos entre atributos específicos y relaciones (reflejados en los registros como elementos aislados de datos) con las distintas operaciones que realizan los usuarios cuando consultan registros bibliográficos. (IFLA, 1998, pp. 31-32).

Los FRBR, sin duda, rediseñaron el panorama, existente hasta el momento, de la descripción bibliográfica basada en estándares de datos y contenidos. Este modelo proporciona al usuario más facilidades para la búsqueda, recuperación y uso de la información y explica el universo bibliográfico de una manera más dinámica tanto para usuarios como para bibliotecarios.

El Consejo Internacional de Museos (ICOM), por su parte, desarrolló un modelo conceptual para la representación de informaciones vinculadas con el patrimonio cultural. El *Conceptual Reference Model* (CRM) fue elaborado por el Comité Internacional de Documentación (CIDOC) del ICOM. Este modelo fue aceptado por el Organismo Internacional de Normalización en el Comité Técnico ISO 46 en el año 2000 y convertido en una norma internacional en el 2006; en la actualidad cuenta con varias versiones, la última la 6.2 del 2015.

La función principal del CRM es facilitar el intercambio de información y la integración entre fuentes heterogéneas vinculadas al patrimonio cultural. Su objetivo es proporcionar las definiciones semánticas y las aclaraciones necesarias para recuperar y localizar las fuentes de información dispersas ya sea dentro de una institución o en Internet (CIDOC-CRM, 2015).

El *Conceptual Reference Model*, aunque se centró inicialmente en los museos, llegó a ser concebido como un modelo de referencia que podría servir para unos objetivos más amplios permitiendo un acceso integrado al patrimonio cultural, abarcando así el acceso a bibliotecas, archivos y museos. En este sentido, se formó en 2003 el Grupo de Trabajo Internacional sobre Armonización FRBR /CIDOC- CRM. El Grupo se centró en los conceptos FRBR de mapeo para conceptos CRM y en mejorar y afinar los conceptos CRM para facilitar el mapeado y la elaboración de un modelo semántico general para todas las instituciones de información (Gueguen, et al., 2013).

En el contexto archivístico, los profesionales motivados por el auge normativo y los beneficios tecnológicos para la recuperación y el acceso a la información, también, han desarrollado modelos nacionales de representación. Según la Comisión de Normas Españolas de Descripción Archivística (CNEDA), un modelo conceptual de descripción archivística es “la representación de la realidad a un alto nivel de abstracción en el que quedan descritos los tipos de entidad archivísticos, sus relaciones y atributos.” (CNEDA, 2017, p. 19).

Australia, Nueva Zelanda, Finlandia y España han sido los países que han elaborado y publicado modelos conceptuales para la descripción archivística. El modelo con más trayectoria es el *Australian Government Recordkeeping Metadata Standard* (AGRKMS), los neozelandeses, por su parte, publicaron un modelo que ha sido ampliamente difundido: las *Technical Specifications for the Electronic Recordkeeping Metadata Standard*. En Finlandia, los Archivos Nacionales (Arkistolaitos) publican, en el 2013, la versión borrador del *Finnish Conceptual Model for Archival Description*. Por último, en España, la CNEDA divulgó, hace algunos años, la versión del Modelo Conceptual de Descripción Archivística y Requisitos de Datos Básicos de las Descripciones de Documentos de Archivo, Agentes y Funciones (Entidades y Relaciones), recientemente, en el 2017, publicaron una versión final donde se incluyen, además, los atributos del modelo.

El desarrollo de estos patrones nacionales archivísticos, unido a la experiencia práctica de la aplicación de los estándares internacionales, trajo consigo que el *Internacional Council of Archives* (ICA) comenzará, hace unos tiempo, la realización de un modelo conceptual internacional que incluyera los datos y contenidos del conjunto normativo y, a su vez, redefiniera entidades, relaciones y atributos. A finales del 2016, el consejo publica el borrador del *Records In Contexts. A Conceptual Model For Archival Description* (RiC-CM). Con este modelo y su posterior desarrollo ontológico se facilita y viabiliza la interoperabilidad estructural y semántica de los contenidos archivísticos en la plataforma web.

Existen otros modelos conceptuales vinculados al patrimonio cultural, por ejemplo, la Asociación de Medios de Comunicación Europea EBU (*European Broadcasting Union*) también ha publicado, en el 2016, una primera versión de un modelo *Class Conceptual Data Model*¹ (CCDM-EBU) con una propuesta de ontología en OWL. Este modelo define un conjunto de entidades, relaciones y propiedades para describir objetos de información vinculados a los medios de comunicación. En este texto no se incluye el análisis de este modelo, próximamente será abarcado, por las autoras, en estudios más profundos que se realizarán sobre este tema.

El trabajo que se presenta tiene como objetivo analizar algunos de estos modelos y resaltar los puntos coincidentes y divergentes entre los esquemas de representación publicados en el área de la Ciencias de la Información. Se realiza un estudio de los estándares conceptuales de bibliotecas, museos y archivos y se determinan las entidades, sub-entidades y relaciones definidos. Todo ello con la intención de mostrar los beneficios de una descripción basada y aplicada en principios conceptuales, principalmente, aquellos relacionados con la interoperabilidad, el acceso y la reutilización de la información en el contexto de la web semántica.

¹ EBU. (2016). EBU Class Conceptual Data Model (CCDM). Recuperado de <https://tech.ebu.ch/docs/tech/tech3351.pdf>. Consultar ontología del modelo en: <https://www.ebu.ch/metadata/ontologies/ebucore/>

METODOLOGÍA

El estudio es, fundamentalmente, de naturaleza exploratoria y cualitativa y con un carácter descriptivo. Consideramos que una metodología de tipo cualitativo es apropiada para el tipo de investigación que se presenta, con el fin de cumplir los objetivos planteados. Se emplea, además, la técnica del análisis comparativo, basada en el examen de las diferencias y semejanzas que se establecen entre las FRBR, CIDOC-CRM, RiC-CM, AGRKMS, TSERMS, RDBDDAF y FCMAD² en relación a un conjunto de cuestiones (entidades, sub-entidades, relaciones). En un primer momento se realiza una contextualización del tema, partiendo del análisis de los modelos conceptuales mencionados. A partir del análisis se procede a exponer las ventajas de los modelos para facilitar la interoperabilidad semántica entre sistemas de información de diversas instituciones y la publicación de datos abiertos.

RESULTADOS

Los FRBR y CIDOC-CRM establecieron una nueva forma de representación del conocimiento, revolucionaron el contexto descriptivo en la Ciencia de la Información; bibliotecarios y museólogos se apoyaron para el desarrollo de estos análisis conceptuales en las normas de descripción publicadas en cada área como las ISBD, GARE, MARC, AITF, entre otras. En el siguiente cuadro se ejemplifican algunas de las entidades definidas en ambos modelos.

Tabla 1. Ejemplos de entidades en FRBR y CIDOC – CRM

<i>MODELOS CONCEPTUALES BIBLIOTECAS Y MUSEOS</i>	
<i>ENTIDADES</i>	
<i>BIBLIOTECAS (FRBR)</i>	<i>MUSEOS (CIDOC-CRM)</i>
Obra	Entidad CRM
Expresión	Entidad Temporal
Manifestación	Estado Material
Ítem	Documento
Personas	Ítem
Entidades Corporativas	Imagen
Concepto	Colección
Objeto	Objeto biológico
Acontecimiento	Objeto físico
Lugar	Objeto lingüístico
	Objeto conceptual
	Cosa

² FRBR: Requisitos Funcionales de los Registros Bibliográficos para Bibliotecas.

CIDOC-CRM: *International Committee for Documentation - Conceptual Reference Model*.

RiC-CM: *Records in Contexts. A Conceptual Model For Archival Description*.

AGRKMS: *Australian Government Recordkeeping Metadata Standard*.

TSERMS: *Technical Specifications for the Electronic Recordkeeping Metadata Standard* de Nueva Zelanda.

RDBDDAF: Modelo Conceptual de Descripción Archivística y Requisitos de Datos Básicos de las Descripciones de Documentos de Archivo, Agentes y Funciones de España.

FCMAD: *Finnish Conceptual Model for Archival Description*.

	Actor (Persona o Grupo)
	Autoridad documental
	Autor institucional
	Actividad
	Evento
	Periodo
	Lugar
	Destrucción
	Actividad
	Adquisición
	Transferencia de custodia
	Modificación
	Producción
	Identificador
	Movilidad
	Características
	Creación

Fuente: *Elaboración propia*

Tanto para bibliotecas como para museos se han identificado los conceptos principales que estarán definidos, representados e interrelacionados en los sistemas de información publicados en ambiente web; cada una con sus particularidades respondiendo al área del conocimiento representada. Algunas incompatibilidades que se aprecian entre las entidades de los dos modelos son resultado de la estructuración de las informaciones en el *Conceptual Reference Model* del ICOM, donde las entidades temporales son elementos adicionadores de metadatos. Según Lima (2008), una de las innovaciones del CRM es el uso de varias entidades temporales, aspecto que lo diferencia de la mayoría de los modelos que tienen un recurso como objeto central de interés y atributos que los caracterizan. A partir de esta idea, los eventos temporales son definidos como agentes, hechos, objetos, lugares, designaciones, intervalos de tiempo, entre otros (Lima, 2008, citado en Santo, 2016, p. 2).

Si se analizan las entidades definidas en ambos modelos, se puede decir que, si bien existen diferencias proporcionadas por los rasgos distintivos de estas instituciones culturales, hay entidades como documento, ítem, personas, lugar, objeto entre otros que pueden ser comunes a bibliotecas y museos, lo que facilita la proyección de un modelo común que integre los objetos de varias instituciones de información en el entorno web.

En la archivística, hasta finales del siglo pasado los modelos conceptuales presentaban un enfoque unidimensional, en la actualidad, responden a la categoría de modelos pluridimensionales de descripción. En los cuatro modelos nacionales publicados se identifican varias entidades, establecen relaciones y se definen atributos para el conjunto de entidades y sus sub-tipos.

Una entidad archivística, según la comisión española es: “Cualquier objeto real o abstracto que existe, ha existido o puede llegar a existir” (CNEDA, 2017, p. 18). Los modelos australianos, neozelandés, finlandés y español han definido explícitamente sus entidades, en la siguiente tabla se pueden consultar por países.

Tabla 2. Entidades archivísticas en los modelos nacionales de descripción.

<i>MODELOS CONCEPTUALES NACIONALES DE ARCHIVOS</i>			
<i>ENTIDADES</i>			
<i>AUSTRALIA (AGRKMS)</i>	<i>NUEVA ZELANDA (TSERMS)</i>	<i>ESPAÑA (RDBDDAF)</i>	<i>FINLANDIA (FCMAD)</i>
Documento	Documento	Documento	Manifestación
Agente	Agente	Agente	Expresión
Función (Business)	Función (Business)	Función y sus divisiones	Ítem
Norma (Mandato)	Norma (Mandato)	Norma	Agente
Relación	Relación	Concepto	Funciones
		Asunto/ Acontecimiento	Norma (Mandato)
		Lugar	Recursos de Información
			Ciclo de Vida
			Lugar
			Período
			Materia

Fuente: *Elaboración propia*

En todas las propuestas se definen cuatro entidades comunes: documento, agente, función y norma. Los modelos nacionales han contemplado los temas aceptados por la comunidad internacional en su conjunto normativo y así se manifiestan en la correspondencia entre entidades y los estándares internacionales.

A su vez, se puede apreciar la enunciación de otras entidades tales como concepto, acontecimiento, lugar, tiempo, materia, manifestación, expresión e ítem que se corresponden con definiciones reguladas en los FRBR y en CIDOC-CRM, lo que demuestra la influencia de estos modelos en los archivos. Por último, entidades como relación, mandato, ciclo de vida, recursos de información son propuestas innovadoras para la descripción en los entornos archivísticos.

Además de las entidades, en los modelos nacionales de archivo se delimitan los subtipos de entidades. Según la CNEDA, un subtipo es “una categoría de un subconjunto de entidades pertenecientes al mismo tipo que poseen los mismos atributos y relaciones” (CNEDA, 2017, p. 21). En el próximo cuadro se ejemplifican los subtipos propuestos en los modelos conceptuales australiano, neozelandés y español para las entidades archivísticas. El modelo finlandés no define estas categorías manifiestamente.

Tabla 3. Subtipos de entidades en los modelos nacionales de descripción

<i>MODELOS CONCEPTUALES NACIONALES DE ARCHIVOS</i>		
<i>SUB-TIPOS DE ENTIDADES</i>		
<i>AUSTRALIA (AGRKMS)</i>	<i>NUEVA ZELANDA (TSERMS)</i>	<i>ESPAÑA (MCDA)</i>
<i>DOCUMENTO</i>		
Archivos (<i>Archives</i>)	Archivos (<i>Archives</i>)	Grupo de Fondos
Fondo (<i>Archive</i>)	Fondo (<i>Archive</i>)	Fondo
Series (<i>Series</i>)	Series (<i>Series</i>)	División de Fondo
Documento Compuesto (<i>File</i>)	Documento Compuesto (<i>File</i>)	Serie
Documento Simple (<i>Item</i>)	Documento Simple (<i>Item</i>)	Sub-serie
Documentos ordenados (<i>Transaction Sequence</i>)	Documentos ordenados (<i>Transaction Sequence</i>)	Fracción de serie
-	-	Unidad documental
-	-	Colección
-	-	División de Colección
-	-	Componente Documental
<i>AGENTE</i>		
Institución	Institución	Institución
Organización	Agencia	Familia
Grupo de Trabajo	Grupo de Trabajo	Persona
Persona	Persona	
Sistema físico o electrónico que realiza transacciones comerciales (<i>Mechanism</i>)	Sistema físico o electrónico que realiza transacciones comerciales (<i>Mechanism</i>)	-
<i>FUNCIÓN</i>		
Función contextual (<i>Ambient fuction</i>)	Función contextual (<i>Ambient fuction</i>)	Función
Función	Función	División de función
Actividad	Actividad	Actividad/Proceso
Transacción	Transacción	Acción
<i>NORMA (MANDATO)</i>		
Legislación	Legislación	-
Regulación	Regulación	-
Políticas	Políticas	-
Procedimientos de negocios (<i>Business Rule</i>)	Procedimientos de negocios (<i>Business Rule</i>)	-
Requerimientos de negocios (<i>Stakeholder Requirement</i>)	Requerimientos de negocios (<i>Stakeholder Requirement</i>)	-
Expectativas de la comunidad de negocios (<i>Community Expectation</i>)	Expectativas de la comunidad de negocios (<i>Community Expectation</i>)	-
Normas	Normas	-
Instrumentos	Instrumentos	-
Códigos de Conducta	Códigos de Conducta	-
Sistemas de requerimientos	Sistemas de requerimientos	-

<i>RELACIONES</i>		
Relaciones de Procedencia	Relaciones de Procedencia	-
Eventos de gestión documental (Recordkeeping Event)	Eventos de gestión documental (Recordkeeping Event)	-

Fuente: *Elaboración propia*

Los documentos, los agentes, las funciones, los mandatos y las relaciones son las entidades para las cuales estos modelos definen categorías de subtipos. Las comunidades archivísticas especifican subtipos de acuerdo a sus realidades teóricas y prácticas. Se puede apreciar que los modelos de Australia y Nueva Zelanda presentan las mismas entidades y subtipos, esta similitud puede estar dada porque ambas normativas se basaron en revisiones del modelo australiano *Recordkeeping Metadata Standard for Commonwealth Agencies* en su versión 1.0 (2000).

El modelo español por su parte, define con mayor profundidad las sub-entidades del documento, es decir amplía los niveles de descripción, respondiendo a la tradición archivística española y a los niveles de organización representados en los diversos instrumentos de descripción publicados. Este modelo aunque presenta el objeto norma como una entidad independiente no establece subtipos como sucede en los modelos australiano y neozelandés.

Como se mencionó anteriormente, el ICA, también ha elaborado un modelo con alcance internacional, *Records In Contexts. A Conceptual Model For Archival Description* se basa en los modelos nacionales de archivos y en los modelos de bibliotecas y museos. Documentos en contexto es la idea central de este modelo; a partir del análisis de la procedencia se plantean nuevas dinámicas para la descripción. La procedencia debe respetar los fondos y el orden original pero desde la comprensión de sus interrelaciones con las dinámicas sociales, culturales y económicas que dan origen a la documentación y su conservación. La descripción, este medio, representa las complejidades y convergencias de los contextos de producción y acumulación, desmarcándose de la fisicalidad y comprendiendo la procedencia como un constante devenir y no como una acción estática e invariable.

En el RiC-CM se han identificado y diferenciado las principales entidades que estarán presentes en los sistemas de descripción archivísticos. En el modelo internacional se definieron un total de catorce entidades, a continuación se ejemplifican:

Tabla 4. Tipos de entidades definidas en el RiC-CM

<i>MODELO CONCEPTUAL INTERNACIONAL (RiC-CM)</i>	
<i>ENTIDADES ARCHIVÍSTICAS</i>	
RiC-E1	Documento
RiC-E2	Componente Documental
RiC-E3	Conjunto Documental
RiC-E4	Agente
RiC-E5	Ocupación (Profesión)
RiC-E6	Posición
RiC-E7	Función
RiC-E8	Función Abstracta
RiC-E9	Actividad
RiC-E10	Mandato (Norma)

RiC-E11	Forma Documental
RiC-E12	Período/Fechas
RiC-E13	Lugar
RiC-E14	Concepto/Asunto

Fuente: *Elaboración propia*

Las entidades presentadas constituyen los objetos archivísticos que serán descritos e interrelacionados en los sistemas integrados de información. El Consejo, basado ISAD(G), ISAAR(CPF), ISDF, ISIAH, en los modelos de Australia, Nueva Zelanda, España, Finlandia, los FRBR y el CIDOC-CRM, ha propuesto una serie de entidades que representan la realidad archivística y con ello llega a un consenso internacional.

Asimismo, si analizamos las tablas 1 y 2 donde se recogen las entidades definidas en modelos para bibliotecas, museos y sistemas de archivos nacionales, se puede aseverar que RiC-CM tiene elementos coincidentes con estas propuestas. Las entidades documento, agente, función, norma, lugar, período y concepto son denominadores comunes en los modelos conceptuales desarrollados en las Ciencias de la Información y constituyen entidades principales.

En el RiC no se incluyen entidades como manifestación, expresión, ítem, colección u objeto que fueron definidos en el Modelo Conceptual Archivístico de Finlandia, en FRRB o en el CIDOC-CRM, lo que demuestra que se pueden, también, establecer diferenciaciones en los esquemas descriptivos.

En el modelo archivístico internacional no se presentan, directamente, subtipos de estas entidades; se explica que cada una de las entidades se tratará como tipos distintos y deben ser descritas de forma diferente aunque muchas propiedades serán compartidas. En el RiC se aboga por una separación de los principales componentes de la descripción aunque interrelacionados entre sí, tal y como sucede con el CRM y los FRBR.

La identificación de las relaciones entre las entidades de información es otro de los aspectos determinantes en estos modelos conceptuales. Según la CNEDA (2017), una relación “es una correspondencia, asociación o vinculación de cualquier tipo entre entidades”, p. 20. En el contexto bibliotecario, por ejemplo, se definen las diversas relaciones que se pueden establecer entre una obra y otra, entre una expresión y otra, entre una manifestación y un ítem, etcétera. En los archivos, por su parte, se pueden definir relaciones entre documentos y productores, entre productores y funciones, entre documentos y conceptos y muchas más; así mismo los museos definen las interrelaciones que se suceden entre las distintas entidades que se han definido en el CRM.

En el modelo conceptual para bibliotecas (FRBR) se establecen varias relaciones a partir de los tres grupos de entidades definidos. El primer grupo comprende los productos de la creación intelectual o artística: obra, expresión, manifestación e ítem. El segundo grupo comprende las entidades responsables del contenido intelectual: persona y entidad corporativa. El tercer grupo comprende un conjunto adicional de entidades que junto con las entidades del primer y el segundo grupo puede servir como materia de una obra: concepto, objeto, acontecimiento y lugar. (IFLA, 1998).

La siguiente figura ejemplifica las relaciones primarias que se establecen con las entidades del primer grupo de FRBR, proporcionadas a través de las siguientes relaciones: realizado mediante, materializado en y ejemplificado por. Así mismo, para el segundo grupo se definen otras: realizado mediante, producido por o poseído por, a su vez, las entidades de los tres grupos pueden vincularse a través de las

relaciones de materia. Estos son sólo algunos de los ejemplos relacionales, el modelo expone detalladamente todas las posibles vinculaciones entre las entidades bibliográficas.

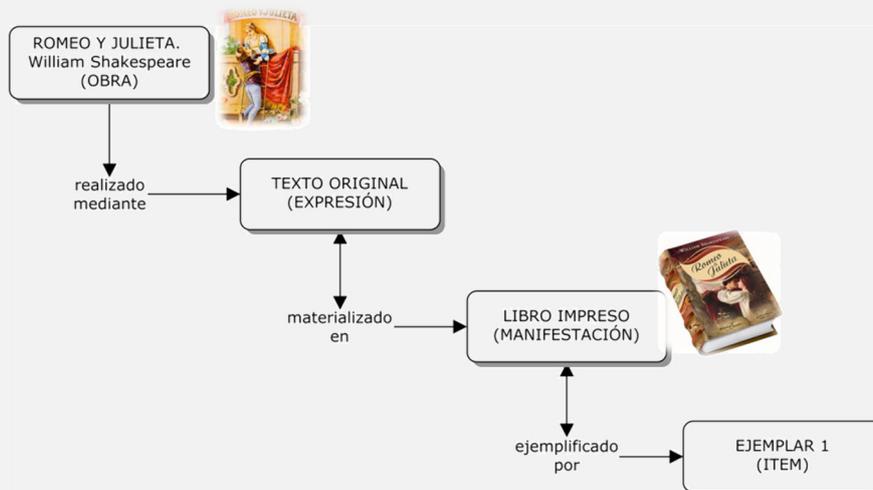


Figura 1. Relaciones establecidas entre las entidades Obra/Expresión/Manifestación/Ítem

Fuente: Basada en los FRBR, 2014

Por su parte, en el RiC-CM se establecen un total de 792 relaciones; esta cantidad de relaciones puede aumentar cuando se publique la versión final del modelo. En el texto se muestran un conjunto de relaciones binarias simples para cada una de las entidades, por ejemplo para la entidad productor: es creado por/ es creador de. Muchas de estas relaciones tienen una semántica similar; esto puede ejemplificarse, también, con la relación es miembro de, con esta relación, los documentos pueden vincularse con conjuntos de documentos, los conjuntos de documentos con otros conjuntos de documentos y los productores con otros productores del tipo grupo. Esto sucede con otros tipos de relaciones definidas en el modelo. En las figuras 2 y 3 se ejemplifican algunos esquemas relacionales que se pueden establecer en la Archivística.

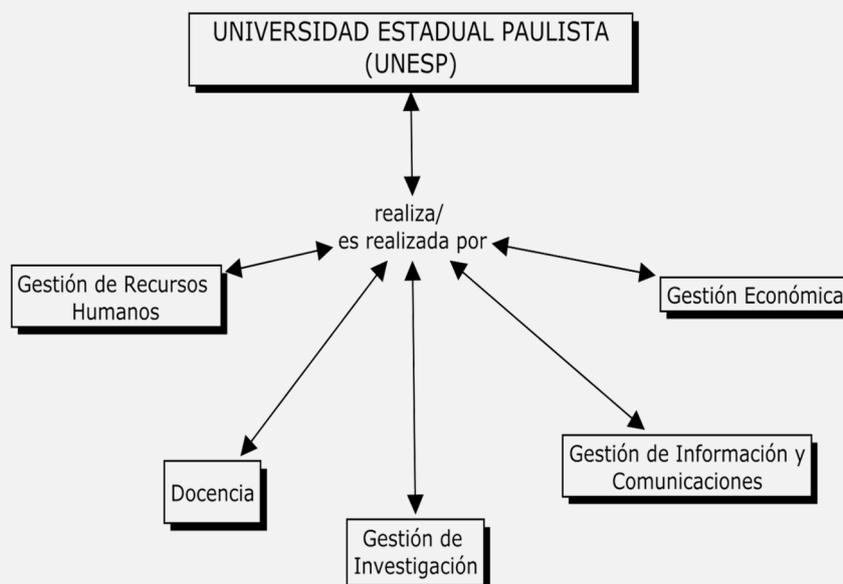


Figura 4. Relaciones entre las entidades Productor/Funciones en un Modelo de Archivos.

Fuente: LLanes. D. (2016), p.45.

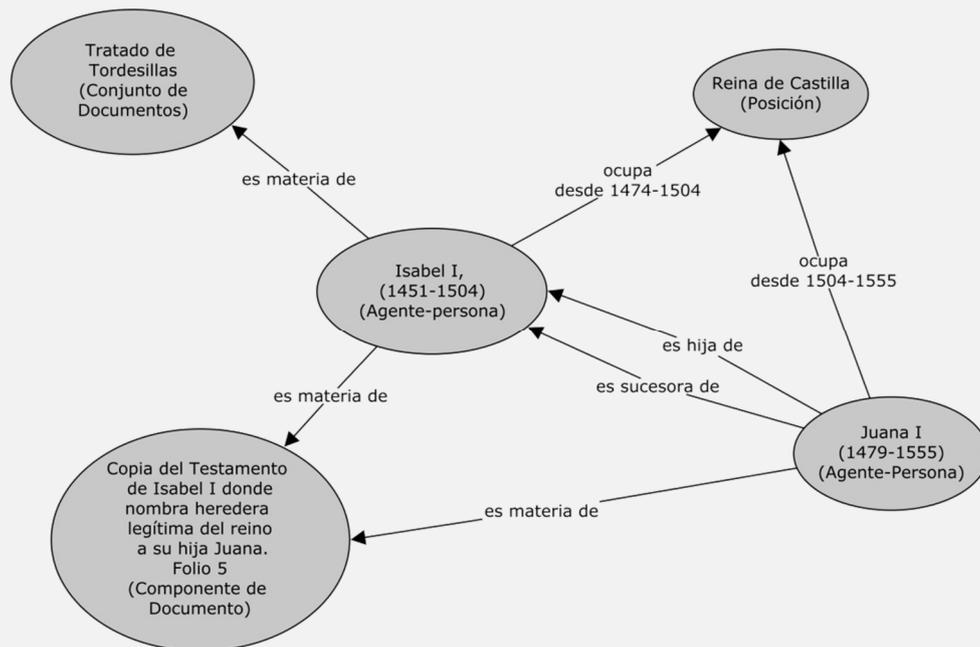


Figura 5. Relaciones archivísticas establecidas a partir del RiC-CM.

Fuente: Elaboración propia

La determinación de una amplia variedad de relaciones garantiza la interrelación de los contenidos de los recursos de información con los diversos ambientes en los que fueron producidos. Asimismo, el establecimiento de estas redes facilita la exposición gráfica de los contenidos descriptivos en el entorno tecnológico, materializado en los softwares que se desarrollan a partir de las ontologías y de los modelos y, además, garantiza el incremento de la capacidad de búsqueda de los sistemas así como la recuperación, las interrelaciones, la reutilización de los datos entrelazado y el uso de la información entre sistemas diversos.

CONCLUSIONES

El paradigma posmoderno en el entorno de la Ciencia de la Información está influyendo en el proceso de transición hacia nuevas formas de representación más alineadas con las necesidades y los enfoques actuales. Los modelos de conceptualización son el camino a seguir hacia la interoperabilidad semántica descentralizada y no fundamentada en el intercambio de esquemas de metadatos. La creación de ontologías OWL son el otro aspecto determinante para implementar los modelos y poder publicar los datos entrelazados en la web. Además, la aplicación de estos modelos facilita la reutilización de información entre los sistemas de información de instituciones culturales diversas.

La aplicación de conceptos como web semántica, ontologías o datos abiertos entrelazados son determinantes para las buenas praxis en el proceso de descripción de recursos y extiende la colaboración y el intercambio de datos entre bibliotecas, museos, archivos y otras instituciones culturales. La descripción realizada, a partir de los modelos conceptuales, facilitará la consulta de recursos heterogéneos entrelazados a partir de un enfoque único dado por los modelos y como resultado permitirá descubrir y compartir más información.

Los modelos conceptuales de representación instituyen herramientas de integración y distribución de la información. A partir de la aplicación de estos modelos, las instituciones podrán publicar conjuntos de datos fiables, accesibles y reutilizables y conseguirán preservarlos a corto, medio y largo plazo. La comunidad de profesionales de la CI tiene ante sí un nuevo reto en cuestión de descripción de contenidos, la elaboración de un modelo semántico integrador que unifique todos los recursos que forman parte de esta ciencia. Los FRBR, CIDOC-CRM, RiC-CM, pueden ser concertados y unificados con la finalidad de lograr un acceso integrado, distribuido y entrelazado del patrimonio cultural mundial en pro de la democratización de la información.

REFERENCIAS BIBLIOGRÁFICAS

CNEDA. (2017). NEDA-MC. Modelo conceptual de descripción archivística: Entidades, relaciones y atributos. Recuperado de <https://www.mecd.gob.es/dms/mecd/cultura-mecd/areas-cultura/archivos/mc/cneda/documentacion/NEDA-MC-2017/NEDA-MC-2017.pdf>

EBU. (2016). Class conceptual data model. (CCDM). Recuperado de <https://tech.ebu.ch/docs/tech/tech3351.pdf>

Gueguen, G., Fonseca, V. M. M. da, Pitti, D. V., & Sibille-de Grimoüard, C. (2013). Hacia un Modelo Conceptual Internacional de Descripción Archivística: Un informe preliminar del Grupo de Expertos en Descripción Archivística del Consejo Internacional de Archivos. *Métodos de información*, 4(7), 155-172. <https://doi.org/10.5557/IIMEI4-N7-155172>.

ICA. (2016). Records in contexts: A conceptual model for archival description (Consultation Draft v0.1). Recuperado de <http://www.ica.org/sites/default/files/RiC-CM-0.1.pdf>

ICOM. (2015). Conceptual reference model. Recuperado de http://www.cidoc-crm.org/sites/default/files/cidoc_crm_version_6.2.pdf

IFLA. (1998). Requisitos Funcionales de los Registros Bibliográficos: Informe final. Recuperado de <http://www.ifla.org/files/assets/cataloguing/frbr/frbr-es.pdf>

LLANES, D. (2016). La descripción archivística en los tiempos posmodernos: Conceptos, principios y normas. Recuperado de https://www.marilia.unesp.br/Home/Publicacoes/la-description_ebook.pdf



ORGANIZAÇÃO E REPRESENTAÇÃO DA INFORMAÇÃO CORPORATIVA

Maurício B. Almeida¹, Christiano P. Pessanha²

¹Universidade Federal de Minas Gerais, Belo Horizonte, Brasil, mba@eci.ufmg.br

²Universidade Federal de Minas Gerais, Belo Horizonte, Brasil, chrisspess@gmail.com

RESUMO A presente pesquisa discute os fundamentos de uma arquitetura baseada em ontologia para gestão da informação em corporações, a partir de três esforços principais, os quais discutem o estatuto ontológico das corporações; as relações entre as unidades da corporação; e as obrigações que uma corporação deve gerenciar. Duas dimensões são utilizadas para a análise: a primeira, descritiva, aborda a estrutura das unidades e subunidades corporativas usando teorias da ontologia formal; a segunda, prescritiva (normativa), se vale de teorias da ontologia do social para organizar direitos e obrigações. A relevância desta análise reside na perspectiva de delinear arquitetura da informação bem fundamentada para corporações através de uma melhor compreensão da sua natureza, de sua estrutura e dos poderes deônticos envolvidos. Dessa forma, é possível introduzir um *framework* adequado aos modernos sistemas de informação no contexto da Web Semântica.

PALAVRAS-CHAVE *corporações, ontologias, sistemas de informação, gestão do conhecimento.*

ABSTRACT The present work discusses the fundamentals of an ontology based architecture for the information management within corporations from three main efforts: the ontological status of corporations; the relations among the units of corporations; and the obligations that corporations have to manage. Two dimensions are used for the analysis: the first one, descriptive, approaches the structure of corporative units and subunits using theories of formal ontology; the second one, prescriptive (normative), uses theories of social ontology to organize rights and obligations. The importance of such analysis rests on the perspective of delineating a well-founded information architecture for corporations through of a better understanding from their nature, their structure and the involved deontic powers. Thus, we hope to introduce a framework that is suitable to the modern information systems in the scope of Semantic Web.

KEY-WORDS *corporations, ontologies, information systems, knowledge management.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Corporações têm investido significativamente em informação e gestão do conhecimento, particularmente, no desenvolvimento de sistemas de informação. Entre as técnicas empregadas o uso de ontologias tem recebido cada vez mais atenção (Almeida, Barbosa, 2009 & Fonseca, 2007 & Guizzardi, 2005). Ontologia é um termo da Filosofia, que tem sido utilizado em diferentes perspectivas na Ciência da Computação, Ciência da Informação, Lógica, Filosofia, entre outros (Obrst, Hughes, Ray, 2006).

Ontologia é um termo que se origina em Filosofia também empregado para descrever uma estrutura hierárquica composta por entidades e relações para fins de representação. A questão da definição de ontologias reside no fato de que diferentes comunidades de pesquisa adotaram diferentes perspectivas: Ciência da Computação, por exemplo, Inteligência Artificial, Bancos de Dados e Engenharia de Software; Biblioteca e Ciência da Informação; Lógica e Filosofia, para mencionar alguns (Obrst, Hughes e Ray, 2006).

Uma diversidade de iniciativas para o uso de ontologias em corporações podem ser encontradas na literatura desde a década de 1990 (Bernus, Nemes e Williams, 1996; Fillion et al., 1995; Fox, 1992; Schlenoff, 1996; Uschold et al., 1998) . No entanto, a pesquisa em ontologias não possui uma abordagem de ontologia aplicada para corporações que espelha duas principais aplicações de ontologias: ontologia como um inventário de entidades para modelagem de sistemas de informação e ontologia como uma teoria formal para fins de raciocínio automático. Este artigo apresenta uma pesquisa em andamento que busca cobrir esta lacuna. Nesse ponto da pesquisa em andamento, nos concentramos em três esforços: i) lançar alguma luz sobre o status ontológico das corporações; ii) compreender as relações entre as unidades corporativas; e iii) abordar as obrigações que uma corporação precisa gerenciar. Para alcançar esses propósitos, apresenta-se um *background* de teorias formais sobre a natureza da corporação. Em seguida, apresenta-se uma análise da corporação através de duas dimensões, à saber, uma descritiva e uma prescritiva (normativa). Na dimensão descritiva, aborda-se a estrutura da corporação do ponto de vista de unidades e subunidades; na dimensão normativa, aborda-se a estrutura da corporação do ponto de vista de seus direitos, deveres e obrigações. No viés descritivo da investigação, empregam-se princípios e noções de ontologias formais de alto nível; no viés normativo, utiliza-se a abordagem da ontologia social, incluindo teorias de atos sociais, atos de fala e atos dos documentos.

O delineamento de uma teoria ontológica para corporações oferece a base de uma arquitetura para organizar informações e conhecimento, capaz de integrar e coordenar uma extensa variedade de sistemas de informação.

METODOLOGIA

A pesquisa em andamento baseia-se em três passos metodológicos já mencionados: i) o estatuto ontológico das corporações, realizado uma revisão da literatura das principais teorias sobre a natureza das corporações; ii) as relações entre as unidades subunidades das corporações, realizado pela aplicação de princípios ontológicos na composição da corporação; iii) os direitos e as obrigações de uma corporação utilizando, para tanto, teorias bem fundamentadas da ontologia do social.

Originário da Metafísica, o termo ontologia se refere a ontologia natural que diz respeito a classificação exhaustiva dos tipos naturais e suas relações. Artefatos de cunho social particularmente importantes para essa pesquisa – como a moeda, governos, nações, contratos, entre outros – são abordados em teorias da ontologia do social.

Na Ciência da Computação e Ciência da Informação o termo ontologia designa atividades de modelagem para representar a realidade, a qual é especificada em linguagem formal. Visando o raciocínio automático e a recuperação da informação, novos sistemas de categorias têm sido desenvolvidos, por exemplo, a DOLCE – Descriptive Ontology for Linguistic and Cognitive

Engineering (Gangemi, 2002) e a BFO – Basic Foundational Ontology são denominadas ontologias de alto nível, unindo artefatos computacionais especificados em linguagem formal e fundamentados em teorias filosóficas. A BFO, largamente aplicada em sistemas de informação de domínios como medicina, direito, bioinformática, entre outros segue o realismo ontológico, provendo categorias genéricas para categorizar entidades em domínios de conhecimento. A BFO, apresentada na figura 1, é uma ontologia formal de alto nível. Detalhes sobre os níveis de categorização da BFO estão disponíveis na literatura (Grenon, Smith, Goldberg, 2004 & Smith, 2003)

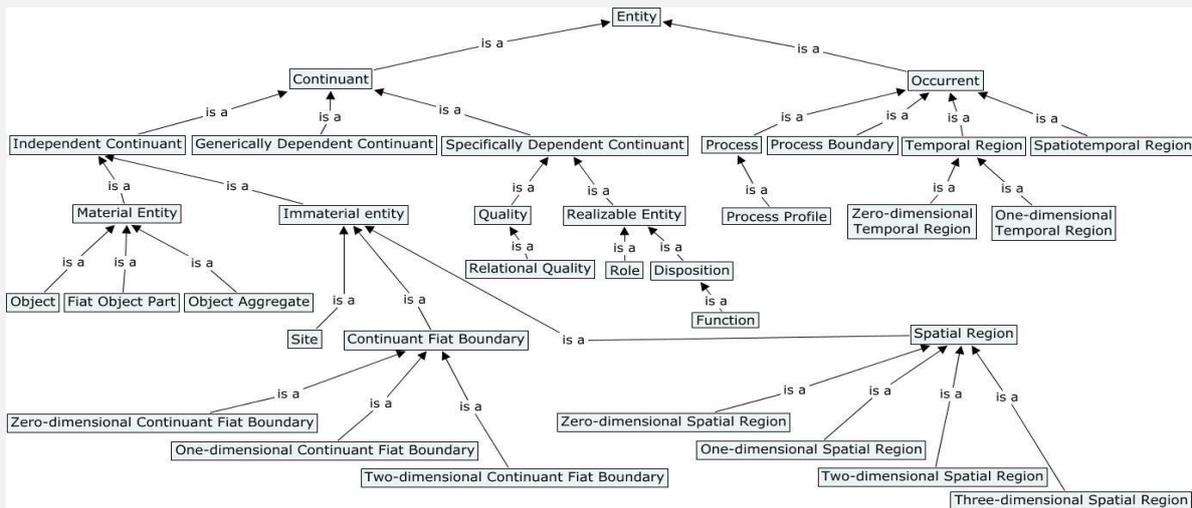


Figura 1. BFO, seus níveis e categorias. Fonte: Smith et al. (2015)

Diversos formalismos podem ser utilizados para expressar as noções intuitivas que as pessoas fazem uso em atividades de modelagem. Dois exemplos são a teoria da mereologia e a teoria das partições granulares. Nessa pesquisa, as duas teorias são utilizadas para explicar a dimensão descritiva das corporações.

Mereologia é uma teoria que trata das relações entre partes e todo e entre as partes dentro de um todo, utilizando-se de dois grupos de princípios: i) princípios de decomposição, que vão do todo para as partes; e ii) princípios de composição, que vão das partes para o todo. A mereologia, apesar de conter vários princípios relevantes, trata de entidades materiais e não parece ser a melhor opção para explicar as corporações e suas unidades. É difícil acreditar que a relação entre uma corporação e seus membros é uma relação parte-todo, a medida em que em mereologia, a relação parte-total é transitiva. Consequentemente, pode-se obter declarações inconsistentes, por exemplo: se João faz parte de uma corporação, qualquer parte de João, como seus olhos ou sua boca, também faz parte da corporação. Além disso, a mereologia não pode explicar o fato de uma corporação preservar sua identidade ao longo do tempo mesmo quando perde ou ganha membros (Vizenor, 2006). Para maiores informações, referências consolidadas sobre o assunto são por exemplo (Simon, 1987) e (Varzi, 1996).

Partição granular, por outro lado, é o nome dado à artefatos cognitivos que as pessoas utilizam para nomear, listar, ou catalogar, coisas do mundo (Bittner, Smith, 2003). Todo tipo de partição considera a existência de objetos de dois tipos, a saber, objetos *bona fide* e objetos *fiat*: o primeiro existe independentemente de qualquer atividade partitiva, enquanto o segundo é criado via projeção de partições na realidade. A teoria das partições granulares é composta pela Teoria A, que reflete a noção de que unidades de partição podem reconhecer objetos *fiat* baseados na cognição humana; e Teoria B, que explica como objetos *fiat* são criados via projeção de partições em uma porção da realidade.

Para explicar a dimensão normativa das corporações, utiliza-se aqui a ontologia social, envolvendo em particular, os atos sociais, os atos mentais, os atos da fala, a intencionalidade e os atos dos documentos.

A teoria dos atos sociais fundamenta conceitos jurídicos a priori pela noção de ato social, de cunho intencional, que necessita ser percebido por outra pessoa. Os atos sociais fundamentam os atos da fala e atos dos documentos. Estados mentais são, por sua vez, necessários para a compreensão dos atos da fala e da intencionalidade. Os fenômenos mentais são de vital importância para a Filosofia da Linguagem, de onde se origina a teoria dos atos da fala, classificada como um ramo da Filosofia da Mente (Searle, 1984).

A teoria dos atos da fala foi proposta inicialmente pelo filósofo britânico John L. Austin como método de análise para questões filosóficas via linguagem. O ato da fala seria a unidade básica de sentido da linguagem constituída das dimensões locucionária (dimensão linguística), illocucionária (contém o ato performativo em si) e perlocucionária. Entre os diversos atos da fala classificados por Searle (Searle, 1976), o declarativo é relevante para a dimensão normativa das corporações. Devido à evanescência dos atos da fala, que os restringe temporalmente, Smith (2012) propõe os atos dos documentos, uma vez que estes podem ser registrados e manipulados, permitindo novos tipos de relações sociais, poderes sociais e deontico para além das limitações da memória.

Finalmente, a intencionalidade é um conceito que distingue a propriedade dos fenômenos mentais de estar direcionada a um objeto real ou imaginário. A performance de um ato illocucionário especifica um estado mental de quem o profere, pois exhibe as condições de satisfação idênticas ao estado mental intencional associado.

A análise ontológica das corporações pode ser introduzida pelas hipóteses: i) corporações mantêm sua identidade no tempo; ii) corporações são entidades artificiais (entidades *fiat*); iii) corporações são agregados não somativos; iv) corporações são entidades duradouras. Tais hipóteses devem ser verificadas através de uma análise ontológica que revele entidades naturais e sociais que compõem a corporação. A ontologia BFO e os demais subsídios teóricos já apresentados são agora utilizados para descrever as dimensões descritiva e normativa de uma corporação.

A primeira hipótese e verificação correspondente busca identificar se, no contexto da BFO, corporações são continuantes independentes, entidades que se alteram ao longo do tempo enquanto preservam sua identidade. Correspondem, ainda que não totalmente, a substâncias aristotélicas (Smith, 1997). A segunda hipótese e verificação correspondente diz respeito ao tipo de todo: i) as corporações são um todo somativo?; ii) um todo integral? ou iii) um todo agregado?

Todos somativos são objetos cujas mesmas partes são iguais, uma noção que não corresponde a corporações que preservam sua identidade quando ocorrem mudanças em seus membros. Um todo integral possui uma forte conexão estrutural entre seus membros, diferente do que ocorre nas corporações. Finalmente, agregados não são definidos em termos de seus próprios elementos, possuindo elementos independentes, porém ligados. Assim, as corporações se identificam com agregados, entidades materiais que consistem numa pluralidade de objetos que são membros-parte destas últimas pelo tempo que existirem.

Porém, a mereologia não parece o melhor *framework* explicativo para corporações e seus membros. Não seria intuitivo supor como parte-todo a relação entre uma corporação e seus membros, visto ser esta relação transitiva. A teoria das partições granulares, ao contrário, utiliza-se de aspectos cognitivos

para mostrar como as pessoas particionam o mundo baseando-se na distinção entre objetos *bona fide* e objetos *fiat*. Utilizando a Teoria A, pode-se definir unidades e subunidades de uma corporação como objetos *fiat* (uma entidade BFO). De maneira semelhante, através da Teoria B, é possível definir a relação de projeção entre as unidades-subunidades e a realidade, bem como a relação de localidade entre um objeto e uma unidade.

Porém, se uma corporação é um tipo de agregado o que diferenciaria uma corporação de uma colônia de formigas, que também pode ser definida é agregado? A distinção é feita na dimensão normativa das corporações.

Se unidades de uma corporação são *fiat objects*, como estes tipos de objeto passam a existir? Através da cognição, por uma partição mental delimitando, por exemplo, uma unidade de uma corporação. Uma vez que corporações são entidades duradouras, como um objeto *fiat* mantém-se na existência? Através da transformação da forma verbal das normas na forma escrita.

Um ato social é um tipo de ato que precisa ser percebido (Smith, 2012). Um tipo especial de ato social, que é relevante para os propósitos desse artigo é a declaração. As declarações expressam o que “deve ser”, por exemplo, uma promessa é uma declaração de como as coisas deveriam ser. As declarações podem criar ou destruir a realidade: uma promessa cria uma obrigação para uma pessoa, bem como uma reivindicação para outra pessoa; mas uma declaração também pode revogar alguma ordem ou obrigação.

Utiliza-se assim a teoria dos atos dos documentos para registrar as causas dos pedidos e obrigações, pois documentar as causas é uma das forças criadoras dos documentos. Smith (2012, p. 8) explica a conexão entre as três teorias: trata-se de uma teoria para os atos dos documentos suplementando a teoria dos atos da fala com uma abordagem através das forma pelas quais, ao fazer coisas com os documentos, as pessoas podem alterar o mundo e trazer a existência novos tipos de relações de propriedade, de organizações de negocios, dentre outras.

Assim, tem-se um tipo de “partição social” em que pessoas realizam atos sociais preenchendo documentos e as aprovando através de autoridades. A teoria dos atos dos documentos, crucial para representar direitos e obrigações numa corporação, e pode ser relacionada à BFO pela ontologia dos atos dos documentos ou d-acts (Brochhausen, Almeida, Slaughter, 2013). D-acts são tipos de continuantes sociais genericamente dependentes (SGDC), os quais, por sua vez, são continuantes genericamente dependentes (GDC), representativos de entidades sociais.

Na figura 2, a seguir, são mostrados os níveis de representação discutidos: o primeiro sendo a realidade, o segundo sendo concretizado via ontologia dos d-acts e o terceiro via BFO.

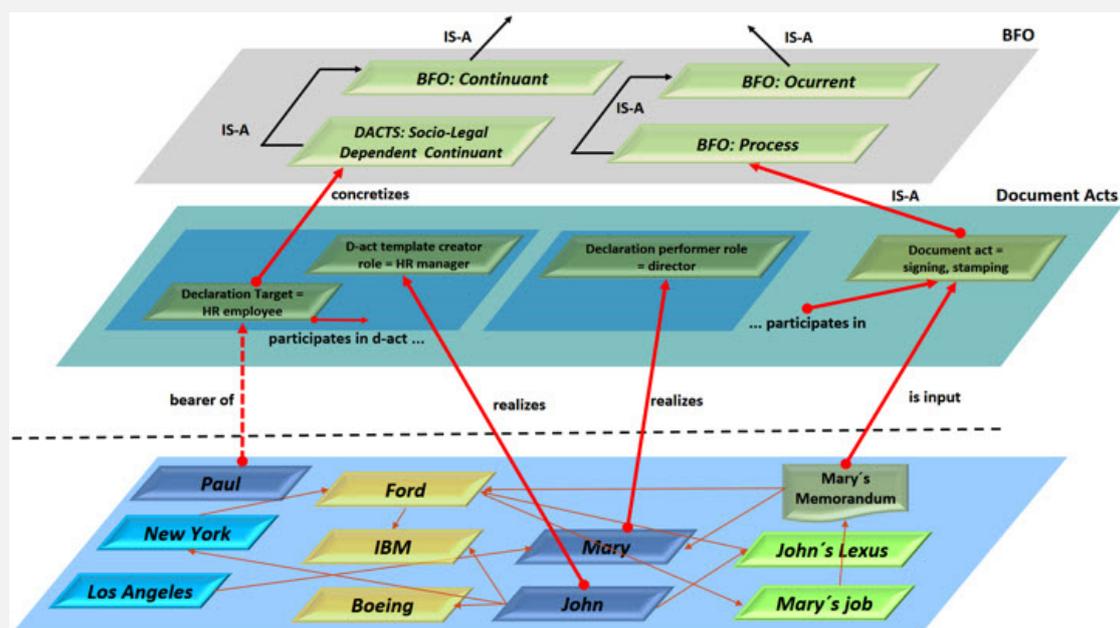


Figura 2. Níveis da realidade aos d-acts, e dos d-acts à BFO. Fonte: Almeida (2016)

Exemplos da relação entre atos de documentos e SGDCs em Ontologia de D-acts são: i) um ato de documento pode criar um SGDC, por exemplo, quando João reivindica um lote de terra; um ato de documento pode transferir um SGDC, por exemplo, quando João transfere sua reivindicação para Maria; e um ato de documento pode revogar um SGDC, por exemplo, quando um juiz assina documentos de divórcio preenchidos por um casal. Além disso, a ontologia de atos dos documentos mantém papéis e portadores específicos: o criador do documento é chamado de “função de criador de modelo de documento original”; o usuário do documento é chamado de “função de executor de declaração”, e o portador de destino do SGDC é chamado de “alvo de declaração”.

Um outro exemplo pode ilustrar a operação da ontologia dos atos dos documentos: um caso fictício de recrutamento em uma corporação. Um diretor assinando e carimbando um memorando oficial para recrutar um faxineiro é um ato de documento; o memorando oficial da diretoria permite legalmente a realização do processo de recrutamento; um memorando é um input específico de um ato de documentos com a ordem do diretor para a contratação; o gerente de recursos humanos responsável pelo processo de recrutamento é o portador do papel de criador do template do ato de documento; o diretor é o portador do papel de executor da declaração; o funcionário de recursos humanos responsável pelos procedimentos de recrutamento é o objetivo da declaração, o qual se torna representante do direito de executar procedimentos que permitem o recrutamento.

RESULTADOS

A pesquisa foi conduzida utilizando teorias das ontologias formais, atos da fala e atos dos documentos como suporte para a análise ontológica da natureza das corporações, através das dimensões descritiva e normativa. Esta análise, por sua vez, mostrou que estes dois níveis de representação mantêm correlações que permitem esquematizar uma arquitetura da informação baseada em ontologias.

Para propósitos de formalização e aplicação nos modernos sistemas de informação, vale lembrar que todos os níveis de representação descritos podem ser mapeados para a BFO, uma ontologia formal de alto nível transcategorial. A relevância desta análise reside na melhor compreensão das corporações e no *framework* formal aplicado a sistemas de informação para o contexto da web semântica.

A dimensão normativa de análise apresentada permite discussões uma vez que processos, ordens, tarefas, nomeações, recrutamentos e demissões, assim como diversas outras atividades relacionadas em corporações, ocorrem através de atos formais via atos da fala e atos dos documentos. Como a pesquisa está em andamento, nesse momento apenas levantam-se algumas especulações que carecem de verificação empírica em trabalhos futuros. Uma dessas especulações diz respeito a relevância dos documentos corporativos e respectivos atos dos documentos para a caracterização das corporações a tipos de corporações.

A hipótese é que pode-se caracterizar um tipo de corporação tanto através dos documentos que servem como *input* para os atos dos documentos, quanto através dos respectivos atos dos documentos. Pesquisaram-se algumas evidências teóricas, provenientes da filosofia, para a premissa que corporações podem ser diferenciadas pelos documentos que produzem e usam.

Ledema (2003) explica que práticas e recursos relacionados a produção e guarda de documentos revelam como se faz a organização em uma corporação. A forma como se estrutura os documentos impõem certo tipo de administração aos trabalhadores e, em última instância, cria índices de como a corporação é gerenciada. De fato, “registros são a base informacional do estado moderno e na organização moderna” (Ledema, 2003, p.151), uma vez que são os meios e os resultados de um processo contínuo de notação, sumarização, e disseminação da informação que objetiva construir uma representação do que acontece no ambiente corporativo.

Ferraris (2003) conjectura sobre a ubiquidade dos documentos em nossa sociedade através de uma teoria chamada “documentalidade”, de acordo com a qual “nada social existe fora do texto” (Ferraris, 2013, p. 318). De fato, um casamento ou um contrato que não são registrados não poderiam existir como um objeto (social), ao passo que uma montanha pode existir sem ser mapeada. Uma vez que nada existe fora do texto, a sociedade seria baseada no registro em documentos. Esse ato de registro seria assim a condição para a criação de documentos, fazendo que com os documentos se constituíssem em um ingrediente fundamental do mundo social.

Smith (2012) também coloca documentos em uma posição importante para explicar a sociedade e as relações sociais, por exemplo, aquelas que são produzidas e manipuladas dentro da corporação. Como a sociedade se torna cada vez mais complexa, “os poderes mnemônicos dos indivíduos tem sido prosteticamente estendidos através de documentos de forma a fazer surgir uma variedade de novos artefatos da realidade social” (Smith, 2012, p. 6).

Dessa forma, poder-se-ia dizer que um tipo de corporação pode ser identificado através dos tipos de atos de documentos executados. Considere-se um documento típico de um tipo de corporação, por exemplo, um registro médico no contexto de unidades de saúde. Um registro médico pode servir para: suportar o cuidado a pacientes na coordenação de processos clínicos, permitir melhor tomada de decisão e a criação de pesquisa demográficas; atender a obrigações externas para reembolso de planos de saúde, auditoria, acreditação, etc.; auxiliar na administração no planejamento, controle, e gestão da qualidade; permitir pesquisa científica e educação clínica (Haux, Knaup and Leiner, 2007). Especula-se que, ao identificar os tipos de atos dos documentos que um registro médico contém ou gera pode fornecer pistas

para formalmente distinguir uma corporação médica de outras. Como mencionado, trata-se de uma especulação em discussão, que precisa ser validada empiricamente em trabalhos futuros.

CONCLUSÕES

A dimensão normativa possibilita a discussão sobre atividades corporativas originadas pela transmissão por atos da fala e registradas por atos dos documentos. Possibilita-se, assim, trabalhos futuros a respeito da relevância dos documentos corporativos e dos atos dos documentos como caracterizadores das corporações e tipos de corporações.

Práticas e recursos que são registrados e produzem registros revelam aspectos organizacionais internos à corporação. A estrutura dos registros impõe modos administrativos, criando índices de como a corporação é gerida, pois resultam de um processo contínuo de notação, sumarização e disseminação da informação representando o que ocorre no ambiente corporativo (Ledema, 2003).

Pode-se, assim, identificar o tipo de corporação pelo tipo de ato dos documentos. A especulação desta pesquisa busca identificar o tipo de atos dos documentos que um registro contém ou gera para prover indícios sobre a distinção entre uma corporação e outra.

Como observações finais, enfatiza-se a conexão entre os níveis de representação existentes ao longo de nossa análise ontológica. O nível de realidade se conecta com dois níveis de representação, ou seja, um nível para partições e o nível de atos de documentos. Como se pode perceber, esses dois níveis de representação, além de descrever as dimensões da análise proposta, mantêm co-relações que, uma vez reunidas, podem oferecer uma visão de qual seria uma arquitetura de informação baseada em ontologias para corporações. Para fins de formalização e aplicação nos modernos sistemas de informação, vale a pena lembrar que todos os níveis de representação descritos podem ser mapeados para BFO, uma ontologia de nível superior formal e transcategorial.

REFERÊNCIAS BIBLIOGRÁFICAS

Almeida, M. B., & Barbosa, R. R. (2009). Ontologies in knowledge management support: A case study. Retrieved February 1, 2010, from https://www.researchgate.net/publication/220434254_Ontologies_in_knowledge_management_support_A_case_study

Almeida, M. B. (2016). An ontological study of healthcare corporations and their social entities. In: ICBO workshop. Retrieved March 21, 2017, from http://mba.eci.ufmg.br/?page_id=84

Arp, R., Smith, B., & Spear, A. D. (2015). *Building ontologies with basic formal ontology*. Cambridge: MIT Press.

Bernus, P., Nemes, L., & Williams, T. J. (Eds.). (1996). *Architectures for enterprise integration*. Springer Science & Business Media.

Bittner, T., & Smith, B. (2003). A theory of granular partitions. Retrieved August 27, 2010, from <http://ontology.buffalo.edu/smith/articles/partitions.pdf>.

Brochhausen, M., Almeida, M. A., & Slaughter, L. (2013). *Towards a formal representation of document acts and the resulting legal entities. Johanssonian Investigations*. Frankfurt: Ontos, 120-39.

Ferraris, M. (2013). *Documentality: Why It Is Necessary to Leave Traces*. New York: Fordham University Press.

Fillion, F., Menzel, C., Blinn, T., & Mayer, R. (1995). An ontology-based environment for enterprise model integration. Retrieved August 22, 2010, from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.162.122&rep=rep1&type=pdf>.

Fonseca, F. (2007). The double role of ontologies in information science research. Retrieved March 12, 2010, from <https://pdfs.semanticscholar.org/3688/4289fbcce8994f5c557d7fd68c0068acd0b3.pdf>.

Fox, M. S. (1992). The tove project towards a common-sense model of the enterprise. Retrieved March 12, 2010, from <https://link.springer.com/chapter/10.1007/BFb0024952>.

Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., & Schneider, L. (2002). Sweetening ontologies with DOLCE. Retrieved January 15, 2011, from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.11.6038&rep=rep1&type=pdf>.

Grenon, P., Smith, B., & Goldberg, L. (2004). Biodynamic ontology: applying BFO in the biomedical domain. Retrieved January 12, 2009, from <https://www.ncbi.nlm.nih.gov/pubmed/15853262>.

Guizzardi, G. (2005). *Ontological foundations for structural conceptual models*. Amsterdam: Centre for Telematics and Information Technology.

Ledema, R. (2003). *Discourses of post-bureaucratic organization* (Vol. 5). New York: John Benjamins Publishing.

Obrst, L., Hughes, T., & Ray, S. (2006). Prospects and possibilities for ontology evaluation: The view from NCOR. Retrieved September 24, 2009, from <http://ceur-ws.org/Vol-179/eon2006obrstetal.pdf>

Searle, J. R. (1976). A classification of illocutionary acts. Retrieved October 9, 2007, from https://sites.duke.edu/conversions/files/2014/09/Searle_Illocutionary-Acts.pdf.

Searle, J. R. (1984). *Minds, brains and science*. Cambridge: Harvard University Press.

Simons, P. M. (1987). *Parts: A study in ontology*. Oxford: Oxford University Press.

Smith, B. (1997). On substances, accidents and universals: In defence of a constituent ontology. Retrieved April 14, 2007, from <http://ontology.buffalo.edu/smith/articles/greensboro.pdf>.

Smith B. (2003). Ontology and Information Systems. Retrieved December 30, 2013, from [http://www.ontology.buffalo.edu/ontology\(PIC\).pdf](http://www.ontology.buffalo.edu/ontology(PIC).pdf)

Smith, B. (2012). How to do things with documents. Retrieved March 10, 2013, from <http://ontology.buffalo.edu/smith/articles/howtodothingswithdocuments.pdf>

Uschold, M., King, M., Moralee, S., & Zorgios, Y. (1998). The enterprise ontology. Retrieved April 2, 2008, from <http://www.aiai.ed.ac.uk/project/enterprise/ontology.html>.

Varzi, A. C. (1996). Parts, wholes, and part-whole relations: The prospects of mereotopology. Retrieved January 13, 2007, from <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=2EEC83238E4F9EB21B000686FF7306B3?doi=10.1.1.41.7596&rep=rep1&type=pdf>



DIMENSÕES DAS PRÁTICAS DE AUTORIA EM ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO NO BRASIL

Murilo Artur Araújo da Silveira¹, Sônia Elisa Caregnato²

¹Universidade Federal de Pernambuco, 0000-0002-9708-6001, muriloas@gmail.com,

²Universidade Federal do Rio Grande do Sul, 0000-0002-5676-2763, sonia.caregnato@ufrgs.br

RESUMO Análise das dimensões das autorias das publicações dos pesquisadores brasileiros em Organização e Representação do Conhecimento nas contribuições científicas certificadas. Tem como objetivo mapear e analisar as práticas de autoria nas publicações certificadas relativas ao interesse temático do domínio da Organização e Representação do Conhecimento no Brasil de 2011 a 2014. Utiliza o método bibliométrico para mapeamento, organização dos dados e análise dos resultados, com base nos 315 artigos de periódicos e dos anais do Encontro Nacional de Pesquisa e Pós-graduação em Ciência da Informação. Os principais resultados são: a) o núcleo de autores mais produtivos é constituído por um seletivo grupo, com trajetória científica consolidada; e b) a produção científica do núcleo de autores mais produtivos é sustentada pela convergência de fatores relativos às orientações na pós-graduação e às posições alcançadas em seu percurso no domínio. Os resultados também indicam uma relação possível entre os índices de produtividade e as posições alcançadas no campo pelos pesquisadores, sobretudo os mais experientes.

PALAVRAS-CHAVE *Organização e Representação do Conhecimento, Produção Científica, Autoria, Brasil.*

ABSTRACT Analysis of the dimensions of authorship in publications of the Brazilian researchers in Knowledge Organization and Representation in the certified scientific contributions. It aims to map and analyze the authorship practices in certified publications related to the thematic interest of the domain of Knowledge Organization and Representation in Brazil from 2011 to 2014. It uses the bibliometric method for mapping, organizing data and analyzing the results, based on the 315 articles from journals and annals of the National Conference on Information Science Research and Post-graduation. The main results are: a) the core of more productive authors is constituted by a select group, with consolidated scientific trajectory; and b) the scientific production of the more productive authors is supported by the convergence of factors related to supervision of postgraduate students and the positions reached in its course in the field. The results also indicate a possible relationship between the productivity indexes and the positions reached in the field by the researchers, especially the more experienced ones.

KEYWORDS *Knowledge Organization and Representation, Scientific Production, Authorship, Brazil.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

O volume de pesquisas em Organização e Representação do Conhecimento (ORC) no Brasil, sobretudo na última década, vem acompanhado de um conjunto de elementos e condições que viabilizaram e possibilitaram o aumento considerável da produção e comunicação. Os elementos e as circunstâncias que se destacam para a formação do cenário brasileiro do domínio da ORC são: a) a criação do capítulo brasileiro da *International Society of Knowledge Organization* (ISKO-Brasil) em 2007; b) a realização de três conferências bianuais da ISKO-Brasil (2011, 2013 e 2015); c) o crescimento do número de programas de pós-graduação em Ciência da Informação (CI); e d) a ampliação de recursos humanos qualificados para a condução de pesquisas (Arboit, 2014; Martins, 2014). Além disso, o aumento de títulos de periódicos e de eventos também influenciou no volume de publicações certificadas dos autores brasileiros (Gabriel Júnior, 2014).

Os pontos destacados que favoreceram o crescimento quantitativo e qualitativo das investigações no domínio da ORC no Brasil apontam um quadro expressivo de investimentos de diversas frentes que necessitam ser estudados para obter um quadro fiel e representativo dos percursos trilhados. Dentre os elementos pertinentes das bases epistemológicas de compreensão da dinâmica investigativa do domínio no Brasil, o aspecto autoral é, sem dúvida, um dos fenômenos de maior relevância dos últimos anos, tendo em vista as diretrizes da Coordenação de Aperfeiçoamento de Pessoal do Ensino Superior (CAPES) sobre a matéria. O Documento de Área 2013 (Ciências Sociais Aplicadas I), ao qual o domínio da ORC está vinculado, estimula a produção intelectual de docentes e discentes em regime de coautoria interna e externa ao curso de pós-graduação, com objetivos diversos que vão desde a iniciação ao universo da pesquisa e publicação até o incremento dos índices de produção científica (CAPES, 2013).

Diante do cenário descrito, em que as políticas científicas para a pós-graduação e a produção do conhecimento incidem na dinâmica dos campos e domínios no Brasil, investigar a autoria e suas circunstâncias torna-se oportuno para identificar como tais recomendações repercutiram nas práticas dos pesquisadores. Para tanto, considera-se que as práticas de autoria podem se manifestar em distintas dimensões, as quais se destacam: autoria individual e/ou autoria coletiva; colaboração intrainstitucional e colaboração interinstituições.

Face ao panorama descrito, os problemas desta pesquisa são: como se configura a autoria no domínio da Organização e Representação do Conhecimento no Brasil por meio da produção científica certificada de 2011 a 2014? Em que medida o Documento de Área 2013 influenciou as práticas de autoria da produção relativa ao domínio de Organização e Representação do Conhecimento? Tais questões de pesquisa se inserem nos estudos de institucionalização da atividade científica (Whitley, 1974; 1980) por meio do seguinte objetivo: mapear e analisar as práticas de autoria nas publicações certificadas relativas ao interesse temático do domínio da Organização e Representação do Conhecimento no Brasil de 2011 a 2014.

Devido ao crescimento da produção científica brasileira nos últimos anos, o domínio da ORC necessita de acompanhamento diacrônico e sincrônico de seus produtos certificados, em uma perspectiva crítica que contemple aspectos poucos explorados. Parte-se da concepção de que estudar o domínio por meio das práticas de autoria permitirá visualizar o comportamento e os efeitos de tais atos no plano coletivo, mergulhados no contexto sociocultural do fazer científico. Considera-se que essas práticas de citação e seus desdobramentos participam da lógica produtiva da edição científica (Bolaño, Kobashi & Santos, 2006), sustentados pela integração de aspectos objetivos e subjetivos, forjados por boas e más condutas

e justificados pela manutenção de posições alcançadas (Bourdieu, 2004; 2013; Erikson & Erlandson, 2014).

O recorte temporal estabelecido pela pesquisa advém dos indicadores de produção científica alcançados pelo domínio a ser investigado no âmbito da pesquisa e pós-graduação em Ciência da Informação no país, conforme expressam Arboit (2014) e Martins (2014). O nível de crescimento das atividades de pesquisa e das formas de disseminação do conhecimento produzido decorre da ampliação dos programas de pós-graduação que se inicia no final dos anos 2000 no país, segundo o Documento de Área 2013, além dos outros tópicos evidenciados. Nesta perspectiva de análise, visualizou-se um incremento de projetos e linhas de pesquisas em ORC nos cursos de pós-graduação, um aumento significativo do número de pesquisadores em todos os níveis de formação e, ainda, um acréscimo qualitativo de títulos de periódicos e eventos científicos. Do panorama evidenciado, a análise dos produtos científicos disseminados nos veículos qualificados do período de 2011 a 2014 possibilitará: a) verificar e compreender os caminhos de pesquisa e publicação percorridos; e b) refletir sobre os elementos sobre autoria que incidiram e impactaram as formas de produção e comunicação.

METODOLOGIA

A pesquisa, de caráter descritiva e bibliográfica, se apoiou nas técnicas bibliométricas para coleta, organização (limpeza e sistematização de dados) e análise da produção científica brasileira no domínio no período de 2011 a 2014. O *corpus* da pesquisa é formado por 315 publicações, sendo 162 trabalhos publicados nos periódicos brasileiros em Ciência da Informação e indexados na Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI) e 153 trabalhos publicados nos anais eletrônicos do GT 2 (Organização e Representação do Conhecimento) do Encontro Nacional de Pesquisa e Pós-graduação em Ciência da Informação (ENANCIB) contemplados na coleção do repositório BENANCIB. Juntas, as duas bases de dados contemplam o núcleo principal das contribuições científicas certificadas dos pesquisadores brasileiros do campo da CI e do domínio da ORC.

A seleção e coleta dos artigos na BRAPCI foram planejadas com base em uma estratégia de busca que contemplasse os principais termos vinculados às temáticas do domínio. Contudo, percebeu-se que a estratégia não contemplaria temas e assuntos emergentes ou com outras descrições temáticas atribuídas pelos autores. Desta feita, optou-se por percorrer todos os sumários dos periódicos indexados pela base no período estabelecido, coletando e organizando os artigos de autores brasileiros com base nos seguintes atributos: autoria, primeiro autor, colaboradores, título do artigo, resumo e palavras-chave; não excetuando, em alguns momentos, a leitura da introdução e outras partes do texto.

Para a identificação e captura dos trabalhos referentes às comunicações orais e aos pôsteres do ENANCIB de 2011 a 2014 na coleção BENANCIB, procedeu-se a busca por meio de dois campos que recuperaram todas as contribuições: GT (grupos de trabalho) e ANO (ano de apresentação). Após a coleta dos arquivos, os mesmos foram organizados com referência aos artigos de autores nacionais e modalidades.

Nas etapas de coleta e organização dos dados em uma planilha, em seus respectivos campos, algumas ações compreendidas tiveram que ser contornadas frente aos inconvenientes decorrentes das variações nas indicações de autoria e vinculação institucional e das inconsistências dos resumos e das palavras-

chave, além de erros de digitação e indicações equivocadas da autoria dos artigos citados. Após o término da coleta e organização dos dados e informações, a planilha preenchida foi introduzida no software *Vantage Point* para a padronização das entradas dos autores e seus respectivos colaboradores, instituições e vínculos entre autores para a geração dos indicadores.

O percurso de pesquisa trilhado está detalhado em cinco etapas, descritas a seguir.

- Etapa 1: Definição dos campos e suas estruturas para a inclusão dos dados e informações referentes ao enfoque bibliométrico em uma planilha elaborada no software *Microsoft Excel*;
- Etapa 2: Correção de dados e informações na planilha elaborada segundo as indicações encontradas (erradas e equivocadas) no conjunto de artigos relacionados ao enfoque quantitativo;
- Etapa 3: Migração dos dados e das informações corrigidas da planilha elaborada para o software *Vantage Point*;
- Etapa 4: Sistematização dos dados e das informações coletadas no *template* de análise do software; e
- Etapa 5: Elaboração de quadros, matrizes e representações gráficas com base nos elementos contemplados pela relação entre autoria e coautoria.

RESULTADOS

Para a análise dos 315 artigos que compõem o *corpus* da pesquisa, recorre-se ao Gráfico 1 que determina as primeiras posições entre o conjunto de 359 autores identificados.

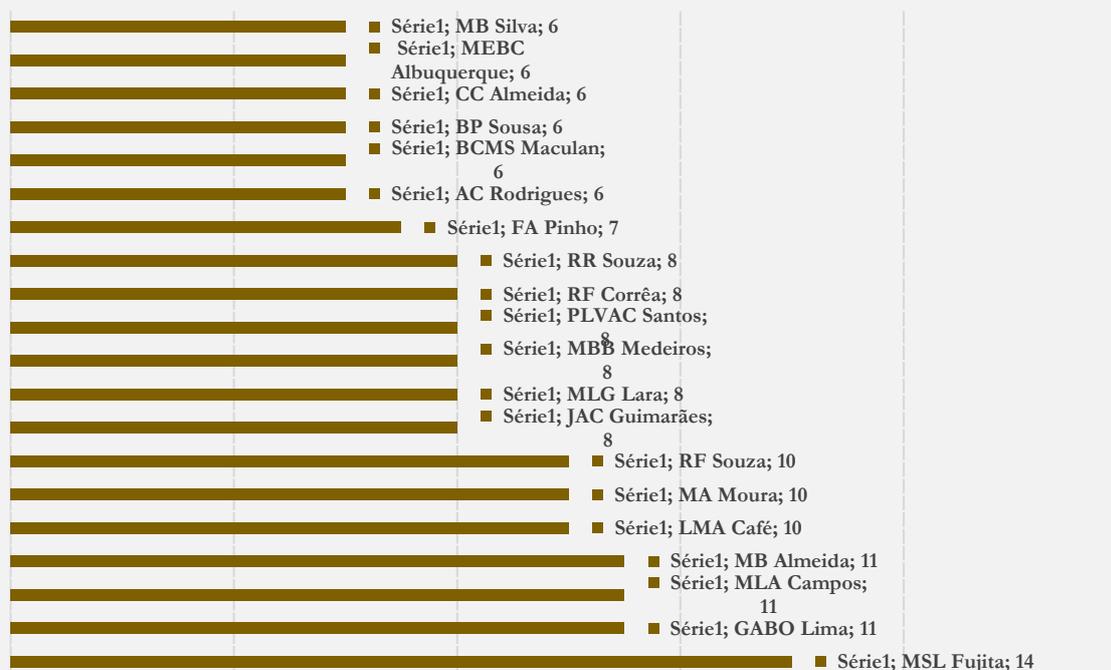


Gráfico 1: Distribuição das Autorias dos Artigos sobre ORC no Brasil (2011-2014)

Fonte: Dados da Pesquisa, 2016.

Os vinte autores mostrados formam o grupo dos mais produtivos no período por meio do recorte documental proposto pela pesquisa. Os números apontados se referem à autoria principal e às colaborações dos pesquisadores destacados, dentro do conjunto de 359 autores identificados nas 315 contribuições analisadas. Sobre a inserção desses autores no domínio, pode-se constatar que mais de 70% deles têm formação no campo da Ciência da Informação em nível de doutorado e mais de 80% tem formação em Biblioteconomia em nível de graduação. Todos os autores presentes na representação gráfica acima estão vinculados às universidades públicas do país, além de pertencerem aos principais programas de pós-graduação em Ciência da Informação. Outro ponto a se destacar é que há uma predominância significativa de autores experientes, quando se observa a trajetória e as contribuições dos mesmos, em um recorte temporal de mais de dez anos. Do conjunto mostrado, confirma-se que oito autores pertencem ao seletivo grupo de pesquisadores produtivos que recebem bolsa do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) para realização de pesquisa de alto nível, formação de recursos humanos, publicação científica regular e qualificada, entre outras exigências.

Em um estudo similar sobre a questão da autoria no domínio da Organização e Representação do Conhecimento no Brasil de 2003 a 2012, Bufrem, Freitas e Nascimento (2014) destacam cinco autores que também estão presentes no Gráfico 1. Ao comparar a lista de autores e os recortes temporais e documentais entre esta pesquisa e a mencionada, nota-se que a proporção de convergência entre a presença de autores nos dois estudos demonstra que há um grupo de pesquisadores com larga experiência no domínio.

Com o propósito de visualizar as circunstâncias que condicionam as práticas de autoria no domínio no país, o Gráfico 2 aponta os dezesseis autores com os maiores índices de ocorrência de coautoria, do conjunto de 188 coautores identificados, nos 315 artigos analisados.

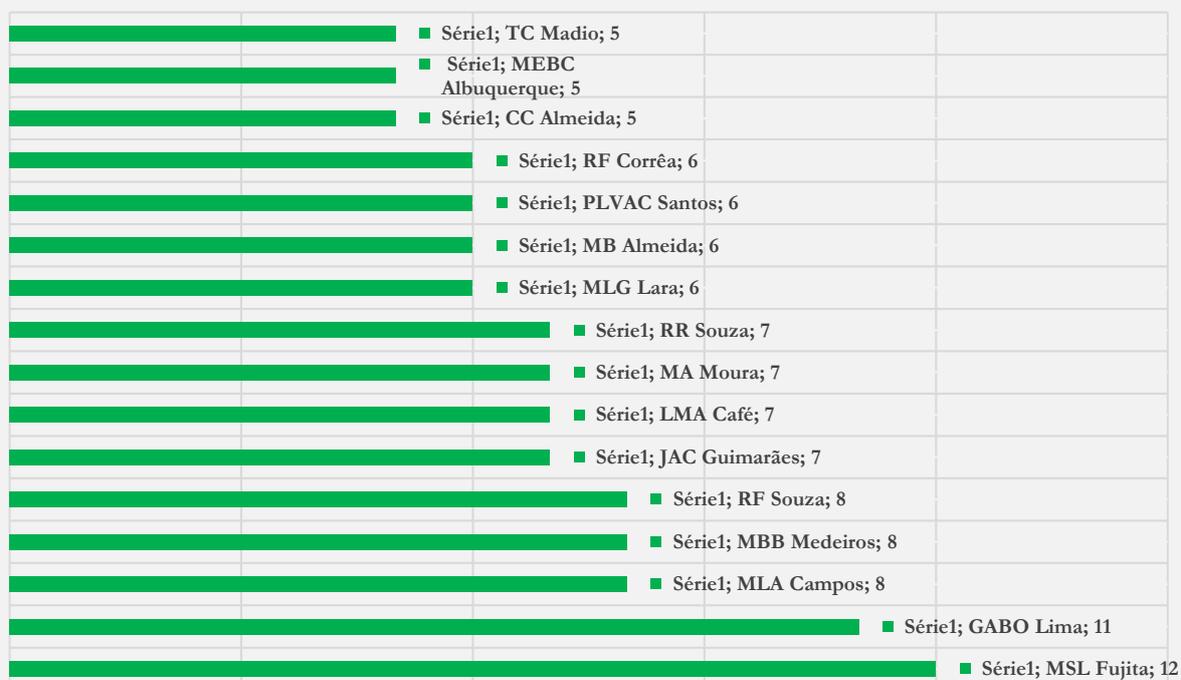


Gráfico 2: Distribuição dos Coautores dos Artigos sobre ORC no Brasil (2011-2014)

Fonte: Dados da Pesquisa, 2016.

O gráfico indica a posição dos autores em relação à assinatura de artigos na condição de coautores advinda das múltiplas relações com os primeiros autores no período. A representação acima destaca que, dos dezesseis autores presentes, apenas um dentre todos não é visualizado no Gráfico 1: TC Madio. Ao comparar as posições dos quinze autores em comum nos Gráficos 1 e 2, percebe-se que há variações, ficando claro que, no recorte desta pesquisa, a produtividade deles está vinculada a relações colaborativas.

As relações de coautoria visualizadas no gráfico acima, em sua maioria, são estabelecidas pelas atividades de orientação de mestrado e doutorado, amplamente difundidas e estimuladas pelas instâncias e instituições ligadas à pós-graduação no Brasil, conforme se verifica no Documento de Área 2013 (CAPES). A leitura dos dois gráficos destaca que os índices de produção alcançados por dois autores no período foram sustentados, exclusivamente, por sua condição de coautor: MBB Medeiros e GABO Lima. Aliado a isto, verifica-se que quatro coautores visualizados na representação gráfica acima desempenham atividades junto às instâncias governamentais e instituições científicas brasileiras que o domínio está vinculado: MSL Fujita (Comitê de Assessoramento do CNPq), MBB Medeiros (Coordenadora Adjunta da Área de Ciências Sociais Aplicadas I da CAPES), JAC Guimarães e RR Souza (Membros da Diretoria da ISKO Brasil).

Para complementar a descrição realizada, traz-se à discussão o Quadro 1 que discrimina os atores e as posições alcançadas por eles nas diversas instâncias de poder relacionadas diretamente com o domínio no período coberto pela pesquisa.

Quadro 1: Posição dos Autores nas Instâncias de Poder

Instituições	Pesquisadores	Posições Assumidas
CAPES	Johanna Wilhelmina Smit (USP) Marisa Bräscher Basílio Medeiros (UFSC) Nair Yumiko Kobashi (USP) Vera Lúcia Doyle Louzada de Mattos Dodebei (UNIRIO)	Coordenadoras Adjuntas da Área de Ciências Sociais Aplicadas I
CNPq	José Augusto Chaves Guimarães (UNESP) Mariângela Lopes Spotti Fujita (UNESP)	Membro do Comitê de Assessoramento das Áreas de Artes, Ciência da Informação e Comunicação
ISKO-Brasil	José Augusto Chaves Guimarães (UNESP) Fabio Assis Pinho (UFPE) Gercina Ângela Borém de Oliveira Lima (UFMG) Renato Rocha Souza (UFMG) Vera Lúcia Doyle Louzada de Mattos Dodebei (UNIRIO)	Membros da Diretoria da Entidade Científica

Fonte: Site das Instituições (CAPES, CNPq e ISKO-Brasil), 2016.

A discussão empreendida até agora sobre os atores e as posições alcançadas por eles, como também os índices de produção científica no domínio, em especial as práticas de autoria, revelam o papel de liderança e o reconhecimento do trabalho desenvolvido. Percebe-se que apenas seis autores figuram no Quadro 1 e no Gráfico 1, os quais estão destacados em negrito: José Augusto Chaves Guimarães (em duas instâncias), Marisa Bräscher Basílio Medeiros, Mariângela Lopes Spotti Fujita, Fabio Assis Pinho, Gercina Ângela Borém de Oliveira Lima e Renato Rocha Souza (uma instância). Nota-se ainda que dos seis autores em destaque, quatro são bolsistas de produtividade em pesquisa do CNPq e estão vinculados aos programas de pós-graduação mais bem qualificados do país, segundo a categorização da CAPES no triênio 2010-2012.

Todavia, pondera-se que a trajetória percorrida não pode ser relegada apenas para o exercício da crítica indiscriminada. Tais constatações e reflexões frente a uma realidade concreta são necessárias porque se tipificam as circunstâncias e se determinam as condições que resultam em uma conjuntura momentânea do período investigado, fornecendo outros elementos e pontos de vista a serem considerados para o desenho de quadro sinóptico das relações estabelecidas no domínio. Além disso, fica evidente que a permanência e o revezamento nas instâncias de poder destes e de outros atores científicos possibilitaram o crescimento qualitativo dos elementos que compõem o domínio. No entanto, a perpetuação pode expressar o confronto entre os capitais científico e político identificados por Bourdieu (2004), que cristalizam ações de renovação e ampliação do escopo do domínio e naturalizam práticas e condutas científicas.

O cenário apontado sobre as dimensões da autoria no período de 2011 a 2014 no domínio da Organização e Representação do Conhecimento no Brasil destaca as formas utilizadas pelos autores para publicação nos veículos reconhecidos e institucionalizados e alcance dos índices de produtividade aos quais são submetidos. Nesse espectro de análise, questiona-se:

- a) em que medida os índices de produtividade alcançados por esses autores refletem suas trajetórias e seus interesses de pesquisa?
- b) as instâncias e instituições brasileiras influenciam, de fato, nas formas de produção científica dos autores vinculados ao domínio?

Diante das discussões realizadas que se ampararam nos dados obtidos do conjunto de 315 artigos de periódicos e do ENANCIB de 2011 a 2014 relativos ao domínio da Organização e Representação do Conhecimento no Brasil, verifica-se que as práticas de autoria estão pautadas nos seguintes fenômenos:

- a) o núcleo de autores mais produtivos é constituído por um seletivo grupo, com trajetória científica consolidada (tempo e experiência de pesquisa);
- b) a produção científica do núcleo de autores mais produtivos é sustentada pela convergência de fatores relativos às orientações na pós-graduação e às posições alcançadas em seu percurso.

É necessário considerar que o crescimento quantitativo da produção nos últimos anos também resulta de esforços na formação de novos recursos humanos que, necessariamente, passa pela produção e comunicação do conhecimento. Todavia, verifica-se que a política científica nacional, que se direciona ao domínio, amplia e redimensiona as formas de produção colaborativa entre os pares, sobretudo nos níveis hierárquicos de relação científica, oriundos do vínculo entre os orientadores e seus orientandos. Outro ponto que repercute diretamente nos índices de produtividade visualizados é a convergência temática entre as pesquisas dos orientadores e a de seus orientandos, na medida em que tais relações conduzem as práticas de pesquisa no âmbito da organização dos programas de pós-graduação no Brasil.

Tais fenômenos percebidos pelos resultados analisados e discutidos endossam as ideias lançadas por Bourdieu (1983; 2013) sobre capital científico e seus desdobramentos nas atividades de produção e comunicação do conhecimento, envoltos por elementos simbólicos que conferem distinção e perpetuam práticas de reprodução. Além disso, a existência desses fenômenos que são amplamente justificados pelos praticantes, os aspectos condicionantes de manutenção e preservação cultural que regem as práticas e seus produtos também necessitam de legitimação social (Eagleton, 2011). O cenário visualizado por meio dos resultados, então, expressa essa lógica, em que história e trajetória são sempre evocadas para fundamentar e esclarecer como os pilares sociais construídos se mantiveram e explicitar

e consolidar as regras e os mecanismos de controle utilizados (Fleury, 2009; Bourdieu, 2011). Nesses termos, ressalta-se que as mais diversas práticas desenvolvidas afetam e são afetadas pelas diferentes forças, sendo os agentes sociais determinantes para que se concretizem. Dessa relação resultam processos e produtos que explicitam competições, cooperações, conflitos e desigualdades, desdobrando-se na legitimação de atores, domínios e práticas, além de instituir uma cultura científica baseada nas relações de trocas.

CONCLUSÕES

O recorte temporal da pesquisa revela que as práticas de autoria e os índices de produtividade no domínio da ORC são sustentadas e mantidas pelas ocorrências de coautoria. Os resultados no período enfatizam que os pesquisadores mais experientes têm produção significativa decorrente das atividades de orientação na pós-graduação, em regime de colaboração científica. Além disso, percebe-se que há uma forte influência das agências reguladoras da pós-graduação brasileira sobre o quadro de autoria, coautoria e produtividade no domínio da ORC. Essa influência normaliza e regulamenta um conjunto de práticas colaborativas que intensificam a produção científica e a cooperação entre os atores mais e menos experientes. Desse modo, constata-se que o crescimento e aumento dos quadros de cursos, recursos humanos e veículos de disseminação qualificados, bem como as diretivas institucionais da agência reguladora da pós-graduação no Brasil, influenciaram para o alcance do cenário encontrado no que se refere às práticas de autoria e o volume de produção científica publicado.

Os resultados também indicam uma relação possível entre os índices de produtividade e as posições alcançadas no campo pelos pesquisadores, sobretudo os mais experientes, configurando-se em um problema de pesquisa passível de investigação futura. Dessa forma, este trabalho sinaliza um conjunto de possibilidades de investigações futuras, dos pontos de vista teórico e metodológico, que se materializa em um conjunto de produtos científicos, interpretados à luz de construtos epistemológicos dialéticos que problematizam e questionem a lógica que integram as relações de produção e comunicação de conhecimentos. Registra-se também a limitação para a generalização dos resultados obtidos quando se considera o recorte de tempo desta pesquisa, em que se apresentam um cenário parcial da realidade brasileira no domínio analisado.

REFERÊNCIAS BIBLIOGRÁFICAS

- Arboit, A. E. (2014). *O processo de institucionalização sociocognitiva do domínio de Organização do Conhecimento a partir dos trabalhos científicos dos congressos da ISKO*. (Tese de Doutorado) - Universidade Estadual Paulista, Marília, SP.
- Bolaño, C. R. S., Kobashi, N. Y. & Santos, R. N. M. (2006). A lógica econômica da edição científica certificada. *Encontros Bibli*, n. esp., 119-131.
- Bourdieu, P. (1983). O campo científico. In Ortiz, R. (Org.). *Pierre Bourdieu* (cap. 4, pp 122-155). São Paulo: Ática.
- Bourdieu, P. (2011). *A distinção: crítica social do julgamento*. São Paulo: EDUSP; Porto Alegre: Zouk.
- Bourdieu, P. (2013). *Homo academicus*. 2. ed. Florianópolis: Ed. Universitária da UFSC.

Bourdieu, P. (2004). *Os usos sociais da ciência: por uma sociologia clínica do campo científico*. São Paulo: Ed. Universitária da UNESP.

Bufrem, L. S., Freitas, J. L. & Nascimento, B. S. (2014). Autoria e pesquisa em Organização e Representação do Conhecimento: análise da produção científica em Ciência da Informação. *Em Questão*, 20(3), 150-165.

Coordenação de Aperfeiçoamento de Pessoal do Ensino Superior (2013). *Documento de Área 2013: Ciências Sociais Aplicadas 1*. Brasília: CAPES.

Eagleton, T. (2011). *A ideia de cultura*. 2. ed. São Paulo: Ed. Universitária da UNESP.

Erikson, M. G. & Erlandson, P. (2014). A taxonomy of motives to city. *Social Studies of Science*, 44(1), 1-13.

Fleury, L. (2009). *Sociologia da cultura: e das práticas culturais*. São Paulo: Ed. Senac.

Gabriel Junior, R. F. (2014). *Geração de indicadores de produção e citação científica em revistas de Ciência da Informação: estudo aplicado à base de dados BRAPCI* (Tese de Doutorado) - Universidade Estadual Paulista, Marília, SP.

Martins, G. K. (2014). *Institucionalização cognitiva e social da Organização e Representação do Conhecimento na Ciência da Informação no Brasil* (Tese de Doutorado) – Universidade Estadual Paulista, Marília, SP.

Whitley, R. (1974). Cognitive and social institutionalization of scientific specialities and research areas. In: Whitley, R. (Ed.). *Social processes of scientific development* (pp. 69-95). London: Routledge and Kegan.

Whitley, R. (1980). The context of scientific investigation. In: Knorr, K. D., Krohn, R. & Whitley, R. *The social process of scientific investigation* (pp. 297-321). London: D. Reidel.



ELEMENTOS DE INDEXAÇÃO PARA ARQUIVOS PERMANENTES: APROXIMAÇÕES TEÓRICAS PRELIMINARES

Natália Bolfarini Tognoli¹, Suellen Oliveira Milani², José Augusto Chaves Guimarães³

^{1,2,3}Universidade Estadual Paulista Júlio de Mesquita Filho – Marília – SP – Brasil –
nataliabtognoli@marilia.unesp.br, suellenmilani@hotmail.com, guima@marilia.unesp.br

RESUMO Esse estudo apresenta algumas convergências metodológicas percebidas entre a indexação na Biblioteconomia e na Arquivística e, para tanto, elementos referentes à análise e representação de assunto foram examinados. Essa reflexão exploratória demonstrou que as discussões teóricas promovidas pela organização do conhecimento nutrem ambas as áreas e são complementares pois enquanto a Arquivística organiza prioritariamente para testemunhar a produção documental a partir do princípio da proveniência, a Biblioteconomia organiza para promover o acesso e a apropriação da informação a partir dos princípios que envolvem o uso ou a demanda das suas comunidades de usuários. Por fim, exemplos da instrumentalidade da análise de assunto, da leitura documental, bem como do conhecimento da tipologia documental e do princípio da proveniência foram expostos.

PALAVRAS-CHAVE *Representação arquivística, Indexação, Análise de assunto, Representação de assunto, Arquivos permanentes*

ABSTRACT This paper presents some methodological convergences between indexing in Library and Archival Science. For that, elements related to the subject analysis and subject representation were addressed. This exploratory discussion demonstrates that the theoretical discussions promoted by the knowledge organization support both areas and they are complementary because once Archival Science organizes information primarily to represent the records production from the principle of provenance, Library Science represents information to promote their access and appropriation from principles involving the use and the demand of its communities of users. Thus, examples regarding the instrumentality of the subject analysis, documental reading, as well as, the documental typology and the principle of provenance were highlighted.

PALAVRAS-CHAVE *Archival description, Indexing, Subject analysis, Subject representation, Archives*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A organização do conhecimento é um marco teórico-conceitual para a Ciência da Informação, isto é, um espaço institucionalizado dentro da área de Ciência da Informação para discutir teorias e metodologias sobre os processos que visam *representar para recuperar* ou, como prefere Smit (1987), *reunir e organizar para achar*. Para tanto, catálogos de bibliotecas e instrumentos de pesquisa de

arquivos serão construídos para que os usuários acessem e se apropriem das informações contidas nos documentos dos acervos da biblioteca e dos fundos do arquivo.

No contexto da organização do conhecimento, enquanto o bibliotecário vale-se dos critérios de uso e especificidade temática para tratar documentos múltiplos advindos de um processo de seleção e com uma função investigativa, o arquivista vale-se dos princípios da proveniência e da organicidade para tratar documentos únicos oriundos de uma acumulação natural e com uma função probatória (Guimarães, 2008).

Para a realização dos processos de organização do conhecimento é imprescindível que se identifique o conteúdo do documento que será representado e, conforme Guimarães (2008) esclarece, na Arquivística o conteúdo encontra-se na designação do órgão produtor, da espécie e da tipologia documental, enquanto que na Biblioteconomia o conteúdo refere-se aos assuntos do documento.

Com a aproximação das áreas de Arquivística e Biblioteconomia dada, principalmente, pelo amparo da Ciência da Informação, tornou-se necessário traçar suas convergências e divergências, principalmente, no contexto da organização e recuperação da informação. Apesar de ser um tema ainda incipiente na literatura nacional, algumas tentativas recentes foram observadas nos estudos de Vital (2017), Barros (2016), Silva, Sousa e Bandeira (2012) e Ribeiro (2006).

Nesse sentido, esse estudo tem o objetivo de discutir as convergências metodológicas entre a indexação na Arquivística e na Biblioteconomia, bem como os possíveis subsídios que poderiam ser fornecidos para a prática de análise e representação em ambas as áreas. Para tanto, a reflexão teórica e de caráter documental proposta configura-se como exploratória uma vez que visa proporcionar maior familiaridade com o tema e se valerá de uma abordagem qualitativa.

INDEXAÇÃO NA BIBLIOTECONOMIA

Considerando que o principal objetivo da Biblioteconomia é promover o acesso e apropriação da informação criando verdadeiras pontes entre as necessidades informacionais das comunidades de usuários da biblioteca e os documentos do seu acervo, os processos de organização do conhecimento são mediadores imprescindíveis. Esses processos de organização do conhecimento promovem o controle bibliográfico, isto é, “[...] o processo sistemático de criação de substitutos que ocupem o lugar dos itens que de fato contêm informação” (Olson, 2002, p. 100). O propósito do controle bibliográfico ou controle de autoridade é:

[...] assegurar consistência ao representar um valor – o nome de uma pessoa, o nome de um lugar ou um termo ou código representando um assunto – nos elementos usados como pontos de acesso na recuperação da informação. Por exemplo, “Guerra Mundial, 1939-1945” foi estabelecido como um cabeçalho de assunto [também denominado termo de indexação ou descritor] autorizado na Lista de Cabeçalhos de Assunto da Biblioteca do Congresso (LCSH). Quando usamos a LCSH para catalogação ou indexação, é atribuído esse cabeçalho a todas as publicações sobre a II Guerra Mundial independentemente se a publicação se refere a “Guerra Europeia, 1939-1945”, “Segunda Guerra Mundial”, “Guerra Mundial 2”, “Guerra Mundial II”, “WWII”, “Guerra Mundial Dois”, ou 2º Guerra Mundial”. As expressões sinônimas levam ao cabeçalho autorizado. Isso assegura que todas as publicações sobre a II Guerra Mundial possam ser recuperadas e dispostas sob o mesmo cabeçalho de assunto em catálogos locais, bases de dados ou catálogo coletivo (International..., 2010, p. 8).

Diferentemente da representação descritiva de um documento com fins de representação e individualização garantidos pela catalogação, a representação de assunto refere-se ao conjunto de processos, instrumentos e produtos destinados a representar tematicamente o documento. Para tanto, os processos de indexação e classificação valem-se de instrumentos, tais como esquemas de classificação, lista de cabeçalhos de assunto, tesouros etc. para construir os índices, resumos e notações de classificação.

Dessa forma, a indexação é “[...] uma operação essencial para que se possam recuperar documentos do acervo documentário e então responder, de forma clara e eficaz, a todo pedido ou questão dos usuários, sem que haja “RUÍDOS” (isto não corresponde ao que eu procurava), nem “SILÊNCIOS” (o documento existe, mas está perdido)” (Chaumier, 1988, p. 74).

A indexação como processo de construção de índices ocorre desde a Antiguidade quando, por exemplo, no terceiro milênio antes de Cristo, a Biblioteca de Ebla, na Síria, abrigava uma coleção composta por “[...] textos administrativos, literários e científicos, registrados em 15 mil tábuas de argila, as quais foram dispostas criteriosamente em estantes segundo o tema abordado, além de 15 tábuas pequenas com resumos do conteúdo de documentos” (Ortega, 2004).

A indexação como processo de representação de um documento é realizada “[...] mais intensamente desde o aumento de publicações periódicas e da literatura técnico científica de modo geral, que impulsionaram a necessidade de criação de mecanismos de controle bibliográfico em centros de documentação especializados” (Fujita, 2003, p. 61), aproximadamente após 1850.

O processo de indexação é que ditará o valor de um sistema de busca e recuperação da informação. Uma vez que não seja realizado de forma defensável, um enorme esforço será exigido do serviço de referência para que os usuários encontrem os documentos sobre o assunto procurado. Para fins de didáticos, o processo de indexação é composto por três momentos: análise de assunto, síntese e representação.

A análise de assunto, também denominada análise temática, análise documentária, análise conceitual ou análise de conteúdo, pode seguir três concepções de acordo com Albrechtsen (1993): simplista, orientada para o conteúdo ou orientada pela demanda. Apesar de considerar apenas a informação explícita ao documento, a concepção simplista tem custos mais baixos. Já a concepção orientada para o conteúdo vai além dos limites da estrutura superficial léxica e gramatical do conteúdo do documento, mas não considera seus possíveis usos. A concepção orientada pela demanda evoca a responsabilidade dos bibliotecários ao escolher quais são os aspectos de um determinado documento que devem ser representados visando seu acesso pelos usuários atuais e futuros. Albrechtsen (1993, p. 223) esclarece que a concepção orientada pela demanda “[...] envolve um alto nível de subjetividade e responsabilidade ao escolher dentre os atributos de um documento.” Nesse contexto, estruturas sociológico-epistemológicas provenientes da indexação também são consideradas.

Uma vez ciente da concepção de análise de assunto que deverá guiar o processo de indexação e classificação, o bibliotecário procederá à compreensão do conteúdo do documento, identificação dos conceitos que representam este conteúdo e seleção de conceitos válidos que se tornarão pontos de acesso de assunto para a recuperação da informação.

Para a compreensão do conteúdo do documento deve-se identificar o seu tema. Há o conteúdo intrínseco ao documento, determinado pelo autor da obra, que pode ser denominado como aboutness, atinência

extensional, atinência, topicalidade ou tematicidade intrínseca. E há o conteúdo de interesse para as comunidades de usuários, ou seja, ele é mutável e pode ser denominado como atinência intensional, significado, informatividade, tematicidade extrínseca (Fujita, 2003). A política de indexação estabelecida pela instituição decidirá se o tema extraído do documento será mais específico ou mais genérico.

Para a compreensão do tema, identificação e seleção dos conceitos que se tornarão pontos de acesso de assunto, o indexador vale-se da leitura documental pois a ele “[...] raramente é dado o luxo de poder ler um documento atentamente do começo ao fim [...] Usualmente, recomenda-se um misto de ler e ‘passar os olhos’ pelo texto” (Lancaster, 2004, p. 24). Por exemplo, haverá momentos, “[...] em que um trecho difícil, para o leitor, exige que ele leia linear e cuidadosamente, e há outros em que apenas inferências pelo contexto permitem a compreensão sem problemas” (Fujita, Nardi, & Santos, 1998, p. 21).

Durante a leitura palavra por palavra, os “esquemas” do leitor são ativados por dois movimentos complementares: *bottom-up*, quando o leitor caminha da parte para o todo observando o contexto do texto com o objetivo de extrair significado, e o *top-down* quando o leitor caminha do todo para a parte valendo-se de predições e com o objetivo de atribuir significado. Esses movimentos são estratégias cognitivas que compõem a leitura fluida, isto é, são comportamentos automáticos e inconscientes (Cintra, 1989).

Quando um profissional realiza uma leitura frente a um problema e com um objetivo específico, como no caso da indexação onde a intenção é compreender o tema do texto e identificar termos que representem os seus assuntos, pode se valer de estratégias metacognitivas composta por comportamentos desautomatizados, ações conscientes, assim “[...] enquanto o conhecimento prévio viabiliza, por força de ‘esquemas’, uma leitura mais rápida, as estratégias, especialmente as metacognitivas, conduzem à eficácia da tarefa” (Cintra, 1989, p. 36).

A NBR 12676 – Métodos para análise de documentos – determinação de seus assuntos e seleção de termos de indexação¹ fornece algumas questões as quais os termos indenticados pela leitura documental deveriam responder:

- a) qual o assunto de que trata o documento?
- b) como se define o assunto em termos de teorias, hipóteses, etc.?
- c) o assunto contém uma ação, uma operação, um processo?
- d) o documento trata do agente dessa ação, operação, processo etc.?
- e) o documento se refere a métodos, técnicas e instrumentos especiais?
- f) esses aspectos foram considerados no contexto de um local ou ambiente especial?
- g) foram identificadas variáveis dependentes ou independentes?
- h) o assunto foi considerado sob um ponto de vista interdisciplinar? (p.ex.: um estudo sociológico da religião) (Associação..., 1992, p. 2).

Uma vez indenticados os termos representativos do conteúdo do documento, o bibliotecário selecionará quais desses termos de fato representarão aquele documento em sua biblioteca tendo em vista os objetivos para os quais esses pontos de acesso de assunto serão usados. Assim, nem todos os

¹ O manual de política de indexação da Universidade Estadual Paulista Júlio de Mesquita Filho (FUJITA, 2014) oferece um modelo de leitura para para auxiliar a identificação dos termos representativos em livros.

conceitos identificados serão necessariamente selecionados. Esse é o final da análise, quando o indexador está preparado para afirmar sobre o que trata o documento, é quando se define o assunto. Assim, o produto da análise de assunto é a frase de indexação, ainda na linguagem natural do documento. Fujita (2003) sugere que o indexador não faça uso das linguagens de indexação até esse momento nem se deixe influenciar por elas.

A segunda etapa, a síntese, consiste no processo de sumarização, isto é, a construção do enunciado de assunto que é composto pelos termos selecionados e a elaboração do resumo. O resumo é uma apresentação concisa dos pontos relevantes de um texto e serve como instrumento de pesquisa em bases de dados fomentando, inclusive, a divulgação científica. O resumo apresenta um menor nível de condensação quando comparado ao enunciado de assunto.

A escolha dos conceitos que devem ser selecionados para compor o enunciado de assunto depende da finalidade para a qual serão usados os termos de indexação. Nesse caso, as características da indexação mais afetadas são: o grau de exaustividade e de especificidade dos termos selecionados.

A exaustividade refere-se à extensão em que os assuntos serão representados. Todos os conceitos potencialmente úteis serão selecionados? Quanto maior for a exaustividade na representação de assunto de um documento, maior será a revocação e menor será a precisão na recuperação da informação. A especificidade refere-se a quão precisos pode-se ser na representação de um documento. Quanto maior for a especificidade na representação de assunto de um documento, maior será a precisão e menor será a revocação na recuperação da informação (Carneiro, 1985).

A representação é a última etapa, é quando o indexador traduzirá a frase de indexação que está na linguagem natural do documento em descritores de indexação a partir de uma linguagem de indexação adotada pela biblioteca. Dois exemplos tradicionais de linguagens usados para indexação são: a) lista de cabeçalhos de assuntos que consistem em cabeçalhos e sub-cabeçalhos compostos por palavras em linguagem natural, os quais buscam condensar o tema sobre o qual trata o documento (Gil Urdiciain, 2004), e b) tesouro que é “[...] um instrumento que relaciona os descritores/termos de forma mais consistente, apresentando uma estrutura sintética simplificada e uma complexa rede de referências cruzadas” (Dodebei, 2002, p. 67).

Assim, ao elaborar uma estratégia de busca no catálogo da biblioteca serão os cabeçalhos de assunto ou os descritores de indexação que farão a ponte entre a necessidade informacional materializada nas palavras da estratégia de busca do usuário e os itens disponíveis no acervo da instituição.

INDEXAÇÃO NA ARQUIVÍSTICA

No âmbito da representação do conhecimento arquivístico, as funções de classificação e descrição são consideradas nucleares, baseando-se, fundamentalmente, na aplicação do princípio da proveniência e no conceito de fundo documental. Enquanto a primeira função representa as estruturas e funções do órgão produtor de documentos por meio de um plano de classificação, a segunda representa o conteúdo dos documentos e sua relação com o contexto de produção no qual está inserido, seja este social, político, econômico, cultural. Neste sentido, representar para a Arquivística significa contextualizar, ir além do conteúdo imediato do documento.

No entanto, essa definição de representação, nem sempre esteve atrelada aos objetivos da função descritiva. Duranti (1993) destaca três objetivos primordiais da descrição, contextualizando-os em seus respectivos momentos históricos. Para a autora, a descrição teve como objetivo: a) a criação de substitutos para os documentos e preservação de fundos, servindo à memória perpétua da sociedade, b) o fornecimento de prova da existência dos documentos, c) um guia de pesquisa para revelar a ordem intelectual de documentos fisicamente desordenados ou sem sentido, d) auxiliar o arquivista na condução de pesquisas, guiando qualquer tipo de usuário independente de sua busca, iluminando as relações contextuais entre os documentos. Esse último objetivo é, hoje, o mais perseguido pelos arquivistas. Iluminar as relações entre os documentos e seu contexto seria o *leitmotiv* da Arquivística Contemporânea.

A autora destaca, ainda, que a preocupação em definir ou conceituar a função da descrição é recente, datando da década de 1970, associando-a a três distintos conceitos complementares: a descrição enquanto um processo de análise, identificação e organização, a descrição enquanto um objetivo de controle, recuperação e acesso aos documentos e a descrição enquanto um produto final, ilustrando o material arquivístico, sua proveniência e contexto documental, suas inter-relações e os modos como podem ser identificados e utilizados.

Até o final da década de 1980, o segundo e terceiro conceito predominaram nas discussões teóricas e na prática arquivística. Construir instrumentos de referência para controlar e permitir ao usuário o acesso ao conteúdo do documento era a principal razão de ser da descrição. Em outras palavras, o lema era descrever para dar acesso.

Como consequência das novas formas de produção documental e da dinamicidade trazida pelas tecnologias de informação a partir da década de 1990, a Arquivística passa por uma renovação em sua teoria e, conseqüentemente, em sua prática, afetando, sobremaneira, os processos de produção, organização e representação do conhecimento arquivístico, deslocando seu foco do conjunto de documentos para os contextos social, organizacional e funcional em constante criação. Portanto, o ato de representar ao seguir um fluxo dinâmico desloca-se dos processos de arranjo e descrição estáticos e definitivos para arranjos contínuos, relativos e fluidos, com processos representativos sempre em andamento (YAKEL, 2003).

Buscando uma descrição dinâmica e fluida, seguindo a dinamicidade do processo de representação arquivística na contemporaneidade, emergem no contexto do Conselho Internacional de Arquivos, as normas de descrição arquivística, visando à garantia de descrições consistentes, apropriadas e autoexplicativas, possibilitando a padronização e o intercâmbio de informações entre acervos e instituições. Nesse sentido, destacam-se a ISAD (G), Norma internacional de descrição arquivística (1992), a ISAAR (CPF), Norma internacional de registro de autoridade arquivística para entidades coletivas, pessoas e famílias (1996), a ISDIAH, Norma internacional para a descrição de instituições com acervo arquivístico (2008) e a ISDF, Norma internacional para a descrição de funções (2007).

O objetivo da descrição arquivística é identificar e explicar o *contexto e o conteúdo de documentos de arquivo* a fim de promover o acesso aos mesmos. Isto é alcançado pela criação de representações precisas e adequadas e pela organização dessas representações de acordo com modelos predeterminados. Processos relacionados à descrição podem começar na ou antes da produção dos documentos e continuam durante sua vida. Esses processos permitem instituir controles intelectuais necessários para tornar confiáveis, autênticas, significativas e acessíveis descrições que serão mantidas ao longo do tempo (Conselho Internacional de Arquivo, 2000, p. 11).

Ainda segundo a ISAD (G), as normas internacionais estabelecem diretrizes gerais para a preparação de descrições arquivísticas e devem ser usadas em conjunto com as normas nacionais existentes, ou como base para a sua criação. Inspirado nesta sugestão, o Conselho Nacional de Arquivos do Brasil, em 2005, prepara uma versão preliminar da Norma Brasileira de Descrição – Nobrade, que veio a ser publicada oficialmente em 2006.

A Nobrade tem como pressupostos básicos o respeito aos fundos e a descrição multinível, ou seja, com o objetivo de representar o contexto e a estrutura hierárquica do fundo e suas partes componentes, a descrição deve ser feita do geral para o particular, apresentando informações revelantes para cada nível, para representar com rigor o contexto e o conteúdo da unidade de descrição. É importante, ainda, que sejam explicitadas as relações entre as descrições de cada unidade, evitando a redundância de informação em descrições hierarquicamente relacionadas.

A norma brasileira, tendo como base da ISAD (G), prevê a existência de 28 elementos de descrição, distribuídos em 8 áreas: (1) Área de identificação, (2) Área de contextualização, (3) Área de conteúdo e estrutura, (4) Área de condições de acesso e uso, (5) Área de fontes relacionadas, (6) Área de notas, (7) Área de controle da descrição, e (8) Área de pontos de acesso e descrição de assuntos. Dentre os 28 elementos de descrição disponíveis, sete são obrigatórios: código de referência, título, datas, nível de descrição, dimensão e suporte, nome(s) do(s) produtor(es) e condições de acesso.

Um dos importantes diferenciais da Nobrade é o campo descritivo de pontos de acesso e indexação de assuntos. Segundo a Nobrade (2006), ponto de acesso pode ser definido como um elemento de informação, termo ou código que, presente em unidades de descrição, serve à pesquisa, identificação ou localização de documentos. A identificação dos pontos de acesso é necessária para a recuperação do conteúdo de alguns elementos de descrição e são imprescindíveis à indexação de assuntos. Destaca-se que a descrição dos sete elementos obrigatórios expostos acima constituem pontos de acesso naturais da descrição multinível.

Para a indexação de assuntos, o arquivista deve considerar o Produtor, utilizando-se ainda da ISAAR (CPF), que oferece as orientações necessárias para a descrição do *nome do produtor*, e o *título* da unidade de descrição, bem como a *entidade custodiadora* e a *procêdencia*. Assuntos presentes nos campos da *história administrativa/biografia* e *âmbito e conteúdo* podem ser considerados, como, por exemplo, nomes de entidades, eventos, áreas geográficas, períodos e assuntos tópicos. Nesse sentido, na identificação de pontos de acesso para a descrição de Fundo, Seção ou item documental, os elementos relacionados à representação do contexto são essenciais à indexação.

É importante destacar que, embora a descrição não ocorra apenas no arquivo permanente – uma vez que o arquivista e o produtor descrevem os documentos e seus contextos desde sua gênese, respeitando a especificidade de representação para cada fase do documento, a indexação de assuntos é recomendada notavelmente a partir do arquivo permanente, uma vez que o estabelecimento de ponto de acesso servirá aos usuários no momento de recuperação da informação.

Deve-se dizer que a indexação é uma etapa posterior à representação arquivística e não concomitante a ela. Ou seja, somente após a descrição dos elementos elencados pelas normas de descrição, a indexação irá acontecer, isso significa dizer que os processos de representação (classificação e descrição) não deverão sofrer influência da indexação.

DISCUSSÃO: APROXIMAÇÕES TEÓRICAS

A indexação em arquivos já está prevista pela Nobrade e, enquanto a indexação de assuntos nos níveis 1 (fundo), 2 (seção/grupo) e 5 (item documental) está relacionada aos elementos de representação do contexto e conteúdo, a indexação de termos referentes à série (nível 3) pode ser um pouco mais complicada.

Segundo Ribeiro (2006), a indexação de conjuntos documentais, compostos por um número elevado de espécies e contendo uma variedade de assuntos, deve prescindir de métodos apropriados para disciplinar a análise do seu conteúdo, a identificação e a seleção dos conceitos a indexar. No contexto das séries documentais que, geralmente, possuem uma quantidade considerável de itens documentais, uma análise minuciosa de cada item é impossível. Assim, Ribeiro (2006) estabelece alguns elementos que podem ser considerados na identificação dos assuntos a serem indexados, são eles:

Titulo da série — quando exista; nestas circunstâncias ele pode ser retirado das lombadas ou capas dos livros, de folhas iniciais ou do interior dos mesmos; *Termo de abertura* — quando a série é formada por livros, na maioria dos casos, eles têm um termo de abertura, no qual é indicada a função a que se destinam; *Índices* — em muitos casos, os livros ou os maços que constituem uma série contêm, no início ou no final, índices ou listas de conteúdo, que são de extrema utilidade para o indexador; *Tipos documentais* — deverá analisar-se o conjunto da série, a fim de se verificar se há unidade tipológica nos documentos que a compõem, ou se se trata de uma série "miscelânea", constituída por tipos documentais diversos. Esta análise é fundamental, pois a unidade tipológica facilita a indexação, uma vez que, nos vários documentos, a organização do texto e o tipo de informação que registam são semelhantes; *Elementos que compõem cada documento* — sendo as séries, geralmente, formadas por tipos documentais análogos, é possível, na maioria dos casos, determinar os elementos que integram o seu conteúdo e, a partir daí, identificar os assuntos a indexar.

A análise da tipologia documental precede o estabelecimento das séries e, portanto, fornece elementos para o conhecimento das espécies documentais e das atividades desempenhadas pelos documentos. Assim, conhecer a tipologia documental permite ao arquivista conhecer o conteúdo dos documentos, que, no âmbito dos arquivos, engloba, além do assunto, sua proveniência e atividade.

Para os bibliotecários, o conhecimento da tipologia documental auxiliaria no processo de leitura documental pois, ao conhecer as áreas de maior conteúdo informativo de cada tipo de documento, a identificação de termos que poderão servir como pontos de acesso de assunto será mais rápida.

Guimarães (2008) defende a complementariedade das lógicas adotadas pela Arquivística e pela Biblioteconomia no contexto da Ciência da Informação, pois na Arquivística, a análise é feita priorizando o contexto de produção a partir dos princípios de proveniência e organicidade, enquanto que na Biblioteconomia a análise se faz priorizando o uso da informação a partir dos princípios de uso ou de demanda, por exemplo.

O princípio da proveniência advindo da Arquivística é fundamental para garantir a confiabilidade do documento a ser tratado, uma vez que consiste na identidade de um documento a partir do estabelecimento do seu produtor. Na Biblioteconomia, ao trabalhar com o documento em conjunto, no contexto do seu domínio, estamos valendo-nos do princípio da proveniência. Por exemplo, a organização de coleções de periódicos em bibliotecas se dá, em um primeiro momento, a partir da proveniência que é a instituição publicadora e da organicidade pois um fascículo não faz sentido sozinho.

CONSIDERAÇÕES FINAIS

Para a Arquivística, o conceito de assunto é importante, principalmente nas pesquisas realizadas em documentos no arquivo permanente. Diante de uma necessidade informacional, não seria defensável exigir dos usuários do arquivo permanente o conhecimento de todos os fundos ou séries de um arquivo para poderem realizar uma pesquisa. Uma vez que a indexação seja realizada nos documentos dos arquivos permanentes, os usuários poderão realizar pesquisas por assunto e recuperarão todos os documentos que, em algum momento, estabeleceram relações com o termo usado em sua estratégia de busca.

A indexação é um processo intelectual e, por não ser neutra é importante que antes de utilizá-la, os alunos de Arquivologia e arquivistas a estudem em sua totalidade, apreciando-a criticamente. O entendimento das convergências e divergências entre a Arquivística e a Biblioteconomia é imprescindível, inclusive para que fique claro que a indexação, como foi aqui apresentada, seria um processo adicional à prática arquivística que traria benefícios para o acesso e recuperação da informação em arquivos permanentes.

REFERÊNCIAS BIBLIOGRÁFICAS

- Albretchtsen, H. (1993). Subject analysis and indexing: from automated indexing to domain analysis. *The Indexer*, 18(4), 219-224.
- Associação Brasileira de Normas Técnicas. (1992). *NBR 12676: Métodos para análise de documentos - determinação de seus assuntos e seleção de termos de indexação*. Rio de Janeiro.
- Barros, T. H. B. (2016). A Indexação e a Arquivística: aproximações iniciais no universo teórico da organização e representação do conhecimento. *Encontros Bibli*, 21(46), 33-44.
- Carneiro, M. V. (1985). Diretrizes para uma política de indexação. *Revista da Escola de Biblioteconomia da UFMG*, 14(2), 221-241.
- Chaumier, J. (1988). Indexação: conceito, etapas, instrumentos. *Revista Brasileira de Biblioteconomia e Documentação*, 21(1/2), 63-79.
- Cintra, A. M. M. (1989). Estratégias de leitura em documentação. In J. W. Smit (Coord.), *Análise documentária: a análise da síntese* (p. 30-37) (2nd ed.). Brasília, DF: IBICT.
- Conselho Internacional de Arquivos. (2000). ISAD (G): *Norma geral internacional de descrição arquivística* (2nd ed.). Rio de Janeiro, RJ: Arquivo Nacional.
- Conselho Internacional de Arquivos. (2004). ISAAR (CPF): *Norma internacional de registro de autoridade arquivística para entidades coletivas, pessoas e famílias*. (2nd ed.). Rio de Janeiro, RJ: Arquivo Nacional.
- Conselho Nacional de Arquivos. (2006). *Norma Brasileira de descrição arquivística*. Rio de Janeiro, RJ: Arquivo Nacional.

- Dobebei, V. L. D. (2002). *Tesouro: linguagem de representação da memória documentária*. Rio de Janeiro, RJ: Interciência.
- Duranti, L. (1993) Origin and development of the concept of archival description. *Archivaria*, 35, 47-54.
- Fujita, M. S. L. (2003). A identificação de conceitos no processo de análise de assunto para indexação. *Revista Digital de Biblioteconomia e Ciência da Informação*, 1(1), 60-90.
- Fujita, M. S. L. (Coord.) (2014). *Manual de política de indexação para as bibliotecas universitárias da UNESP*. São Paulo, SP: UNESP.
- Fujita M. S. L., Nardi, M. I. A., & Santos, S. (1998). A leitura em análise documentária. *Transinformação*, 10(3), 13-31.
- Gil Urdiciain, B. (2004). *Manual de lenguajes documentales* (2nd ed.). Gijón, AS: Ediciones Trea.
- Guimarães, J. A. C. (2008). Ciência da Informação, Arquivologia e Biblioteconomia: em busca do necessário diálogo entre o universo teórico e os fazeres profissionais. In M. S. L. Fujita, & J. A. C. Guimarães (Org.), *Ensino e Pesquisa em Biblioteconomia no Brasil: a emergência de um novo olhar* (p. 33-44). São Paulo, SP: Cultura Acadêmica.
- International Federation of Library Associations and Institutions (2010). *Functional Requirements for Subject Authority Data (FRSAD): A conceptual model*. Netherlands: IFLA. Retrieved from: <http://www.ifla.org/node/5849>
- Lancaster, F. W. (2004). *Indexação e resumos: teoria e prática* (2nd ed.). Brasília, DF: Briquet de Lemos.
- Ison, H. A. (2002). *The power to name: locating the limits or subject representation in libraries*. Dordrecht: Kluwer Academic Publisher.
- Ortega, C. D. (2004). Relações históricas entre Biblioteconomia, Documentação e Ciência da Informação. *DataGramaZero – Revista de Ciência da Informação*, 5(5).
- Ribeiro, F. (1996). *Indexação e controlo de autoridade em arquivos*. Porto: Câmara Municipal do Porto/Arquivo histórico.
- Silva, M. B. da, Sousa, D. E. L. de, & Bandeira, P. M. (2012). A representação temática em documentos arquivísticos: o caso da indexação documental realizada pelos alunos de Arquivologia da UFPB. *InCID: Revista de Ciência da Informação e Documentação*, 3(2), 124-141.
- Smit, J. W. (1987). *O que é documentação* (2nd ed.). São Paulo, SP: Brasiliense.
- Vital, L. P. (2017). *Representação temática de documento arquivístico: as contribuições do modelo conceitual FRSAD*. (Doctoral Dissertation in Information Science). Universidade Federal de Santa Catarina. Retrieved from repositorio.ufsc.br/bitstream/handle/123456789/174305/345863.pdf?sequence=1&isAllowed=y
- Yakel, E. (2003). Archival Representation. *Archival Science*, 3, 1-25.



MAPEAMENTO SISTEMÁTICO DA PRODUÇÃO CIENTÍFICA NO ÂMBITO DA BASE DE DADOS EM CIÊNCIA DA INFORMAÇÃO: PERIÓDICOS CIENTÍFICOS EM DISCUSSÃO

Jorge Santa Anna¹, Maria Elizabeth de Oliveira Costa²

¹Universidade Federal de Minas Gerais, 0000-0002-0709-3639, professorjorgeufes@gmail.com

²Universidade Federal de Minas Gerais, mabethcosta@gmail.com

RESUMO Traz à baila alguns indicadores sobre as características intrínsecas quanto extrínsecas dos periódicos científicos. Objetiva apresentar mapeamento da produção científica sobre o tema periódicos científicos, com foco nos artigos de periódicos publicados na Base de Dados em Ciência da Informação, no âmbito brasileiro. O método adotado foi o mapeamento sistemático, o qual demonstrou haver heterogeneidade na quantidade de artigos publicados ao longo dos anos; homogeneidade no número de artigos publicados nas revistas, com discrepância acentuada em apenas uma revista; prevalência da autoria coletiva com alto índice de diferenciação com a autoria individual; preferência por métodos mistos de pesquisa; e, por fim, percebeu-se diversidade de assuntos abordados, sendo que aspectos relacionados à comunicação científica, periódico eletrônico, portais de periódicos, usabilidade e sistema eletrônico de editoração de periódico manifestam-se como assuntos recorrentes. Conclui-se que, mesmo sendo um tema em voga, haja vista sua contribuição para as ciências e diante dos novos paradigmas impostos com o aparecimento do periódico eletrônico, esforços devem ser empreendidos no sentido de garantir a qualificação e melhoria contínuas dessas publicações. Ações também devem ser realizadas, haja vista proporcionar mais legitimidade, confiabilidade e segurança nos processos editoriais, de modo a aumentar a visibilidade, uso e reconhecimento dos periódicos como principais e confiáveis meios de comunicação científica para o século XXI.

PALAVRAS-CHAVE *Produção científica, Periódico científico, Periódico científico eletrônico, Avaliação de periódico, Indicadores de qualidade em periódicos.*

ABSTRACT It brings to light some indicators about the intrinsic and extrinsic characteristics of scientific journals. It aims to present a mapping of scientific production on the topic of scientific journals, focusing on articles from periodicals published in the Database on Information Science in the Brazilian context. The method adopted was systematic mapping, which showed heterogeneity in the number of articles published over the years; Homogeneity in the number of articles published in the journals, with a marked discrepancy in only one journal; Prevalence of collective authorship with high index of differentiation with individual authorship; Preference for mixed methods of research; And finally, a diversity of subjects was observed, and aspects related to scientific communication, electronic periodicals, periodicals portals, usability and electronic system of periodical publishing appear as recurring subjects. It is concluded that, even though it is an issue in vogue, given its contribution to the sciences and the new paradigms imposed by the appearance of the electronic journal, efforts must be made to ensure the continuous qualification and improvement of these publications. Actions should also be taken to provide more legitimacy, reliability, and security in editorial processes, so as to increase the visibility, use, and recognition of journals as the principal and reliable means of scientific communication for the twenty-first century.

KEYWORDS *Scientific production, Scientific periodical, Electronic scientific periodical, Periodic evaluation, Periodical quality indicators.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Desde sua gênese, no século XVII, os periódicos científicos têm contribuído para o progresso das ciências, principalmente por facilitar a divulgação dos resultados de pesquisas desenvolvidas por cientistas em diversas áreas do conhecimento, o que tem caracterizado esse tipo de publicação como um dos mais importantes instrumentos de comunicação científica, ao longo dos tempos.

É bem verdade que, o que assistimos hoje, denominado por muitos de explosão da informação, momento marcante pelo aumento do número de publicações disseminadas na sociedade, é consequência, em grande parte, da rapidez requerido no processo editorial dos periódicos, como apresentado na obra de Solla Price (1993).

Portanto, o periódico científico, em comunhão com a tecnologia da imprensa de Gutenberg, ainda no século XV, são os dois fatores mais impactantes que influenciaram a rapidez dos processos de editoração, por conseguinte, viabilizaram o aumento do número de fontes de informação publicadas, o que proporcionou, também, como resultado dessas ocorrências, o crescimento da produção científica (Solla Price, 1993).

Além dos reflexos oriundos com o uso da tecnologia da imprensa e com a proliferação dos periódicos científicos, outro fator que contribui demasiadamente com o aumento da produção editorial diz respeito ao surgimento da editoração eletrônica, a partir do uso das tecnologias digitais, as quais promoveram novas formas de interação e trocas de informações entre cientistas, além de ampliar a capacidade de acesso e visibilidade das descobertas científicas, principalmente com o uso, pela comunidade científica, do periódico científico eletrônico (Oliveira, 2008).

O periódico científico revolucionou os rumos da comunicação científica, ao eliminar possíveis barreiras no âmbito dos achados e comunicados de pesquisa. De forma semelhante, o periódico eletrônico amplia as possibilidades de interação, proporcionando rapidez na divulgação das pesquisas (Mueller, 2003) e, por outro lado, gera problemas mais abrangentes, como aspectos ligados à preservação dos dados digitais, e a necessidade de se elaborar políticas específicas para nortear os fazeres editoriais (Café & Fachin, 2007, Sayão, 2010).

Embora assiste-se na segunda década do século XXI, o desaparecimento do periódico impresso, uma tendência que se consolida a cada dia (Sayão, 2010), importante frisar, conforme apontado no estudo de Oliveira (2007), que a preservação e o arquivamento da informação, a garantia de acesso ao longo do tempo, o desenvolvimento de interfaces confiáveis e a disponibilização de coleções retrospectivas representam os grandes desafios ainda impostos na aceitação definitiva do periódico eletrônico, como novo canal de comunicação científica, na sociedade contemporânea.

Além desses desafios, manifestam-se outros, tais como a elaboração de instrumentos de controle, armazenamento e gestões das coleções digitais, haja vista permitir a preservação da memória científica (Ohira *et al.* 2003). Esse fato tem gerado a elaboração dos portais de periódicos, tema recorrente na

literatura e que viabiliza, nesse contexto, o estabelecimento de mecanismos e indicadores que apontem a qualidade das publicações periódicas, logo o conhecimento acerca das características da produção científica divulgada em periódicos (Garrido & Rodrigues, 2010), além das técnicas utilizadas para entender o uso dessas plataformas e a satisfação dos utilizadores desses espaços (Cendón, Souza & Ribeiro, 2011).

Portanto, que os periódicos científicos, no âmbito da comunicação científica, representam uma temática pertinente de investigação, não resta dúvida. Que eles são fontes de informação essenciais para proporcionar o acesso aos achados de pesquisa, por conseguinte, possuem árdua contribuição na produção científica, também representa uma informação incontestável. Com efeito, é pertinente realizar um mapeamento da produção científica no conjunto de publicações periódicas de uma base de dados especializada em armazenar periódicos.

Destarte, o estudo que ora se apresenta objetiva mapear a produção científica sobre o tema periódicos científicos, com foco nos artigos de periódicos publicados na Base de Dados em Ciência da Informação (Brapici), no decorrer dos anos. Para tanto, o estudo realiza levantamento das características principais (intrínsecas e extrínsecas) dos periódicos científicos presentes na Brapci, tais como: ano de publicação, autorias dos artigos, periódicos utilizados na divulgação dos artigos, métodos de pesquisa e assuntos principais apresentados.

No que se refere aos procedimentos metodológicos, utilizou-se pesquisa exploratória, uma vez que foram analisadas as principais tendências estudadas em um contexto específico de investigação (Gil, 2010). Ademais, a fim de explorar a literatura, recorreu-se ao mapeamento sistemático, o qual, segundo Santos e Kobashi (2009), esse método tem se estabelecido como uma tendência significativa para aprofundar o debate sobre as dinâmicas da ciência e sua comunicação, permitindo enfoques cognitivos que ultrapassam a análise quantitativa.

Em linhas gerais, o percurso metodológico visou, em alguns aspectos, identificar características e indicadores relativos aos periódicos científicos, de modo a viabilizar a pertinência da contribuição desse tema para o avanço do conhecimento científico, o que só é possível com recurso do mapeamento sistemático da literatura (Proença Junior & Silva, 2016).

No âmbito desse mapeamento, levantar as informações presentes na amostra de pesquisa, de um modo geral, pode conduzir à análise da produção científica de um determinado campo do conhecimento, área temática ou mesmo assuntos específicos, consistindo, dessa forma, como uma ferramenta fundamental no sentido de se “[...] perceber ou não a ocorrência de estruturas, tendências, dinâmicas e movimentos sociais que podem estar expressos nessa produção, seja analisada de forma específica ou em estudos comparados com outras áreas, temas e assuntos [...]” (Martins, 2013, p. 71).

BREVES CARACTERIZAÇÕES DOS PERIÓDICOS CIENTÍFICOS

Os periódicos, assim como os livros, são fontes de informação primária, devidamente publicadas e que tramitam pelos processos editoriais, cujo objetivo desse tipo de fonte é registrar informações atualizadas, sem a necessidade de fontes intermediárias ou mediadoras para acesso ao conteúdo informacional, fato esse que caracteriza os periódicos como publicações correntes (Miranda & Carvalho, 2014).

Essas fontes possuem como característica principal o fato de serem editadas em intervalos de tempo, que podem ser regulares ou não, o que confere a fidelização a uma periodicidade, proporcionando a continuidade da publicação, por meio de números, volumes e fascículos (Cunha, 2001).

As fontes de informação periódicas, para muitos também denominadas de seriadas, podem se apresentar em diferentes formatos, mantidos por organizações diferenciadas, possuir objetivos e políticas específicas, como também direcionar-se a públicos gerais ou especializados. Logo, existem diversas modalidades de periódicos, sendo que quando essas publicações constituem elementos fundamentais na disseminação e progresso da ciência e tecnologia em um país, manifesta-se o periódico científico (Fachin & Hillesheim, 2006).

Assim, os periódicos científicos são um tipo de publicação periódica que tem como um de seus principais objetivos divulgar os resultados de pesquisas científicas, sobre os mais variados assuntos nas mais diversas áreas do conhecimento humano. Representam o tipo de fonte de informação científica mais utilizada para recuperar e manter os cientistas atualizados acerca das descobertas e avanços das ciências (Fachin & Hillesheim, 2006).

A utilização dessas publicações no meio científico viabiliza rapidez na divulgação do que está sendo discutido e pesquisado dentro de um campo ou vertente de conhecimento, fato esse que muito difere das publicações no formato de livros ou longos tratados, os quais demandam um tempo maior no processo de investigação quanto de editoração. Portanto, a partir da criação do periódico científico, em meados do século XVII, esse meio de comunicação é consagrado como o elemento transferidor de conhecimento, quanto o de oportunizador da comunicação entre os pares de uma comunidade científica (Miranda & Pereira, 1996).

Assim, o periódico científico caracteriza-se como um meio ou instrumento formal de divulgação científica, o qual tem sido considerado o mais influente e utilizado pelos cientistas (Fachin & Hillesheim, 2006, Oliveira, 2008), sofrendo transformações ao longo dos tempos, de modo a atender as necessidades demandadas pelas comunidades científicas, sem, contudo, perder a essência dos rigores e métodos científicos (Mueller, 2003). Logo, esse tipo de publicação transferiu-se “[...] de um veículo, cuja finalidade consistia em publicar notícias científicas, em um veículo de divulgação do conhecimento que se origina das atividades de pesquisa” (Miranda & Pereira, 1996, p. 375).

Em decorrência disso, esse instrumento de comunicação participa intensamente do fluxo da comunicação científica, a qual pode ser definida como "o conjunto de atividades associadas à produção, disseminação e uso da informação, desde o momento em que o cientista concebe uma ideia para pesquisar, até que a informação acerca dos resultados seja aceita como constituinte do conhecimento científico" (Garvey, 1970 apud Miranda & Pereira, 1996, p. 375).

Com efeito, o periódico científico está relacionado ao processo de comunicação científica, uma vez que possui como principal característica fornecer informações selecionadas e atualizadas de assuntos representativos e de interesse para o melhoramento e o avanço das pesquisas, por conseguinte, proporciona a evolução do conhecimento (Fachin & Hillesheim, 2006). Destarte, segundo Mueller (2003), a editoração de periódicos obedece a rigores científicos, adentrando-se aos princípios básicos da ciência, qual seja, garantir confiabilidade, credibilidade e legitimidade do que está sendo descoberto e disponibilizado para uso.

Uma análise à historiografia dessas fontes de informação permite esclarecer que, ao longo dos séculos, os periódicos foram se modificando, contemplando desde notícias de diversas áreas do conhecimento (como o *Journal de Sçavans*), a relatos sobre descobertas de áreas específicas (como os periódicos especializados), até incorporar as tecnologias digitais, de modo a promover novas formas de interação e novos direcionamentos no fluxo da comunicação científica, tal como presenciamos hoje, com a consolidação do periódico científico eletrônico (Oliveira, 2008; Barbosa et al., 2013).

O surgimento do periódico científico eletrônico representou mudanças radicais na forma de se comunicar, seja entre cientistas, como demais partícipes da comunicação científica, além da relação estabelecida e da redefinição dos papéis atribuídos a editores e outros interessados (Mueller, 2003). Todavia, as características e os rigores metodológicos que caracterizam a essência do fazer científico continuaram inalteráveis, como nos ensina a autora citada.

A priori, evidencia-se que uma das características mais visíveis do periódico eletrônico está em seu formato. Sendo assim, essas fontes eletrônicas são compostas por

[...] quaisquer publicações que tenham a intenção de disponibilizar artigos científicos de forma subsequente ou continuada (não interrompida, em intervalos regulares ou não) e que adotam alguma forma de procedimento de controle de qualidade (não necessariamente avaliação prévia) em **meio eletrônico** (Gomes, 1999, p. 10-11, grifo nosso).

Nas últimas décadas do século XX e mais intensamente, nos primeiros anos do século sucessor, as facilidades proporcionadas com o uso das tecnologias digitais fizeram com que os periódicos impressos, de forma muito rápida, migrassem para o formato eletrônico. Mesmo com essa migração, alguns permaneceram durante bom tempo com o formato impresso, caracterizando as versões híbridas. No entanto, nos últimos anos, observa-se um crescimento e uma tendência cada dia mais frequente, do abandono à versão impressa, condicionando a disponibilização, apenas, do formato eletrônico (Barbosa et al., 2013).

METODOLOGIA

No contexto deste estudo, o mapeamento foi realizado em uma das mais importantes bases de dados que indexa periódicos científicos pertencentes à área da Ciência da Informação, no Brasil, a Brapci. Ressalta-se que o acervo dessa base é composto por 16.555 trabalhos distribuídos em 35 revistas cadastradas, de abrangência nacional, 2.244 trabalhos em eventos, além de uma tese e de um livro indexado. Ademais, dos periódicos disponíveis, 27 estão ativos e 8 históricos (descontinuados) (Brapci, 2017).

No que se refere à natureza do tipo de trabalho a ser analisado, optou-se pelo levantamento apenas dos artigos de periódicos, haja vista a maior quantidade desse tipo de publicação existente no acervo da referida base de dados. Os critérios seletivos das fontes recuperadas para serem mapeadas podem ser visualizados na tabela 1.

Tabela 1. Critérios seletivos para formação da amostra a ser mapeada

Do tipo de publicação a ser analisada	Artigos de periódicos
Do ambiente de investigação	BRAPCI
Da delimitação do tempo	De 1972 a 2017 (conforme limite estabelecido no próprio catálogo da Brapci)
Dos descritores utilizados na busca dos artigos	1 - periódico científico, 2 - periódico eletrônico, 3 – portais de periódicos
Das categorias de análise	1 - ano de publicação, 2 - autorias, 3 - periódicos em que os artigos foram publicados, 4 - métodos de pesquisa, 5 – assuntos principais
Das estratégia para identificação das categorias de análise	Leitura ao título, resumo, introdução e conclusão dos artigos

Referência: dados da pesquisa (2017)

Importante destacar o número de publicações recuperadas, após levantamento no catálogo eletrônico, e o número de publicações selecionados após o processo de leitura, considerando cada um dos descritores utilizados durante a busca e recuperação dos artigos (tabela 2).

Tabela 2. Comparativo entre a quantidade de artigos recuperados e os selecionados após a prática da leitura

Descritores	Artigos recuperados	Artigos selecionados após leitura
Periódico científico	40	25
Periódico eletrônico	34	22
Portais de periódicos	6	5

Referência: dados da pesquisa (2017)

Sendo assim, após as etapas iniciais do mapeamento sistemático (busca, coleta e filtragem das referências), concretizou-se uma amostra composta por 52 artigos, distribuídos em variadas revistas da Ciência da Informação no Brasil e que se encontram cadastradas na Brapci. Logo em seguida, manifestou-se a etapa referente ao relato dos resultados obtidos com o mapeamento da amostra de pesquisa.

PRINCIPAIS RESULTADOS

Analisando a amostra de pesquisa para mapeamento sistemático, no que se refere aos anos de publicação, nota-se considerável discrepância quando ao número de pesquisas ao longo dos anos, o que confirma heterogeneidade nessa categoria de análise, conforme demonstrado na figura 1.

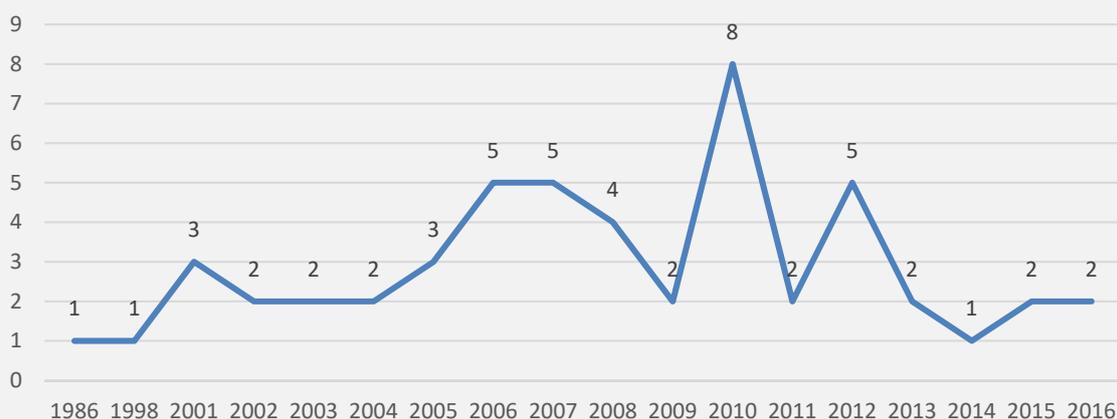


Figura 1. Distribuição do número de artigos por ano

Fonte: dados da pesquisa (2017)

Considerando o recorte temporal adotado para esta pesquisa, qual seja, de 1972 a 2017, nota-se que o primeiro artigo sobre periódico científico aparece em 1986 e o seguinte, apenas 11 anos depois, em 1998, sendo que não houve publicação no ano de 1999 e 2000. No início do século XXI, o número de publicações oscila, havendo publicações em todos os anos, com crescimento mais elevado no ano de 2010 (com oito artigos) e, por fim, nos dois últimos anos, em 2015 e 2016, manifestam-se duas publicações em cada um desses anos. Esses dados demonstram heterogeneidade na quantidade de artigos ao longo dos tempos, mesmo que a temática tenha se fortalecido com a chegada do periódico eletrônico, a partir da década de 1970, conforme discorrido no estudo de Oliveira (2008).

Ao mapear os periódicos que publicaram os 52 artigos da amostra, considerando os meios de comunicação utilizados para divulgação dos trabalhos, percebe-se a presença de grande parte das revistas de Ciência da Informação existentes em território brasileiro indexadas na Brapci, ou seja, o tema periódico é apresentado em 19 periódicos dessa área de conhecimento (figura 2).

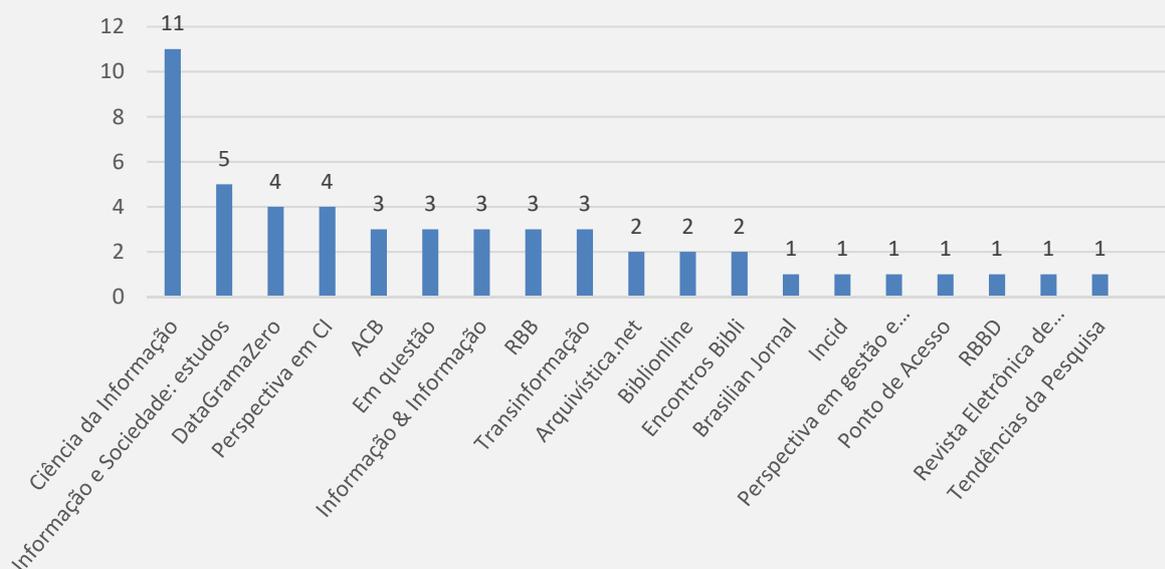


Figura 2. Distribuição das revistas por quantidade de artigos

Fonte: dados da pesquisa (2017)

A leitura ao gráfico supracitado demonstra o maior número de publicações na Revista Ciência da Informação, com 11 artigos, seguida da revista Informação e Sociedade, com cinco publicações e, empatadas com quatro artigos cada, aparecem as revistas DataGramaZero e Perspectiva em Ciência da Informação. Portanto, com base nesses dados, é possível identificar discrepância acentuada entre uma revista no conjunto das demais revistas, sendo que, entre as 18 revistas, percebe-se uma certa homogeneidade no número de artigos, em um intervalo compreendido entre um a cinco artigos por revista. Essa discrepância poderá estar associada à quantidade de fascículos publicados por cada revista, considerando sua periodicidade, regularidade ou ano de existência, dentre outros fatores.

No que se refere à categoria autoria, importante destacar que foi analisado cada artigo considerando o número de autores que participaram da elaboração de cada trabalho, com vistas aos seguintes aspectos: autoria individual, dois autores, três autores, quatro ou mais. Os 52 trabalhos mapeados apresentam diferenças quanto ao número de autores por artigos, sendo que a maioria dos artigos (22) foi desenvolvida por dois autores, seguido de 17 artigos, os quais apresentam autoria individual, 10 artigos contendo três autores, e, por fim, em um total de três trabalhos houve a participação de quatro ou mais autores. Esse resultado permite, de alguma forma, confirmar a preferência dos autores publicarem trabalhos de forma coletiva, a chamada autoria múltipla (35 trabalhos).

Outro aspecto a ser considerado no processo de mapeamento diz respeito ao método de pesquisa empregado pelos autores a fim de se atingir as propostas por eles levantadas. A análise aos 52 artigos constatou a preferência por métodos variados de pesquisa, principalmente no que se refere à combinação de dois ou três métodos diferentes em muitos artigos. Assim, constatou-se a preferência por métodos mistos e devidamente combinados, procedimentos esses utilizados em 32 trabalhos da amostra. Em linhas gerais, esses métodos contemplaram a aplicação de técnicas quantitativas de coleta de dados, sobretudo análise estatística, sustentada por softwares específicos para tabulação de dados. Portanto, constata-se forte tendência para análise de citações e técnicas bibliométricas, utilizadas, principalmente, em trabalhos que objetivavam levantar e mapear a produção científica de portais e demais ambientes de armazenamento de periódicos.

Ademais, merecem destaques os artigos que adotaram um único método de condução do estudo, sendo identificados o uso da pesquisa bibliográfica em dez trabalhos, incluindo-se nessa categoria estudos teóricos, ensaios e revisões de literatura, além de quatro artigos adotarem apenas o relato de pesquisa, sobretudo no que se refere à implantação de portais e migração dos periódicos impressos para os eletrônicos. Também se identificou em quatro artigos apenas o uso da pesquisa documental, geralmente realizada em documentos institucionais, administrativos e relatórios formulados a partir da consulta às plataformas digitais e, por fim, em dois artigos concretiza-se apenas a pesquisa de campo, em grande parte, realizada por meio da aplicação de entrevista ou questionário a usuários dos sistemas editoriais eletrônicos.

Por fim, a partir da análise ao conteúdo abordado nos artigos, constata-se que os assuntos mais recorrentes, publicados na literatura analisada nesta pesquisa são: comunicação científica, periódico eletrônico, portais de periódicos, usabilidade e sistema eletrônico de editoração de periódico, o que infere afirmar que a editoração eletrônica vem se consolidando nos novos tempos, representando, nesse contexto, novas práticas de comunicação. As demais palavras juntas somam um total de 171 palavras, evidenciando a diversidade de assuntos presentes no âmbito dos periódicos científicos.

Portanto, semelhante ao estudo de Mueller (2003, 2006), presume-se uma redefinição nos processos editoriais nos periódicos, a partir da incorporação das tecnologias digitais, no entanto, a essência do

fazer científico permanece o mesmo. Por outro lado, algumas questões ainda não foram integralmente solucionadas, sobretudo no que se refere à preservação do digital e a filosofia do acesso aberto, conforme ensinado por Oliveira (2008), questões essas que requerem soluções, para que o periódico eletrônico seja efetivamente e consolidado na sociedade moderna.

CONCLUSÕES

Ao findar este artigo, acendemos o conceito de periódico científico, o qual se manifesta como uma publicação de caráter sequencial direcionada a divulgar os resultados de pesquisa e contribuir para a comunicação e divulgação do conhecimento produzido pelas diversas áreas do saber humano. Reiteramos as funcionalidades e contribuições desse recurso para o progresso das ciências e as transformações que vem atravessando, ao longo dos tempos, atingindo, na atualidade, sua inserção em plataformas e ambientes digitais, os quais permitem a ampliação do fluxo da comunicação, como também garantem o armazenamento, recuperação e divulgação das descobertas científicas.

A partir dos resultados obtidos com esta pesquisa, foi possível apresentar as principais características da produção científica da Brapci, sobre o tema periódicos científicos, considerando alguns aspectos intrínsecos das publicações analisadas, o que consolidou o mapeamento dessa literatura, a qual pode ser considerada como um dos fragmentos da Ciência da Informação no Brasil.

Considerando as categorias para análise, em linhas gerais, constatou-se heterogeneidade na quantidade de artigos publicados ao longo dos tempos; homogeneidade no número de artigos publicados nas revistas, com discrepância acentuada em apenas uma revista; prevalência da autoria coletiva com alto índice de diferenciação com a autoria individual; preferência por métodos mistos de pesquisa; e, por fim, percebeu-se diversidade de assuntos abordados, sendo que aspectos relacionados à comunicação científica, periódico eletrônico, portais de periódicos, usabilidade e sistema eletrônico de editoração de periódico manifestam-se como assuntos recorrentes.

A partir desses indicadores, conclui-se que, mesmo sendo um tema em voga, haja vista sua contribuição para as ciências e diante dos novos paradigmas impostos com o aparecimento do periódico eletrônico, esforços devem ser empreendidos no sentido de garantir a qualificação e melhoria contínuas dessas publicações. Ações também devem ser realizadas, haja vista proporcionar mais legitimidade, confiabilidade e segurança nos processos editoriais, de modo a aumentar a visibilidade, uso e reconhecimento dos periódicos como principais e confiáveis meios de comunicação científica para o século XXI.

Essas constatações evidenciam a realização de estudos futuros, haja vista mapear quanto avaliar os periódicos, considerando outras categorias de análise e outros ambientes e contextos de investigação, no intuito de melhor qualificar e disseminar essas publicações. Além disso, recomenda-se a elaboração de estratégias de melhoria para os periódicos eletrônicos, sobretudo no que tange à preservação digital e ao acesso aberto, como também a concretização de modelos de gestão, planejamento e políticas voltadas a garantir a sobrevivência dos periódicos e sua melhor utilização pela comunidade científica nas décadas vindouras.

REFERÊNCIAS BIBLIOGRÁFICAS

- Barbosa, A. G. et al. (2013). Evolução das funções dos periódicos científicos e suas aplicações no contexto atual. *Múltiplos Olhares em Ciência da Informação*, Belo Horizonte, 3(1), 1-10. Recuperado em: 20 março, 2017, de <http://portaldeperiodicos.eci.ufmg.br/index.php/moci/article/view/1970>.
- Brapci (2017). *Sobre a Brapci*. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/>.
- Cendón, B. V., Souza, J. L., & Ribeiro, N. A. (2011). Satisfação dos usuários do Portal de Periódicos da Capes: um estudo sobre a obtenção de sucesso no uso do sistema. *Perspectivas em Ciência da Informação*, 16(2), 67-100. Recuperado em 20 março, 2017, de <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/1236>.
- Cunha, M. B. (2001). *Para saber mais: fontes de informação em ciência e tecnologia*. Brasília: Briquet de Lemos.
- Fachin, G.R B., Hillesheim, A.I.A. (2006). *Periódico científico: padronização e organização*. Florianópolis: Ed. da UFSC.
- Garrido, I. S. & Rodrigues, R. S. (2010). Portais de periódicos científicos online: organização institucional das publicações. *Perspectivas em Ciência da Informação*, 15(2), 56-72. Recuperado em 20 de março de 2017, de <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/943/732>.
- Gil, A. C. (2010). *Como elaborar projetos de pesquisa* (5a ed.). São Paulo: Atlas.
- Gomes, S. H. A. *Inovação tecnológica no sistema formal de comunicação científica: os periódicos eletrônicos nas atividades de pesquisa dos acadêmicos de cursos de pós-graduação brasileiros*. 1999. 465 f. Tese (Doutorado) – Faculdade de Estudos Sociais Aplicados, Universidade de Brasília, Brasília, DF, 1999.
- Martins, D. L. (2013). Mapeamento da produção científica e participação dos autores nos anais do ENANCIB: uma análise a partir da correlação entre as redes de coautoria e redes de participação nos grupos de trabalho. *RICI*, 6(2), 71-88. Recuperado em 20 março, 2017, de <http://periodicos.unb.br/index.php/RICI/article/view/9385/7662>.
- Miranda, D.B., Pereira, M.N.F. (1996). O periódico científico como veículo de comunicação: uma revisão de literatura. *Ciência da Informação*, Brasília, DF, 25(3), p. 375-382. Recuperado em 20 março, 2017, de <http://revista.ibict.br/ciinf/article/view/636/640>.
- Miranda, A.C.C., Carvalho, M.M.C. (2014). Desenvolvimento de coleções de fontes de informação eletrônicas em bibliotecas universitárias. *Biblionline*, João Pessoa, 10(1), 15-28, 2014. Recuperado em 20 março, 2017, de <http://periodicos.ufpb.br/index.php/biblio/article/view/17030>.
- Mueller, S. P. M. (2003). O periódico científico. In: Campello, B. S., Cendón, B. V. & Kremer, J. M. (Org.). (2003). *Fontes de informação para pesquisadores e profissionais* (2a ed.) Belo Horizonte: Editora UFMG.

Mueller, S. P. M. (2006). A comunicação científica e o movimento de acesso livre ao conhecimento. *Ciência da Informação*, 35(2), 27-38. Recuperado em 20 março, 2017, de <http://www.scielo.br/pdf/ci/v35n2/a04v35n2.pdf>.

Ohira, M. L. B., Prado, N. S., Oliveira, F., Rosa, L. G., Nazario, V. M. & Correia, C. (2003). Análise dos periódicos eletrônicos (full text) em ciência da informação: América Latina, Caribe, Portugal e Espanha. *Informação & Informação*, 8(1), 1-24. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/4345>.

Oliveira, R. B. P. M. (2007). Uso de periódicos científicos eletrônicos por docentes e pós-graduandos do Instituto de Geociências da USP. *Ciência da Informação*, 36(3), 59-66. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/5137>.

Oliveira, R. B. P. M. (2008). Scientific electronic journals: definitions and history. *Informação & Sociedade: Estudos*, 18(2), 1-10. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/16664>.

Proença Júnior, D., & SILVA, E. R. (2016). Contexto e processo do Mapeamento Sistemático da Literatura no trajeto da Pós-Graduação no Brasil. *TransInformação*, 28(2), 233-240. Recuperado em 20 março, 2017, de <http://www.scielo.br/pdf/tinf/v28n2/0103-3786-tinf-28-02-00233.pdf>.

Santos, R. N. M., & KOBASHI, N. Y. (2009). Bibliometria, cientometria, infrometria: conceitos e aplicações. *Pesquisa brasileira em Ciência da Informação*, 2(1), 155-172. Recuperado em 20 março, 2017, de <http://basessibi.c3sl.ufpr.br/brapci/index.php/article/download/7870>.

Sayão, L. F. (2010). Repositórios digitais confiáveis para a preservação de periódicos eletrônicos científicos. *PontodeAcesso*, 4(3), 68-94. Recuperado em 20 março, 2017, de <https://portalseer.ufba.br/index.php/revistaici/article/view/4709/3565>.

Solla Price, D. J. (1993). *Little science, big science*. New York: Columbia University Press.



O CONTROLO DE AUTORIDADE DE DOCUMENTOS FOTOGRÁFICOS NOS ARQUIVOS

Joana Rodrigues

*Faculdade de Engenharia da Universidade do Porto - Faculdade de Letras da Universidade do Porto,
Mestrado em Ciência da Informação, joanasousarodrigues.14@gmail.com*

RESUMO Embora ainda não seja vastamente estudado no domínio da Ciência da Informação, o documento fotográfico constitui um desafio para as práticas de representação da informação nas estruturas documentais, nomeadamente nos arquivos. As novas tecnologias e a presença assídua de dispositivos de captura de imagem podem levar à banalização do tratamento da fotografia e é nesse sentido que devem ser assegurados mecanismos de gestão de informação que permitam, não só, a descrição fiel do documento fotográfico, mas que incutam ferramentas de trabalho que potenciem uma recuperação da informação célere e eficaz. O estabelecimento de uma política de controlo de autoridade para documentos fotográficos pode ser o ponto de partida para a gestão uniforme e coerente da informação inerente aos documentos desta tipologia documental, seja através do estabelecimento de diretrizes para a criação de pontos de acesso, seja através da utilização consciente e exaustiva das fontes de informação que potenciarão uma descrição completa e representativa. Os arquivos, como detentores de informação, comportam um papel fundamental, não só na preservação da memória que envolve o documento fotográfico, mas também na sensibilização e consciencialização das questões inerentes ao controlo de autoridade, devendo ser construtores de mentalidades que sustentem o tratamento de informação criterioso e unívoco.

PALAVRAS-CHAVE *arquivo fotográfico, controlo de autoridade, documento fotográfico.*

ABSTRACT Although it is not widely studied in the field of the Information Sciences, the photographic document constitutes a challenge to the practices of representing information in documental structures, namely in archives. New technologies and the constant presence of devices that capture images can lead to the trivialization of photo treatment and, for that reason, information management mechanisms must be guaranteed so they can allow not only the exact description of the photographic document but also instill work tools that enhance the recovery of information in a fast and effective way. The establishment of an authority control policy for photographic documents can be the starting point for the uniform and coherent management of the information inherent in the documents of this documentary typology, either through the establishment of guidelines for the creation of access points or through conscious and exhaustive use of the sources of information that will enable a complete and representative description. The archives, as information holders, play a fundamental role, not only in preserving the memory involved in the photographic document, but also in the awareness and awareness of the issues inherent to the control of authority, and should be constructors of mentalities that sustain the information treatment criterious and univocal.

KEY-WORDS *authority control, photographic archive, photographic document.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

As dinâmicas inerentes à gestão de documentos fotográficos através de mecanismos de controlo de autoridade em instituições de domínio arquivístico levantam inúmeras questões, sobretudo porque esta tipologia documental ainda não é vastamente estudada, circunstância essa que motiva o escasso entendimento destas áreas de estudo e o seu cruzamento. É neste sentido que surge este trabalho de investigação que visa o estudo do documento fotográfico através de uma abordagem alicerçada na Ciência da Informação, com o intuito de alcançar o papel que estes devem desempenhar em instituições de conhecimento, como arquivos. Além disso, este projeto pretende elevar a importância do controlo de autoridade em variadíssimas vertentes da gestão da informação e da sua aliança com os sistemas de informação.

O projeto desenvolveu-se através do estudo de um conjunto de fotografias pertencentes a uma instituição do domínio universitário que continha um arquivo destinado ao acervo fotográfico que foi constituído através da produção noticiosa. Com este estudo foi possível levantar inúmeras questões, nomeadamente no que diz respeito às práticas de informação instauradas quando se trata de imagem.

Sendo o controlo de autoridade um dos pilares deste estudo, foi importante perceber como é que este se enquadrava na dinâmica das instituições arquivísticas e se, aquilo que veiculava, estava de acordo com a normalização existente. Só através desta análise seria possível chegar a conclusões consistentes sobre o tratamento da informação em questão e propor melhorias que promovessem a devida representação da informação que, por sua vez, potenciasse a recuperação de cada um dos registos fotográficos.

CONTEXTUALIZAÇÃO

1. O DOCUMENTO FOTOGRÁFICO

O documento fotográfico permite a confirmação do real, é, no fundo, uma prova do que aconteceu. A fotografia cresceu, não é estanque e o seu crescimento permitiu a conquista da liberdade de expressão e o respeito pela linguagem visual. Barthes (2012, 75) afirma que o objetivo do documento fotográfico sempre foi, acima de tudo, garantir a diversidade na comunicação com o outro, por isso, não é estranho afirmar que a fotografia é um órgão de comunicação tão capaz como outro qualquer.

Segundo Serén (2013, 183-184), o documento fotográfico deve ser visto como um objeto social, pertencente a uma sociedade de informação e comunicação, pois esta caracteriza-se pela sua capacidade de registo, em que são anotadas todas as transações de uma sociedade transversal e global. Serén afirma ainda que, tal como os restantes tipos de documentos, a fotografia é um ato social que exige a presença daqueles que são chamados os seus elementos básicos. Estes são a intenção, relacionada com a vontade, a expressão, que está intrinsecamente ligada à maneira de exteriorizar uma ideia e a inscrição, que é importante para manter os dois elementos anteriores.

No que diz respeito à posição que o documento fotográfico ocupa no tratamento documental, Henrique (2010, 20) vê-a como incerta, prova disso são os lugares duvidosos onde alguns destes documentos são acondicionados, descurando o seu tratamento. Citando esta autora, «considerar este documento enquanto ato fotográfico equivale a considerar este documento como integrante em um processo

documental (ou processo de registo) que se pretende recuperar.». A autora chega mesmo a afirmar que é através dessa recuperação, que é possível encontrar o contexto do documento, constituindo, assim, como uma propriedade do processo arquivístico.

2. O LUGAR DO DOCUMENTO FOTOGRÁFICO NOS ARQUIVOS

Boccatto e Fujita (2006, 85) desenvolveram a problemática do lugar do documento fotográfico nos arquivos e a sua relação com as propriedades e os comportamentos da Ciência da Informação. Segundo estas autoras, o documento fotográfico deve ser visto «como fonte de informação para a geração de conhecimento», uma vez que este cumpre um papel social fundamental que assegura a multiplicação de novos conhecimentos.

Segundo Barthes (2012, 35), a convergência dos meios de comunicação social não permitiu que, até à data, o documento fotográfico adquirisse um papel de destaque nos arquivos. Na verdade, a crescente produção fotográfica desencadeou a criação de agências fotográficas que se consideravam capazes de fazer a gestão de informação fotográfica que produziam, o que, na realidade, nem sempre, ou mesmo em poucos casos acontecia. Para o autor, a coerência visual deve ser proporcional à coerência da sua descrição e é esse o papel que os arquivos devem desempenhar, um papel de organização da informação eficaz, a curto e a longo prazo.

Lacerda (2012, 284) reforça a ideia de que é necessário refletir sobre o aparente domínio de documentos de carácter textual nos arquivos e de que forma são realizados o tratamento, a classificação e a descrição dos documentos fotográficos.

A análise documental de documentos fotográficos tem como finalidade a representação e, com isso, a recuperação de conteúdos através de atributos. Com isto surge a análise de conteúdo do documento fotográfico que é composto por três etapas, nomeadamente a leitura do conteúdo fotográfico, a síntese desse mesmo conteúdo e a sua representação.

Citando Boccatto e Fujita (2006, 98) «a análise documental é um processo instrumental» e é nesse sentido que a Ciência da Informação, através do seu carácter «multidisciplinar pós-moderna», deve incentivar que sejam definidas as orientações que levem ao acesso competente de imagens nos sistemas e nas unidades de informação. É seguindo esta ordem de ideias que surge o profissional e informação que vai «representar o conteúdo da imagem fotográfica para torná-la acessível – socialização do conhecimento – ao usuário».

3. O ARQUIVO FOTOGRÁFICO

Os arquivos fotográficos desempenham um papel fundamental na salvaguarda do património físico e intelectual dos documentos fotográficos, contribuindo para a sua gestão, que inclui o estudo, o tratamento, a organização e a preservação das fotografias.

A descrição de documentos de arquivos fotográficos contempla orientações particulares que auxiliam «uma maior transparência e enquadramento descritivo destes documentos.» (Orientações para a Descrição Arquivística 2011, 31).

Tendo em conta as Orientações para a Descrição Arquivística (ODA) (2011, 76-78), os níveis de descrição têm como principal objetivo a identificação dos níveis de organização arquivística de determinada unidade de descrição, seja ela de que domínio seja. Estes níveis ajudam a entender o posicionamento das unidades de descrição na correspondente hierarquia.

São os níveis de descrição que potenciam a explicitação das relações que cada uma das unidades de descrição têm com as restantes, de ajuste com as regras da descrição multinível. Assim sendo, uma dada unidade de descrição pode depender hierarquicamente de algum outro nível superior, assim com, níveis hierarquicamente inferiores, podem depender desta. Todavia, é de salientar que não é obrigatório que as diversas massas documentais reflitam todos os níveis de descrição.

4. O CONTROLO DE AUTORIDADE

O controlo de autoridade é, cada vez mais, uma prática de consistência no que diz respeito à criação fiável dos pontos de acesso à informação, através da relação entre um ficheiro de autoridade e o conjunto de registos de autoridade que a si dizem respeito.

Barbara Tillett (2003) desenvolveu um artigo intitulado de “Authority control: state of the art and new perspectives” em que aborda a posição que o controlo de autoridade tem no ambiente *web*. Para a autora, a aplicação do controlo de autoridade neste ambiente traz precisão para as pesquisas, possibilita uma navegação mais eficaz e fornece aos utilizadores um resultado equivalente ao esperado. O controlo de autoridade não pode ser dissociado das ferramentas *online*, como os diretórios, dicionários biográficos, bases de dados, repositórios, serviços de indexação, entre outros.

Ainda Tillett (2000), explicita de que forma é que a *web* sairá favorecida com o controlo de autoridade. A autora afirma que a *web* é caótica, carregada de informação dispersa e é por esse motivo que os utilizadores nem sempre conseguem obter aquilo que realmente esperam com as suas pesquisas. Com a utilização do controlo de autoridade essa variação de resultados pode ser anulada e, além disso, o conteúdo recuperável pode ser muito mais que uma mera descrição de dado documento. Também Borbinha (2003, 7-8) vê o controlo de autoridade como um desafio, uma vez que este proporciona uma evolução natural de entidades bem definidas, com interfaces estabelecidas e que auxiliará na geração de um novo conceito informacional que assentará num ambiente mais dinâmico e executivo. A realidade é que o registo de autoridade tem uma relação direta com os sistemas integrados e com os registos bibliográficos que estes contêm, uma vez que é a partir dos ficheiros de autoridade que o processo de pesquisa acontece, potenciando que sejam apresentados os resultados no momento em que o utilizador questiona a base de dados.

5. A IMPORTÂNCIA DA LINGUAGEM CONTROLADA NO CONTROLO DE AUTORIDADE DE IMAGENS

É necessário distinguir aquela que é a linguagem natural, usada pelos autores dos documentos, caracterizando-se como sendo uma linguagem simples e livre, sem restrições e controlo e a linguagem controlada, que deve ser realizada conscientemente e tem um objetivo muito claro: o armazenamento e a recuperação da informação.

Assim sendo, pressupõe-se que o vocabulário utilizado seja alvo de controlo e restrito a um concentrado número de termos. Na verdade, não existe um vocabulário específico para a representação de conceitos

e embora este deva ser, efetivamente, controlado, não pode afastar-se do conteúdo original do documento em análise. Assim sendo, a criação de um léxico controlado e estruturado implica que sejam incluídas relações semânticas entre os termos de indexação definidos.

Para Gorman (2003, 13) o controlo vocabular é vital para o controlo de autoridade, seja para o registo de formas padronizadas de pontos de acesso, para garantir a agregação de todos os registos para o mesmo ponto de acesso, para documentar as decisões tomadas e as fontes do ponto de acesso, ou para habilitar registos de catálogo padronizador.

Também Lancaster afirma que o uso de um vocabulário controlado é essencial no processo de recuperação, chegando mesmo a afirmar que o ciclo informacional tem que incluir o processo de definição de termos e conceitos e que essa seleção não pode ser deixada a cargo de um mero utilizador, cabe sim ao profissional de informação averiguar os termos que condizem com o documento em causa e, posteriormente, utilizar meramente terminologia que saiba que será capaz de recuperar esse documento, mesmo que a pesquisa seja pouco profunda.

METODOLOGIA

A metodologia utilizada será o método qualitativo de investigação, através do estudo de caso, nomeadamente a pesquisa exploratória. A pesquisa exploratória tem o intuito de melhorar a familiarização do investigador com o caso em estudo, assim, este método servirá para que seja feita uma investigação preliminar sobre os temas a abordar na dissertação, para que seja possível refletir sobre estes e melhorar a sua compreensão. O controlo de autoridade de documentos fotográficos em arquivos, sejam eles físicos ou digitais, ainda não é vastamente estudado, por isso, com a pesquisa exploratória pretende-se que sejam geradas ideias e hipóteses que complementem este tema e que auxiliem o trabalho a desenvolver.

A investigação qualitativa assentou sobre as fontes primárias no domínio dos materiais empíricos. Para isso foram realizadas entrevistas semiestruturadas aos responsáveis do Centro Português de Fotografia e do Centro de Documentação da Global Media Group. Com estas entrevistas e visitas pretendeu-se dar credibilidade e profundidade ao estudo.

Foi, ainda, desenvolvido trabalho de pesquisa em livros, artigos de jornal, revistas científicas e outros documentos já publicados. Este método de investigação, segundo estes alicerces, gerará a fundamentação necessária ao trabalho, pois o trabalho de pesquisa em fontes de informação credíveis e especializadas contribuirá para uma maior verosimilidade e consistência ao trabalho.

RESULTADOS

De modo a entender de que forma é que o documento fotográfico pode ser, significativamente, valorizado através de práticas de controlo de autoridade, foi imperativa definição de soluções que visassem, não só, a uniformização de critérios, mas também, a representação real e fidedigna de cada um dos documentos tratados, incentivando, assim, uma descrição fiel e uma recuperação célere proveitosa.

Uma dessas soluções passa pela criação de pontos de acesso que possibilitem a relação entre o documento fotográfico e o sistema de informação que o irá receber. A construção de pontos de acesso é essencial para a prática do controlo de autoridade, uma vez que permite ao utilizador aceder e recuperar os recursos bibliográficos sem equívocos e perdas de tempo, através de um controlo vocabular adequado e representativo do que está a ser descrito. É nessa medida que a existência de orientações específicas para a construção dos pontos de acesso é essencial para a eliminação de desvios, uma vez que a uniformização é a única forma de garantir a coerência na descrição de registos.

A seleção dos pontos de acesso deve ser criteriosa e seguir determinadas diretrizes, sendo que estas alicerçam-se no seguinte:

- **seleção dos termos:** selecionar termos que tenham por base as fontes de referência que estão inerentes à área temática em questão, verificar termos sinónimos, termos homónimos, termos diferentes com significado igual, grafias variantes, abreviaturas e acrónimos, formas alternativas de escrita, termos diferentes em línguas diferentes e termos composto e selecionar termos que se encontrem no singular, visto tratar-se de nomes de pessoas,
- **alteração terminológica:** quando existirem alterações ao nível da terminologia na língua natural é necessário refletir se são importantes para efeitos de indexação. Caso isso se verifique, estas devem refletir-se na linguagem documental, para, dessa forma, assegurar o acesso através da utilização dos novos termos que serão mais apropriados que os anteriores,
- **língua:** devem ser utilizados termos em língua portuguesa, salvo:
 - se não existir uma designação portuguesa para dado conceito,
 - se este for mais conhecido pelo público português por um termo noutra língua, estando nessa forma em fontes de língua portuguesa,
 - se a área temática do conceito exigir que o termo seja expresso noutra língua, devido a uma terminologia científica própria.
- **estrutura física:** devem ser utilizada uma estrutura constante, ou seja, o número de partes do ponto de acesso e a ordem por que estes se apresentam deve ser invariável.

Que os pontos de acesso são elementos necessários para que a recuperação da informação, na medida em que constituem instrumentos de pesquisa que se viabilizam com o processo de descrição, é uma verdade que não pode ser discutida. No entanto, a forma como estes são feitos nem sempre segue as mesmas diretrizes, prova disso são as, ainda, discrepâncias, existentes quanto se fala do tratamento em bibliotecas ou arquivos. Se por um lado, falando das bibliotecas, temos recomendações estabelecidas por organismos como a Biblioteca Nacional, a International Federation of Library Associations ou mesmo a Associação de Bibliotecários, Arquivistas e Documentalistas, por outro, falando dos arquivos, temos princípios dados pelo Conselho Internacional de Arquivos e pela Direção Geral de Arquivos. Não é possível considerar que algum desses princípios estejam incorretos, apenas é possível afirmar que a diversidade propicia dúvida na altura da escolha dos pontos de acesso e, se assim o é, é possível que surjam variações que fogem dos objetivos que, numa primeira instância, foram estabelecidos.

Outro fator passível de ser analisado prende-se com os profissionais destacados para as tarefas de descrição bibliográfica. Estes, na larga maioria das vezes, estabelecem uma política interna de criação de pontos de acesso, com vista a satisfazer as necessidades da própria instituição ou então para criar compatibilidade entre esta e os recursos humanos e não humanos que possuem. Na realidade, este método de trabalho não deve ser descartado, uma vez que todas as decisões tomadas em ambiente organizacional devem ter como base as necessidades da própria organização, bem como da sua visão e missão. Todavia, esta abordagem deve ter sempre em conta princípios normativos já instituídos por organismos competentes. Falamos então de adaptação, não de invenção.

Que os pontos de acesso normalizados constituem uma solução efetiva e considerável para o controlo de autoridade de documentos fotográficos, é um facto irrefutável. Todavia, a sua existência não deve descartar a utilização de outros mecanismos e ferramentas de trabalho. É neste sentido que surge uma outra solução que trará um acréscimo de valor ao registo de autoridade que está inerente e é parte substancial do controlo de autoridade: a criação de notas de apoio ao registo. Estas notas podem ser de domínios diversos, tais como notas biográficas, notas geográficas, notas históricas, entre outras.

As notas de apoio são fundamentais, pois, resultando de trabalho de pesquisa, tem como intuito expor e explicar, de uma forma interessante, determinado assunto ou história de vida. Estas deve conter informações que darão a conhecer pormenores importantes que poucos conhecem e têm como base dois tipos de fontes de informação, nomeadamente documentos publicados e conversas/entrevistas. A primeira dará credibilidade ao estudo, priorizando-se fontes com relevância, a segunda é uma forma mais direta de recolha de informação e concederá profundidade à nota. As fontes de informação constituem uma ferramenta importante na qualidade da nota de apoio, uma vez que uma pesquisa em fontes pouco cuidadas e precisas proporcionará uma nota, também ela, pouco cuidada e com imprecisões.

Por fim, apresentar uma terceira solução que apoiará o registo de autoridade de um documento fotográfico: o estabelecimento de um identificador de descrição único e para cada documento tratado. Este identificador deve seguir as orientações dadas pela norma ISAAR (1995) que afirma que o código de identificador deve ser inequívoco, capaz de identificar o país, a instituição e a pessoa. Também a ISADG (2002) declara que estes três elementos são essenciais para a identificação do registo, pois permitem identificar claramente a unidade de descrição estabelecendo, também, uma ligação com a descrição que por si é representada. Quando não existe a referência a uma pessoa este identificador deve incluir apenas os primeiros dois elementos.

As três soluções apresentadas constituirão, assim, uma política de controlo de autoridade que é essencial na determinação das escolhas técnicas a tomar, bem como na escolha de instrumentos a instituir, de modo a ajustar estas aos princípios e necessidades do serviço onde atuam, com o objetivo de tornar eficaz a representação e a recuperação da informação. Esta política englobará, portanto, todos os registos de autoridade, sendo que, em cada um deles, caberá uma zona de descrição, destinada aos pontos de acesso; uma zona de descrição, destinada às notas de apoio e uma zona de controlo destinada ao identificador de descrição e às fontes de informação utilizadas.

CONCLUSÕES

Vivemos numa época em que a produção fotográfica se tornou uma prática de rotina, através das novas tecnologias (como dispositivos móveis) a fotografia está ao alcance de qualquer um, correndo-se o risco de banalizar este tipo de documentos. Vivendo naquela que é chamada de Era da Informação e que vem alinhada com o avanço tecnológico, é notório que as tecnologias, os métodos de impressão e digitalização e a imposição do digital trouxeram enormes desafios para o documento fotográfico. Este facto não deve ser visto como um fator negativo, mas sim como a possibilidade de gerar novas tendências.

Na verdade, não existem muitos estudos sobre a questão do controlo de autoridade de documentos fotográficos nos arquivos, sejam eles físicos ou digitais. Isto acontece, porque a esmagadora maioria de documentos produzidos e que estão à guarda de arquivos são de carácter textual, deixando, portanto, a esta tipologia documental para segundo plano.

Um acervo fotográfico pode comportar parte da identidade de um organismo ou instituição, ele carrega património físico e intelectual de importância considerável e pode constituir um meio para a valorização daqueles que o gerem. Seja qual for o seu âmbito, um acervo deve ir ao encontro das necessidades culturais, educativas, lúdicas e informacionais da comunidade a que se destina, mesmo que os seus destinatários tenham um perfil heterogéneo.

Com o estudo realizado foi possível verificar a variedade de critérios que existem, nomeadamente quando o assunto é controlo de autoridade. As diferenças são evidentes quando se fala de instituições de tratamento de informação em biblioteca e em arquivo. Na realidade, esta diversidade e instabilidade provocam desvios e pouca afirmação dos princípios que estão na base do controlo de autoridade.

Neste seguimento surgem as Resource Description and Access (RDA) que se dispõem a iniciar a mudança, começando por alterar a terminologia a adotar. O objetivo passará por facilitar a descrição e posterior compreensão e recuperação da informação, eliminando e simplificando termos dúbios e agregando termos que pertençam à mesma classe. Com isto deixará de existir informação redundante e que muitas vezes traz dúvidas ao utilizador, além disso deixarão de existir termos diferentes conforme a tipologia documental que retratassem, ou seja, todos os tipos de documentos são vistos como iguais do ponto de vista terminológico.

Outro exemplo proposto e que incitará uma mudança significativa são os nomes dados aos pontos de acesso que, segundo as RDA, são designados assim mesmo, como “pontos de acesso”. É comum vermos designações muito diferentes, como “cabeçalhos” e “entradas” nas RPC (2000), ou “Forma autorizada de nome” nas ODA (2011), contudo, todas estas designações se concluem numa só e se assim o é porque não uniformizar o vocabulário a empregar e facilitar o processo de tratamento da informação, acabar com a diversidade de critérios utilizados e extinguir as formas paralelas de representação?

Em suma, destacar os profissionais de informação que devem trabalhar juntos, a sua formação deve ser integrada e coincidente, tal como os seus métodos de trabalho, uma vez que, se é objetivo geral criar um sistema normalizado, não deve haver espaço para variações, mas sim para a promoção da investigação, aprendizagem, partilha de conhecimentos e do estabelecimento de um conjunto de princípios lógicos a seguir.

REFERÊNCIAS BIBLIOGRÁFICAS

- Barthes, R. (2012). *A câmara clara*. Lisboa: Edições 70.
- Biblioteca Nacional de Portugal. (2016). *RDA: Resource Description & Access: Introdução ao RDA. Seminários Profissionais Biblioteca Nacional de Portugal*. Lisboa: Biblioteca Nacional de Portugal;
- Biblioteca Nacional de Portugal. Serviços. [Em linha]. [Consult 20 abril 2017] Disponível em: http://www.bnportugal.pt/index.php?option=com_content&view=article&id=106&Itemid=53&lang=pt;
- Bocato, V. R. C., & Fujita, M. S. L. (2006). Discutindo a análise documental de fotografias: uma síntese bibliográfica. *Cadernos de Biblioteconomia, Arquivística E Documentação*, 2, 84–100;
- Borbinha, J. (2003). Authority control in the world of metadata. In *International Conference on Authority Control: definition and international experiences*. Florença;
- Centro Português de Fotografia. [Em linha]. [Consult 25 janeiro 2017] Disponível em: <http://digitarq.cpf.arquivos.pt/details?id=1184961>;
- Conselho Internacional de Arquivos Internacionais. (1998). *ISAAR (CPF): Norma internacional para os registos de autoridade arquivística relativos a instituições, pessoas singulares e famílias*. Lisboa: Instituto dos Arquivos Nacionais, Torre do Tombo e Associação Portuguesa de Bibliotecários, Arquivistas e Documentalistas;
- Direção Geral de Arquivos, & Centro Português de Fotografia. (2007). *Workshop Descrição Arquivística e Arquivos de Fotografia*. Porto: Direção Geral de Arquivos, Centro Português de Fotografia;
- Direção Geral de Arquivos. Grupo de Trabalho de Normalização da Descrição em Arquivo. (2011). *Orientações para a Descrição Arquivística* (3ª ed.). Lisboa: Direção Geral de Arquivos;
- Gorman, M. (2003). Authority control in the context of bibliographic control in the electronic environment. In *International Conference on Authority Control: definition and international experiences*. Florença;
- Henrique, S. I. D. P. (2010). *O Lugar da Fotografia nos Arquivos: Uma Proposta de Reavaliação*. Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa;
- International Federation of Library Associations. (2016). Statement of International Cataloguing Principles. In B. Tillett & A. L. Cristán (Eds.), *IFLA Meetings of Experts on an International Cataloguing Code* (p. 20). Munique: International Federation of Library Associations;
- Lacerda, A. L. de. (2012). A fotografia nos arquivos: produção e sentido de documentos visuais. *História, Ciências, Saúde*, 19(1), 283–302;
- Quivy, R., & Campenhoudt, L. Van. (2008). *Manual de Investigação em Ciências Sociais* (5ª ed.). Lisboa: Gradiva;

Serén, M. do C. (2013). O Documento Fotográfico: Da Mediação Cultural à Mediação Técnica. *Revista do CITCEM*;

Tillett, B. (2000). Authority control on the web. In *Bicentennial Conference on Bibliographic Control for the New Millennium, sponsored by the Library of Congress Cataloguing Directorate*;

Tillett, B. (2003). Authority control: state of the art and new perspectives. In *International Conference on Authority Control: definition and international experiences*. Florença.



PROPOSTA METODOLÓGICA PARA MAPEAMENTO E ANÁLISE DE CONFIGURAÇÕES DE DISCURSOS SOCIAIS: AS REPRESENTAÇÕES DISCURSIVAS DE DILMA ROUSSEFF

Ludmila Salomão Venâncio¹, Renato Rocha Souza²

¹Universidade do Estado de Minas Gerais – UEMG, ludmila.salomao@uemg.br

²Fundação Getúlio Vargas – FGV, renato.souza@fgv.br

RESUMO Propõe-se uma abordagem metodológica para compreensão do universo semântico e dos posicionamentos dos sujeitos em discursos sociais, evidenciando as marcas argumentativas presentes em seus pronunciamentos. Para tanto, analisam-se textos proferidos pela ex-presidente da República do Brasil, Dilma Rousseff. A metodologia consiste em uma análise lexicométrica e uma análise de redes terminológicas que, viabilizadas pela identificação e pela extração automática das unidades lexicais mais frequentes em cada texto e pela construção de redes, apontaram diversas apropriações contextuais e o estabelecimento de significados contingentes diacronicamente. Desse modo, evidencia-se que a formalização explicativa, empreendida pela abordagem proposta, possibilita compreender melhor as configurações e remanejamentos das trocas discursivas em domínios dinâmicos.

PALAVRAS-CHAVE *representação e organização do conhecimento, análise lexicométrica, discursos sociais, discurso político, redes terminológicas.*

ABSTRACT We propose an approach to understand the semantic field and the subjects' positions in social discourses, evidencing the argumentative marks present in their pronouncements. To do so, we analyze texts delivered by the former president of the Republic of Brazil, Dilma Rousseff. The methodology consists of a lexicometric analysis and also an analysis of terminological networks that, enabled by the identification and automatic extraction of the most frequent lexical units in each text and the construction of networks, could lead us to the conclusions that several contextual appropriations and the establishment of diachronically contingent meanings. Thus, it is evident that the explanatory formalization, undertaken by the proposed approach, enables to better understand the configurations and re-assignments of discursive exchanges in dynamic domains.

KEYWORDS *representation and organization of knowledge, lexicometric analysis, social discourses, political discourse, terminology networks.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Ao longo dos últimos anos, a pesquisa da linguagem vem se tornando tema central no campo da representação e organização do conhecimento. Muitos dos estudos nesse tema incluem o uso de diversas técnicas automatizadas de tratamento de corpora textual como, por exemplo, o uso de processamento (automático) da linguagem natural (PLN). O PLN é uma subárea tanto da ciência da computação, quanto da linguística, e consiste no desenvolvimento de modelos computacionais para a consecução de tarefas que dependem de informação registrada em formato textual. Dentre estas tarefas, temos, como exemplo, a tradução e interpretação de textos, a busca de informações em documentos e modelagem de interfaces homem-máquina. De acordo com Marciniak *et al* (2009), a pesquisa em PLN está voltada, essencialmente, a três aspectos da comunicação em língua natural: som (fonologia); estrutura (morfologia e sintaxe) e significado (semântica e pragmática). Na área da ciência da informação, o PLN auxilia a processos de indexação de documentos, análise de assunto e organização de informações, utilizando, para tal, técnicas variadas, como a extração de entidades nomeadas e identificação de estruturas semânticas, como os sintagmas nominais, para a indexação e recuperação da informação. Outra das muitas ferramentas que vêm auxiliando as pesquisas em linguagem são os sistemas de organização do conhecimento (SOC/KOS). Como SOCs, consideramos todos os tipos de esquemas para organização de informação e promoção da gestão do conhecimento (Hodge, 2000) sendo os mesmos representações do conhecimento baseadas em conceitos com diferentes graus de detalhe dentre os relacionamentos entre estes (Souza et al, 2012). SOCs nos auxiliam a mapear e guiar a extração de informações em domínios específicos do conhecimento e na adoção de esquemas representacionais ontológicos em sistemas de organização e recuperação da informação. Ferramentas como os tesouros e as ontologias auxiliam processos de representação de relacionamentos conceituais e semânticos; uma proporção significativa dos estudos bibliométricos, permitindo também que frequência de termos em documentos e construção de índices sejam menos sujeitos a variações lexicais e de sinonímia. Finalmente, a união das técnicas de PLN com os SOCs permite a construção de processos de indexação automática realizada com base em termos escolhidos, dentre aqueles que se mostram relevantes em um determinado domínio.

Em todos esses casos, a linguagem torna-se o fundamento básico para a proposição de metodologias e construção de ferramentas para a representação e a organização do conhecimento, atentando-se para seu emprego pelos diferentes sujeitos sociais mobilizados em práticas discursivas diversas (Venâncio, 2014).

Neste contexto, a presente comunicação tem como objetivo apresentar uma proposta metodológica para mapeamento e análise do universo semântico e dos posicionamentos dos sujeitos implicados no domínio político eleitoral, evidenciando as marcas argumentativas presentes em seus pronunciamentos diacronicamente. Utilizando-se técnicas oriundas do PLN, como a representação estatística de documentos através do cálculo de frequências de palavras; e o uso de SOCs para a identificação de termos-chave no campo do discurso, tal proposta integra as duas dimensões do discurso, quais sejam, uma *linguístico-enunciativa* e outra *discursivo-situacional*.

A dimensão *linguístico-enunciativa* refere-se à organização do discurso em termos de suas propriedades lexicais e de suas relações semânticas, em que se determina a encenação da qual participam os sujeitos, constantemente regulada no próprio processo discursivo, a partir das restrições da dimensão *discursivo-situacional*. Já a dimensão *discursivo-situacional* configura, assim, como a instância em que se produz o acontecimento discursivo, coordenado por uma relação contratual que os sujeitos estabelecem entre

si, como reconhecimento das regras e convenções que organizam o direito às suas enunciações em um dado domínio social, a exemplo do domínio político (Charaudeau, 1996). Assim, a dimensão situacional é vista como base da dimensão linguística e, ao mesmo tempo, constituída por esta.

A análise de um gênero discursivo pressupõe, antes de tudo, a compreensão do conceito do gênero em questão e dos parâmetros que o definem. Emediato (2003, p. 64) compreende o gênero discursivo como um tipo situacional, “*como um produto de uma interação complexa entre as condições de produção dos textos e das variantes textuais e as competências que estariam na base da produção dos discursos sociais*”. Subjaz a esse entendimento o conceito de ancoragem social do discurso, que preconiza que não há produção de texto que não se encontre atrelada às determinações impostas pelo domínio de prática social onde ela se realiza, seja esse religioso, político, jurídico, científico, educativo ou midiático (Charaudeau, 2006; Emediato, 2003). O domínio de prática social surge, então, como “*um determinante maior da natureza comunicacional do discurso*” (Emediato, 2003, p. 65).

Todo ato de linguagem se realiza, dessa forma, em uma situação de comunicação, que determina o espaço da troca linguageira e suas condições de produção, trazendo limitações no seu processo de encenação. O conceito de *situação de comunicação* remete, necessariamente, ao conceito de *contrato de comunicação*. Todo domínio de comunicação propõe a seus parceiros um conjunto de condições nas quais se realiza qualquer ato de linguagem e que normatizam as trocas comunicativas, de modo que os parceiros possam se entender sobre o que constitui a expectativa da troca – uma espécie de contrato de reconhecimento. Tal contrato permite, assim, aos parceiros de uma troca linguageira reconhecerem um ao outro com traços identitários que os definem como sujeitos desse ato – identidade; reconhecerem o objetivo do ato de fala – finalidade; entenderem sobre o que constitui o objeto temático da troca – propósito, e considerarem a relevância das coerções materiais que determinam esse ato – circunstâncias.

No nosso caso, o domínio político, mais especificamente o domínio político eleitoral, determina os rituais linguageiros que podem ser construídos no espaço da política. São exemplos desses rituais a configuração discursiva dos debates instituídos nesse domínio, os papéis ou *scripts* que podem ser desempenhados pelos sujeitos envolvidos de acordo com as expectativas dos seus interlocutores, as apresentações nas reuniões com os comitês de campanha e convenções do partido, a estrutura das declarações na mídia e em comícios, em uma entrevista ou em uma propaganda de campanha e a maneira de abordar temas que constituem as maiores preocupações do eleitorado.

METODOLOGIA

O corpus deste estudo compreende um conjunto de textos proferidos por Dilma Rousseff, ex-presidente da República do Brasil, desde a fase de sua candidatura no primeiro pleito até seu *impeachment*, no segundo mandato. Para efeito de análise, o conjunto de textos abordado foi dividido em séries distintas que possibilitaram a compreensão das diversas apropriações semânticas contextuais nos discursos de Dilma e do estabelecimento de significados contingentes diacronicamente.

A primeira fase da metodologia consistiu na realização de uma análise lexicométrica viabilizada pela identificação e pela extração automática das unidades lexicais (substantivos, adjetivos, verbos, pronomes, sintagmas nominais e nomes próprios) mais frequentes em cada texto analisado, com o

auxílio do *Parser Palavras*¹. A lexicometria é um método estatístico comparativo, que consiste no levantamento da distribuição das unidades lexicais e em comparar ou opor os diversos usos dessas unidades em fragmentos do discurso, com a utilização de ferramentas computacionais. Ela abrange, assim, duas operações principais: (i) a escolha, após a divisão da cadeia textual em unidades que podem ser estudadas; e (ii) a combinação, a comparação ou a oposição das formas textuais identificadas nos discursos que compõem o conjunto da enunciação analisado (Bonnafous; Tournier, 1995; Charaudeau; Maingueneau, 2008).

Na segunda fase foram mapeadas as redes terminológicas – aquelas formadas por termos utilizados pelos sujeitos em processos interacionais em determinado domínio -, que representam a associação entre os termos utilizados por Dilma em sua estratégia argumentativa na dinâmica do debate político. Tais redes são plotadas por meio de um método de análise de coocorrência e avaliadas por meio da contextualização do emprego dos termos e das conformações que elucidam a dinâmica da rede.

Compreendida como um artifício metodológico, a abordagem de redes sociais engloba um conjunto de teorias, conceitos, métodos e modelos das ciências sociais, permeado por diversas perspectivas disciplinares e epistemológicas. A análise de redes sociais é uma abordagem focada na estrutura relacional, que visa a investigar de que forma as relações e os atributos dos atores influenciam o comportamento de cada componente na rede e do todo. Certas unidades básicas compõem as redes. Um ator é uma “unidade discreta individual, corporativa ou social coletiva” (Wasserman; Faust, 1999, p. 17). O *laço relacional* estabelece a ligação entre pares de atores e “são canais de transferência ou fluxo de recursos, sejam eles materiais ou não materiais” (Wasserman; Faust, 1999, p. 4). Um dos objetivos mais importantes da análise de redes consiste na identificação de subgrupos coesos no interior de um conjunto social. Um *subgrupo* coeso é composto por um conjunto de atores que estabelecem relações fortes, intensas, diretas e frequentes (Lazega, 2007). A análise *centro-periferia*, proposta por Borgatti e Everett (2000), auxilia na identificação desses subgrupos. Tal análise pressupõe a existência de dois grandes grupos em uma rede: um grupo de atores fortemente relacionados, denominado de *centro*, e um grupo no qual os atores têm poucos contatos entre si, mas que se ligam ao centro da rede por alguns de seus membros, denominado *periferia*. Nesta região está concentrada a maior parte dos atores da rede. De maneira semelhante, a identificação de *cliques*, subconjunto de membros todos adjacentes uns aos outros que mantêm relações mais próximas entre si, possibilita conhecer quem pertence a um grupo, bem como os tipos e padrões de relacionamentos que os definem e os sustentam como um grupo (Lazega, 2007). Por fim, outra medida importante refere-se à *centralidade de grau* (*degree centrality*) que define a posição de um ator em relação às trocas e às comunicações na rede, considerando-se a quantidade de ligações diretas entre eles (Everett; Borgatti, 2005).

No contexto deste estudo, cada termo representa um ator na rede. Para o mapeamento da rede, foram desenvolvidas algumas rotinas em linguagem de programação *Python*. O algoritmo que possibilita tal mapeamento pode ser descrito pelas seguintes etapas: (i) calcula-se a frequência de cada termo no parágrafo para gerar uma lista de termos; (ii) determina-se a coocorrência entre cada par de termos – entendida como a presença dos dois termos no parágrafo; (iii) atribui-se um peso à coocorrência – esse peso se relaciona inversamente à distância entre os termos no mesmo parágrafo; (iv) sumarizam-se todas as coocorrências em parágrafos distintos com os pesos (atribuindo o peso calculado no passo

¹ Trata-se de um programa, desenvolvido por Eckhard Bick na *Southern University of Denmark*, que incorpora as regras para tratamento morfológico e sintático exigidas pela língua portuguesa, além de realizar a extração dos sintagmas diretamente nos documentos. Uma versão para avaliação está disponível em: <<http://visl.sdu.dk/visl/pt/>>. (Bick, 2000).

anterior a cada ocorrência) e pondera-se pela frequência dos termos no texto; e (v) gera a matriz de coocorrências de um texto a partir dos valores sumarizados.

Os dados gerados pelo algoritmo foram importados para o *software* de análise de redes sociais Ucinet (Borgatti; Everett; Freeman, 2002). Assim, determinaram-se diversas medidas para os termos, como graus de centralidade, coeficientes de agrupamento (cliques) e análise centro-periferia. Esses indicadores forneceram subsídios relevantes para analisar as redes e possibilitaram uma análise comparativa da variação do emprego da terminologia ao longo do tempo e das diferentes conformações assumidas pelas redes.

RESULTADOS

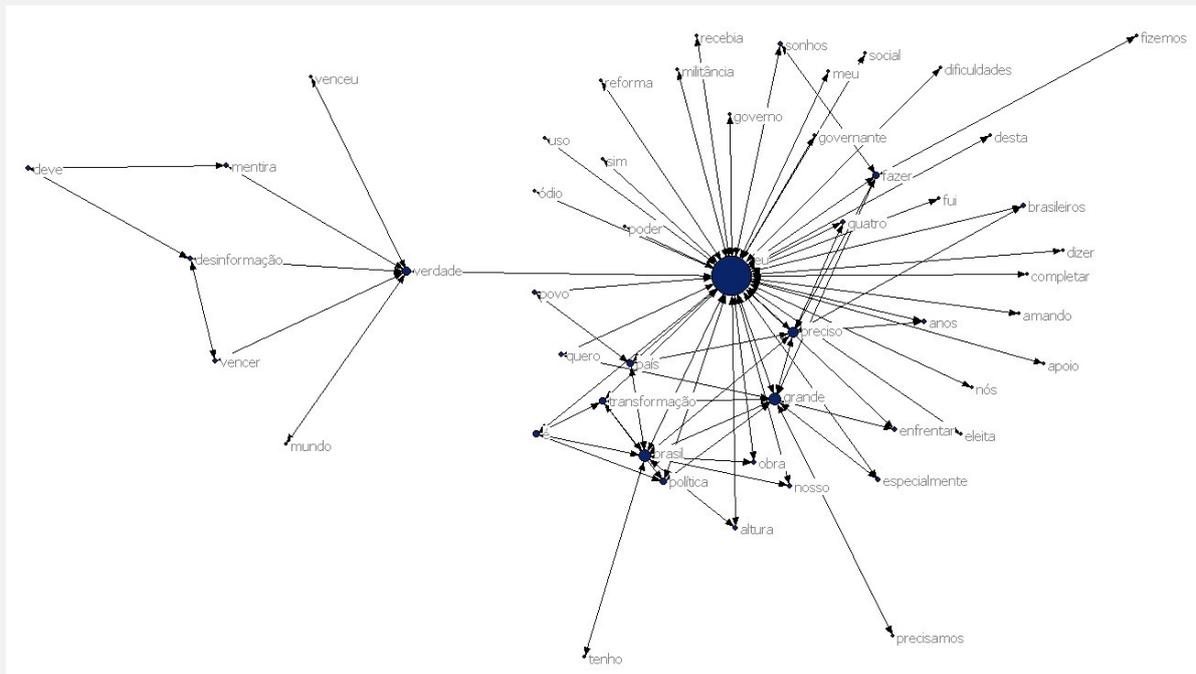
Ao se procurar caracterizar as representações discursivas, logo se percebe que o fazer político é sempre situado e apropriado pragmaticamente em processos interacionais. Desta forma, tentando ajustar o foco sobre a especificidade do domínio em análise, cabe uma descrição da cena enunciativa das eleições presidenciais brasileiras de 2010 e 2014, bem como da candidata Dilma Rousseff.

O cenário das eleições presidenciais de 2010 foi bastante distinto do encontrado nos pleitos anteriores. Pela primeira vez, após a redemocratização do país, Luiz Inácio Lula da Silva não seria lançado como candidato à Presidência da República pelo Partido dos Trabalhadores (PT), pois a Constituição Brasileira não permite o exercício de mais de dois mandatos consecutivos como Chefe do Executivo. Nove candidatos disputaram o cargo presidencial em 2010, dentre eles Dilma Rousseff.

Dilma, representante do governo, é economista e foi Ministra da Casa Civil na gestão do governo do Presidente Lula (2003-2009). Candidata pela coligação “*Para o Brasil Seguir Mudando*”, teve como vice o deputado Michel Temer. Foi candidata pela primeira vez à Presidência da República e indicada por Lula que antecipou o processo de discussão partidária usual nesse estágio. Foi responsável pela criação de programas estratégicos como o *Plano de Aceleração do Crescimento* (PAC) e o programa *Minha Casa, Minha Vida*. Eleita, em 2010, no segundo turno com 46,91% dos votos válidos, alcançou dois fatos inéditos na história política brasileira: ser a primeira mulher presidente do Brasil e exercer o terceiro comando nacional sucessivo de um mesmo partido.

Em 2014, o cenário político-eleitoral do Brasil era preocupante, havia aumentado a apreensão em relação às contas públicas e ao desempenho da indústria no Brasil, e a rejeição popular ao governo de Dilma só crescia. Dilma se reelegeu em 2014, no segundo turno, com 51,6% dos votos válidos e sofreu processo de *impeachment* em 2016, acusada de crimes de responsabilidade fiscal.

Como abordado anteriormente, os textos foram divididos em séries e analisados diacronicamente. O primeiro conjunto de textos refere-se aos discursos de candidatura das eleições de 2010 e 2014. Nessa fase da campanha, Dilma articulou seu discurso em torno das propostas de campanha, pautadas pela agenda dos fatos políticos, em que se busca ampliar o apoio no interior do partido, das alianças estabelecidas com outros partidos, dos movimentos sociais e de outros grupos da sociedade civil em geral. Destaca-se a utilização das formas *Brasil* e suas respectivas derivações, bem como termos correlatos como *país*, *povo*, *população* e *sociedade*, e a recorrência aos termos associados aos projetos e pré-programas de governo, como: *desenvolvimento*, *educação*, *trabalho*, *saúde*, *cultura* e *segurança*.



Obs.: Rede com 50 termos apresentados segundo medida de centralidade de grau

FIGURA 2: Rede terminológica – Discurso de candidatura 2º pleito

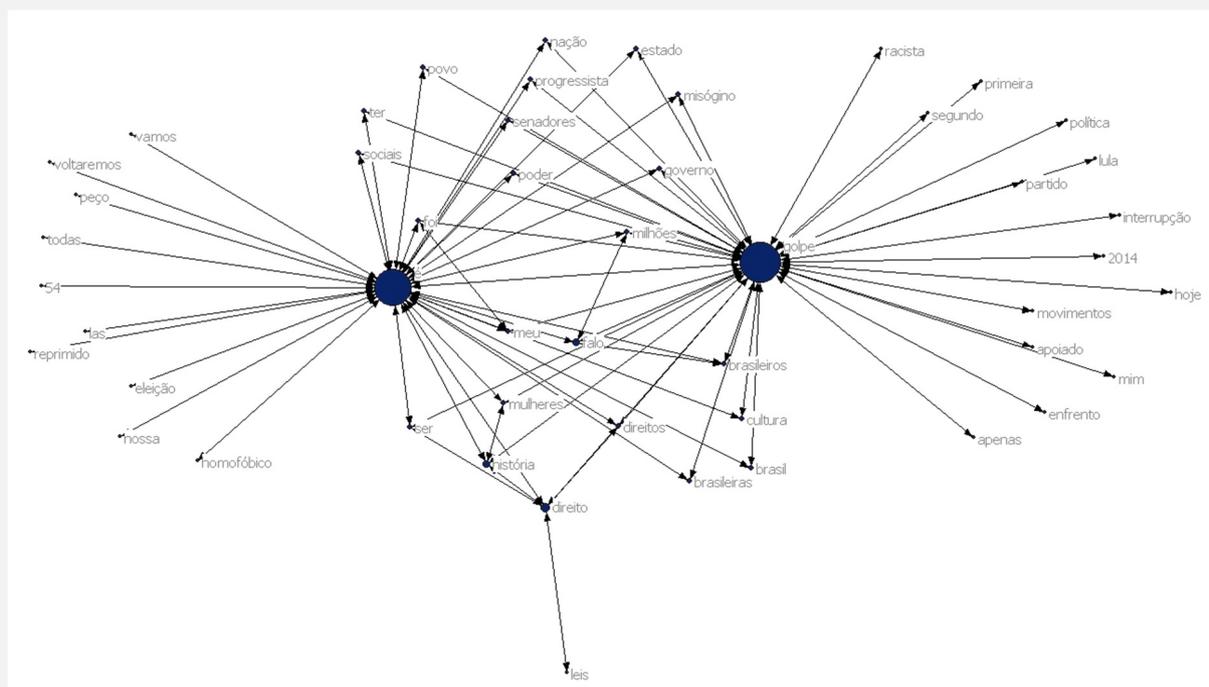
Fonte: Dados do estudo gerados com auxílio da ferramenta Ucinet

Em relação à segunda série, que engloba os discursos de posse, há uma remissão ao governo anterior, principalmente porque Dilma representava o terceiro mandato de Lula, em 2010, uma fala permeada por agradecimentos a aliados e consolidação de alianças futuras, e uma projeção de um país desejado que seria viabilizado pelo novo mandato. Os discursos desta série se assemelham ao da série anterior, uma vez que Dilma fez do discurso de posse uma lista de obras e projetos em andamento e que precisam melhorar ou continuar – estratégia pouco comum em discursos dessa natureza, dado o contrato de comunicação político eleitoral. Assim, destacam-se como termos centrais nas redes: *investimento*, *desenvolvimento*, *compromisso*, *crescimento*, *mudança*, formando subgrupos com os termos *economia*, *política*, *reforma* e *educação*, além dos verbos que projetam ações no futuro, tais como: *ampliar*, *avançar*, *continuar*, *fazer* e *vamos* (ver Figuras 3 e 4). Destaca-se ainda o agrupamento formado pelos termos *Brasil*, *brasileiros*, *povo*, *governo*, *país* e *vamos*, evidenciado pela análise centro-periferia, e que remetem a uma reafirmação do acordo estabelecido entre a candidata e seu eleitorado.

de denúncia e de acusação. Na fala em que reiterou ser vítima de um *golpe* - termo com maior medida de centralidade da rede, Dilma constrói sua estratégia argumentativa com base em dois pontos principais.

O primeiro refere-se à forma indevida de ascensão ao governo, utilizada por alguns políticos, por meio do que ela denomina de um golpe de Estado. Para tanto, faz remissão ao agrupamento dos termos: *governo, golpe, poder, eleição, estado e povo*.

O segundo aspecto diz respeito ao apelo às mulheres e à presença feminina na política, argumentando que a primeira presidente mulher do país foi destituída do cargo devido à misoginia e ao machismo do comando do país. Tal argumento é percebido pela utilização do grupo lexical formado pelo termos *mulheres, machista, misógino, golpe e política* (ver Figura 5).



Obs.: Rede com 50 termos apresentados segundo medida de centralidade de grau

FIGURA 5: Rede terminológica – Discurso de Impeachment

Fonte: Dados do estudo gerados com auxílio da ferramenta Ucinet

A análise das três séries possibilitou perceber a mobilização dos termos empregados por Dilma, ressaltando a situação de uso e as polarizações semânticas dos termos diacrônicos, utilizados no ajuste de sua estratégia argumentativa, dadas as alterações na cena enunciativa, de forma particular, e à dinamicidade do domínio político, de maneira geral.

O agrupamento dos termos permitiu a identificação de termos correlatos utilizados com o mesmo teor semântico, de termos recorrentes e de termos que tendem a ficar próximos uns dos outros, de acordo com o período em que foram proferidos. Assim como uma análise semântico-discursiva dos termos e a análise centro-periferia evidenciaram a ideia de que a centralidade de certos termos e as relações estabelecidas entre eles derivam do assunto em evidência, comportam diferentes nuances de sentido e servem como ponto de entrada para a representação dos tópicos da dinâmica do domínio político. Isso fica evidenciado nas redes mapeadas e apresentadas a seguir, por exemplo, que representam o vocabulário utilizado na fase de candidatura e no discurso de posse pela candidata, no primeiro pleito.

CONCLUSÕES

A metodologia proposta propicia a construção de uma rede destinada a expressar as relações entre os termos, considerando-se o contexto particular onde aparecem, e possibilita o estabelecimento de relações inusitadas entre eles. Evidencia-se que o aprofundamento de estudos de redes terminológicas possibilita descrever e explicar melhor as dinâmicas e multifacetadas relações sociais estabelecidas entre diferentes sujeitos, para o que abordagens, como a empregada nesta comunicação, que combinam técnicas oriundas do processamento (automático) da linguagem natural (PLN), como a representação estatística de documentos através do cálculo de frequências de palavras; e o uso de sistemas de organização do conhecimento (SOC/KOS) para a identificação de termos-chave no campo do discurso, são potencialmente benéficas.

Por fim, a formalização explicativa, empreendida pela abordagem aqui proposta, através da análise de alguns processos discursivo-argumentativos que são constitutivos da maneira como os sujeitos efetivam os discursos sociais, permite melhor compreender as configurações e remanejamentos das trocas discursivas em domínios dinâmicos, tal como o político, e, conseqüentemente, reduzir o gap entre a formulação conceitual e sua representação em diferentes dispositivos no âmbito da organização do conhecimento.

REFERÊNCIAS BIBLIOGRÁFICAS

- Bick, E. (2000). *The Parsing System "Palavras": Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework*. Aarhus University Press.
- Bonnafeous, S., & Tournier, M. (1995). Analyse du discours, lexicométrie, communication et politique. *Langages*, 29(117), 67-81.
- Borgatti, S. P., & Everett, M. G. (2000). Models of core/periphery structures. *Social networks*, 21(4), 375-395.
- Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). *Ucinet for Windows: Software for social network analysis*. Harvard, MA: Analytic Technologies.
- Charaudeau, P. (1996). Para uma nova análise do discurso. Carneiro, A. D. *O discurso da mídia*. Rio de Janeiro: Oficina do Autor, 5-43.
- Charaudeau, P. (2006). *Discurso político*. Contexto.
- Charaudeau, P., & Maingueneau, D. (2008). *Dicionário de análise do discurso*. Contexto.
- Emediato, W. (2003). Os gêneros discursivos como tipos situacionais. Mari, H et al. *Análise do discurso em perspectivas*. Belo Horizonte: Nad/Fale/UFMG, 63-72.
- Everett, M. G., & Borgatti, S. P. (2005). Extending centrality. Carrington, P. J., Scott, J., & Wasserman, S. (Eds.). (2005). *Models and methods in social network analysis*. Cambridge university press, 57-76.
- Hodge, G. (2000). *Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files*. Digital Library Federation, Council on Library and Information Resources.

Lazega, E. (2007). *Réseaux sociaux et structures relationnelles: «Que sais-je?»*. Presses universitaires de France.

Marciniak, M., & Mykowiecka, A. (2009). Aspects of Natural Language Processing. *Lecture Notes in Computer Science*, 5070.

Souza, R. R., Tudhope, D., & Almeida, M. B. (2012). Towards a taxonomy of KOS: Dimensions for classifying Knowledge Organization Systems. *Knowledge organization*, 39(3), 179-192.

Venâncio, L. S. (2014). *Encenações linguageiras, jogos argumentativos e redes terminológicas nas eleições presidenciais brasileiras de 2010: a representação da informação em domínios dinâmicos*. Tese de doutoramento, Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Minas Gerais, Brasil.

Wasserman, S., & Faust, K. (1999). *Social network analysis: methods and applications*. Cambridge: University Press.



A BIBLIOMETRIC STUDY OF ALGERIAN JOURNALS IN AGRONOMY. METHODOLOGICAL AND POLITICAL CONSULTING ASPECTS

Radia Bernaoui¹, Peter Ohly²

¹Ecole Nationale Supérieure Vétérinaire (ENSV), Algiers, Algeria, radiabernaoui@yahoo.fr

²ISKO, Bonn, Germany, 0000-0002-7353-6733, peter.ohly@gmx.de

ABSTRACT Our study deals with an evaluative analysis of the scientific articles in Algerian journals of agronomic sciences in order to identify the most productive authors, the most dominant themes and finally the cooperation between institutions and research laboratories. It is a bibliometric analysis of the *Annals of the INA* (2000-2008) and the *Journal Research Agronomique* of INRAA (2000-2009). The results of our study reveal a relation between the publications in the two national agronomic journals and the programs of the National Agricultural Development Plan (NADP), as well as the National Research Programs (NRPs). In contrast, a deliberate intervention must introduce the two journals into a more academic and collaborative context at the international level.

KEYWORDS *Scientific Journal; Agronomic Research; Research Evaluation; Co-authorship; Algeria.*

COPYRIGHT This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

Scientific publications are considered as means of disseminating knowledge and research results. Academic journals permit such sharing of knowledge between universities, research institutions, and research laboratories. We will evaluate this scientific production in order to judge its effectiveness and efficiency.

But, the evaluation of scientific production should not only be based on the judgment of the scientific and technical quality on the research work itself. Moreover there might be some shortcomings by peer-to-peer reviews, e.g. scientific rivalry (van Raan, 2003). "What is the use of the best quality works on themes finally, if not related to the aims pursued by the program" (Larédo & Vinck, 1991). Further, the involvement in networks enables to establish additional collaborative links (networking), e.g. publishing with other authors and other research laboratories.

Since Algeria has always been concerned with the food security of the country, it is evident that the agronomic research institutions have to be bound to the research programs and the orientations of the economic sector to increase productivity and ensure food security, and to transfer knowledge through research findings for sustainable development of the country.

The evaluation of scientific production in the Algerian agricultural sector via bibliometrics can be analyzed by scientific publications in national journals of agronomy. In this study, we focus to two main journals in this field: the *Annals of National Institute of Agronomy* (INA; now: ENSA) and the *Review of Agronomic Research* of the National Institute of Agronomic Research of Algeria (INRAA). At the particular time of research these two journals were considered as the most important and oldest periodicals in the field of agronomy in Algeria. This evaluation study is seen as an approach for helping decision-making and management of scientific research.

2. METHODOLOGY

There are many descriptive studies that identify the scientific production in the field of scientific networking. Some trace this research field back to the psychologist James McKeen Cattell (1903), who used the counts of scientific publications and citations to make quantitative and qualitative statements about the scientists (as cited after Godin, 2006, as cited after Jovanović, 2012). And in the 30s of the last century the sociologist Robert K. Merton (1973) took bibliographical categorizations of journal articles of the 17th century as data for his research on the development and change of research interests. Tunger (2017, p. 390) states that publications on bibliometrics are more growing than the scientific production as such, not at least as a reaction on the demand for more output oriented research. According to Rostaing (1993), the "characteristic of bibliometrics is to enable studies on publications, on quantitative data and not simply subjective (peer opinion). These quantitative data are calculated from statistical counts of publications or of elements extracted from these publications". Indeed, Okubo (1996) reports that "the analyses on cooperation permit to depict the scientific networks, and highlight the established relations between countries, institutions, or researchers, the analysis of a big program, the structure of scientific disciplines and their mutual relations".

Salaün and Van Cuyck (1999) explain these network relationships by saying "the connection of different individuals in a process of more global communication implies the possibility of developing personal knowledge (in the individual sense) and collective knowledge (that emerges from the group and is brought to the group as a whole [...]). In addition, the way to disseminate information, the nature of the links that are formed between individuals, the density of the relationships that are formed [...] modify the different characteristics of the networks and imply networks shaped according to the social connections that structure the individual members of the network". In this sense we continue an older study for the years 1939-1999 (Slimani Rahima 2005) and take the collaborative partnerships in the field of agricultural research in Algeria for our evaluation study of scientific production from the *Annals of the INA* (2000- 2008 with 72 articles, later stopped) and the *Agronomic Research Review* of INRAA (2000-2009 with 167 articles, 2008 was an unusual year as it contained the proceedings of a seminar on Biotechnology, additionally).

Also, we evaluate whether these themes dealt with correspond to the National Agricultural Development Plan (NADP) and the National Research Programs (NRP). These programs were adopted by the NADP and the actions to be implemented within a five-year perspective 1996-2000 (PNDA 1997). For the NRP, the program focused on the five-year projection on scientific research and technological development 1998-2002 (PNR 2008). Accordingly the publication years were selected.

The earlier scientometric study of Slimani Rahima (2005) had been carried out only on the "*Annals of the INA* of El-Harrach". This analysis was on several levels: the quantification of publications, the

chronological evolution (date of publication), the location of the authors (institutional affiliations), the origin of the authors and finally the indexed themes). Now our main analysis objectives are:

- 1) The enumeration of the scientific output of researchers.
- 2) Identification and enumeration of major themes.
- 3) Co-publications as indicators for scientific collaborations.
- 4) A comparative analysis between the two journals.

3. RESULTS AND DISCUSSIONS

3.1. EVALUATION OF AGRICULTURAL SCIENTIFIC PRODUCTION THROUGH THE *ANNALS OF THE NATIONAL AGRONOMIC INSTITUTE OF EL-HARRACH (2000-2008)*

3.1.1. Distribution of the articles of the *Annals of the INA* by year of production

This distribution by year (Figure 1) allows us to visualize the evolution over time. In the period 2000-2001, we see that the magazine has a mean publication rate, comparable to the period (2006-2008), whereas between 2001-2002 and 2004-2005 there is a small decrease. 2004 is mostly represented (10 articles).

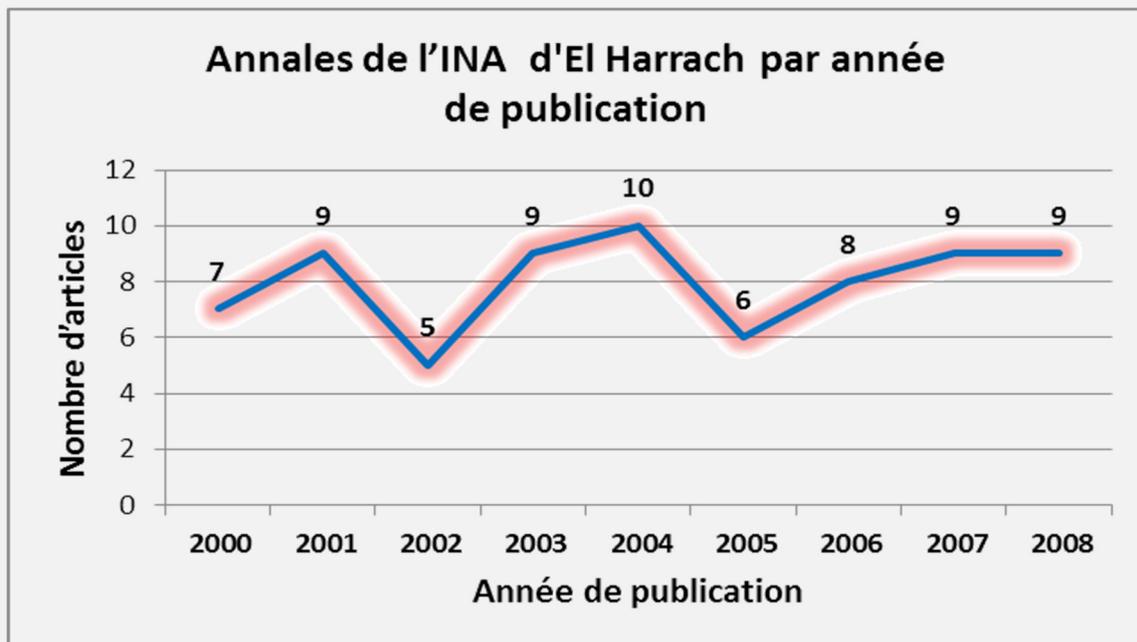


Figure 1: Distribution of the articles of the *Annals of the INA* by year of publication

3.1.2. Scientific production of the *Annals of the INA* by geographical area and by research laboratories

We find that 93.75% of the articles in the *Annals of INA* originate from Algerian laboratories, of which 59.43% result from laboratories of the city of Algiers. 34.32% of the articles are from 16 laboratories in other national cities, but each with a very low rate (Figure 2). Only 5.85% come from foreign laboratories (Tunisia, France and Switzerland). This reveals that during this period of 2000-2008, ENSA (formerly: INA) was little integrated into international scientific cooperation.

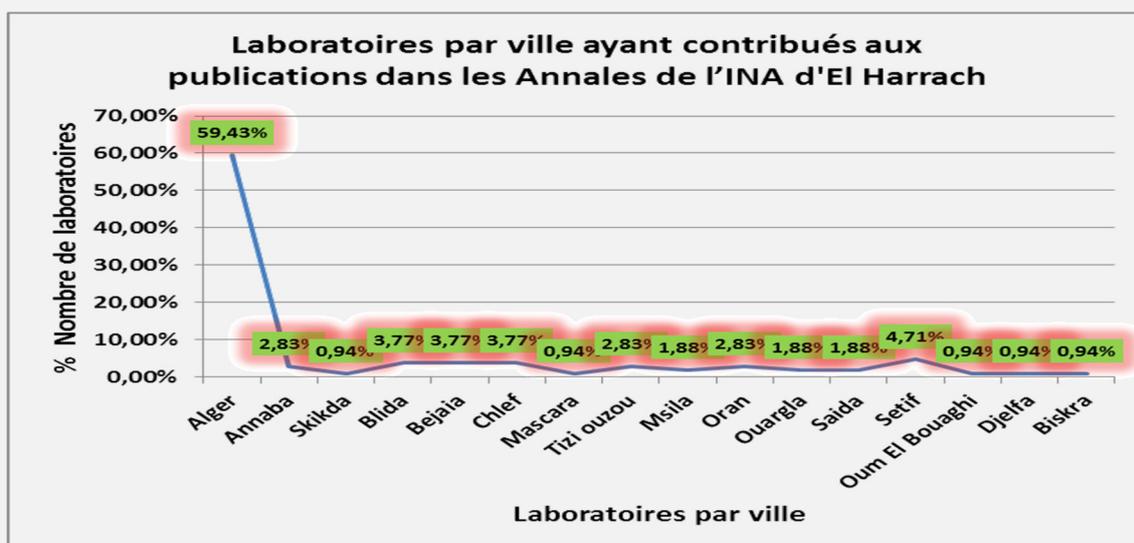


Figure 2: Distribution of laboratories (resp. cities) having contributed to publications in the *Annals of the INA*

3.1.3. Contribution of research laboratories to scientific production

41 laboratories in the nine years period contribute to publications in the *Annales de l'INA*. The Laboratory of Agricultural Machinism is heading in 2007 with four publications and in 2008 with three articles. The Laboratory of Biology and Physiology of Organisms holds the second place, with three publications for each year 2005 and 2006. The Laboratory of Genetic Resources and Biotechnology publishes once in each of three successive years (2002-2003-2004). The other laboratories are only able to launch one publication in this period 2000-2008).

3.1.4. Scientific production by department

The departments of "Animal Breeding" are the most productive ones, with a total of 27 published articles. Next are the departments of "Phytotechnics" and the departments of "Agricultural Engineering", with a respective total of 16 and 14 articles. This is followed by a total of 9 articles for the departments of "Food Technology". The rest of the departments publish only one article each.

3.1.5. Scientific production by subject area

In the *Annales de l'INA* are mostly represented Livestock Production and Crop Production, besides Biology in general. These two subject areas hold 15% and 14%. They are followed by specialized subject areas, which each represent 8 to 10% of published articles. A relatively low profile is represented by the themes Genetic Resources and Biotechnology as well as Biology and Physiology of Organisms with a rate of 7% and 6%. As for other subject areas: they do not exceed 1% of the publications. Examining whether these themes correspond to the National Agricultural Development Plan (PNDA 1997) and the National Research Programs (PNR 2008) shows that the themes addressed by the researchers correspond with the requirements of these two major national programs.

3.1.6. Occurrence of the most productive authors of the *Annals of the INA*

The two most prolific teacher-researchers of ENSA reach a frequency of six papers written in the nine years, for the period (2000-2008). Five other researchers contribute with five articles each. On the other hand, 107 authors out of a total of 138 are represented with only one publication. Based on our

exploratory interviews with the most prolific authors, it is found that they have published mostly in the areas of Animal Production, of Machinery, and of Food Technology.

3.1.7. Publications co-authored with ENSA

In our evaluation of authors' production in the *Annales de l'INA*, we are also looking at collaborative networks. In a first step, the co-author analysis shows that of the 72 articles published in the *INA Annals*, nearly 36 are co-authored. In this group, ENSA, the journal holder and founder, ranks first, with a contribution of 27 co-authored publications. Here the University of Sciences and Technology Houari Boumediene (USTHB) with its departments Biology and Vegetal Ecology shares seven with ENSA.

At the international level, collaboration of ENSA is restricted only to France, which comes at second position with five co-published articles. These contributions concern particularly the National Institute of Agricultural Research (INRA) and the Center for International Cooperation in Agricultural Research for Development (CIRAD) in Montpellier. But curiously, INRAA only ranks third with three collaborative contributions. But the two research institutions ENSA and INRAA are considered as the hub and the central base of agricultural research in Algeria. Shouldn't there be more synergy in building scientific networks?

3.1.8. Synthesis of the quantitative analysis of the *Annals of the INA* of El-Harrach

Publication of articles in the *Annals of the INA* in El Harrach is more or less constant for the period under investigation (2000-2008), with the exception of the year 2002. The publications differentiated by geographical area and by research laboratories reveal the tenacity of research laboratories acting at the national level.

Although scientific production by departments focuses on "Zootechnics", "Crop Farming" and "Rural Engineering", it is primarily specialized in Animal Production and Plant Production. Secondly, it devotes itself to other very diverse disciplines, such as: Food Technology, Mechanizing, Ecology, Hydraulics. In contrast, the themes of Ecology of Marine Environments, Soil Sciences, Economics and Information are mostly marginalized by scientific writings. However, the various themes dealt with are in fact in line with the country's economic development programs.

As for the frequency-based analysis of the authors, there is a limited number of prolific ones together with a general low prolific rate. Finally, the networks of collaborations are expressed by a strong national contribution between the ENSA and the USTHB. On the international level it is only France which is represented with two prominent centers.

3.2. EVALUATION OF AGRONOMIC SCIENTIFIC PRODUCTION BY THE AGRONOMIC RESEARCH REVIEW OF THE NATIONAL INSTITUTE OF AGRONOMIC RESEARCH OF ALGERIA (2000-2009)

3.2.1. Distribution of the articles of the INRAA *Agricultural Research Review* by year of publication

The number of publications in 2000 is significantly lower than in the later years (Figure 3). The outcome more than doubles in 2001 but decreases then somewhat till 2004 and to mount in 2005 again and to rise on a higher level between 2005 and 2008, with an average of about 20 publications. When we add the special issue with proceedings of a seminar on "Biotechnology at the service of the agri-food sector"

in 2008 (in the graph represented as an extra year) it shows a remarkable increase on further 33. As a counter balance, in 2009 it is followed by a decline on only ten articles.

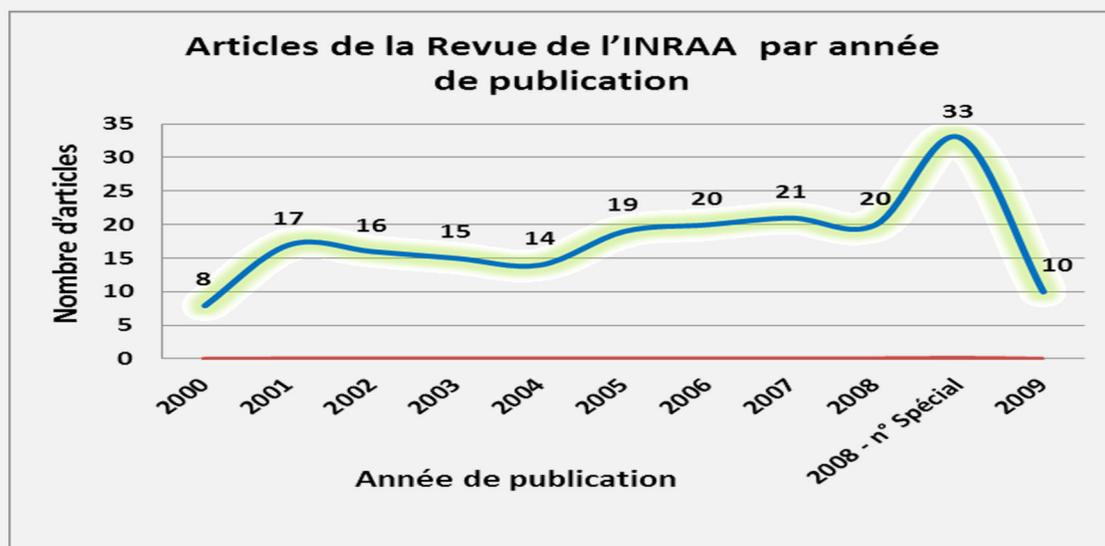


Figure 3: Distribution of articles in the INRAA *Agronomic Research Journal* by year of publication

3.2.2 Scientific production of INRAA's *Agronomic Research Review* by geographical area and by research laboratories

Figure 4 shows that the contribution of laboratories is centralized on the city of Algiers with a total of 67 laboratories (40%). Among the 57% of laboratories belong to other cities in Algeria, University of Blida and Sétif are mostly present. The participation of Algerian researchers is very dominant, with 163 articles published in the period 2000-2009 (Figure 6 b). Only a small number of scientific production comes from the neighboring Maghreb countries (Tunisia and Morocco) with a respective total of 2 articles (1%).

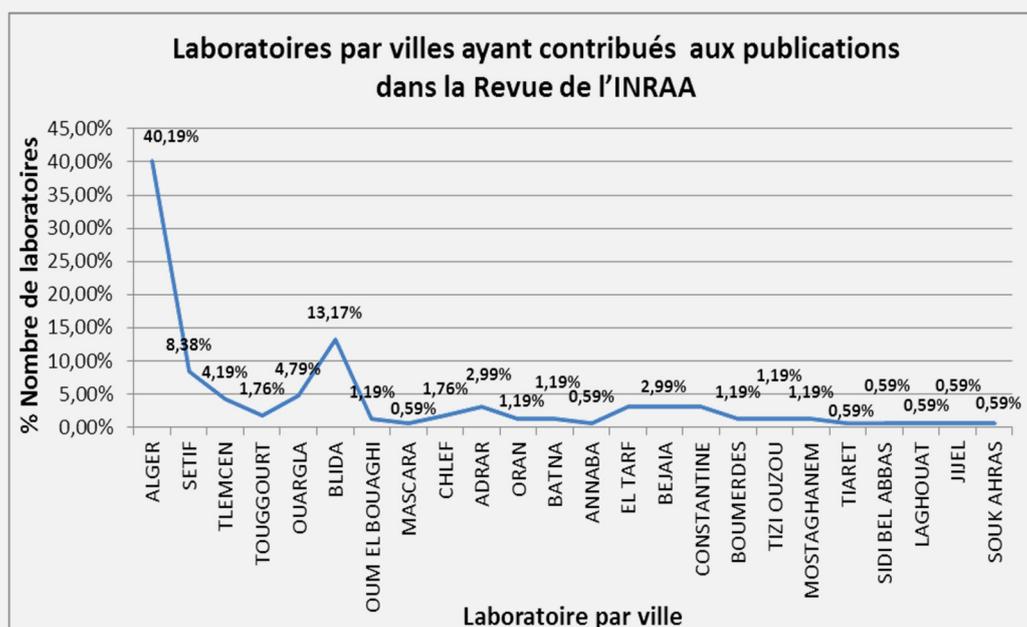


Figure 4: Distribution of laboratories (resp. cities) having contributed to publications in the INRAA *Agronomic Research Review*

3.2.3. Scientific production by specialty

Two disciplines of equal importance with respect to the socio-economic aspect are lying in front. The Crop Production produces the biggest number of articles with a total of 38 (26.5%) and is followed by the Animal Production which 28 items (20%). The theme of Economics (and Rural Sociology) as well as the Physical Environment occupy an average of 10% in terms of productivity. On the other hand, the disciplines of Plant Protection, Zoology, Zootechnics, Bioclimatology, Forestry, Animal Biology are little represented in INRAA's *Agronomic Research Review*. Like the themes in the *Annals of the INA* of El Harrach the themes in the INRAA *Agricultural Research Review* are carried out within the framework of the National Plan of Agricultural Development (PNDA) and the National Programs of Research (PNR).

3.2.4. Frequency of occurrence of the most productive authors of the INRAA *Agronomic Research Review*

Two researchers rank first with a total of seven articles published in the period 2000-2009. They are followed by two authors with a contribution of six articles. The rest of the 319 researchers were only able to publish one article each in the period 2000-2009. The most productive authors in the INRAA *Agronomic Research Journal* are grouped on the same themes: Livestock Production, Crop Production and Food Technology.

3.2.5. Publications co-authored with INRAA

92 articles of 167 articles published in this journal are co-authored. INRAA, the main publisher of its journal, occupies the first place with 35 articles in co-publication, dominated by a collaboration with the Ecole Nationale Supérieure d'Agronomie of El-Harrach (ENSA), with 27 articles. Obviously INRAA needs to cooperate with ENSA in order to carry out research in shared themes.

On the other hand, the University of Science and Technology Houari Boumediene (USTHB) has very little presence in this collaborative networking process. According to our exploratory interviews with INRAA researchers, they indicate that the USTHB is an institution based on education and training in ecology and environment as a priority. Its mission focuses more on basic research, which consists of experimental or theoretical work undertaken mainly to acquire new knowledge of the foundations of phenomena and observable facts, without really considering an application or a particular use. Contrary, INRAA's scientific research is more aimed at an in-depth framework of applied research. Therefore, it is mainly directed towards experimental development, which consists of systematic work based on existing knowledge obtained through research and/or practical experience, with a tendency to launch the manufacturing of new products, to establish new devices, or to improve existing ones significantly.

From a geographical point of view, the international cooperation network seems to be accentuated with France, which appears with seven publications. More particularly, INRAA mostly collaborated with the National Schools of Agronomy, Dijon, Paris and Montpellier. The neighboring countries Morocco and Tunisia are only represented with a total of three articles in co-publication. For Morocco it concerns the Agronomic and Veterinary Institute Hassan II of Rabat and the National Institute of Agronomic Research (INRA) of Kenitra, for Tunisia the National Institute of Agronomy of Tunis.

3.2.6. Synthesis of quantitative analysis of the INRAA Review

We note a certain stability in productivity. Only the years 2000 and 2009 are characterized by a rather low publication rate. The breakdown by geographical area and by research laboratories highlights the predominant focus on the national level in relation to the international level, where only few laboratories from neighboring countries are represented (Morocco and Tunisia).

The scientific specialization is mostly on Animal Production and Plant Production. Other themes are poorly represented, such as Plant Protection, Zoology, Zootechnics, Bioclimatology. But the main themes are in line with the country's economic priorities.

In this scientific production, some most eminent occurring researchers are accompanied by a majority of authors who appear only once. Regarding the share of co-publication between national organizations, we observe ENSA in the writings of the INRAA, but less the USTHB. Articles are co-authored with foreign bodies from France and the Maghreb countries only.

4. COMPARATIVE STUDY OF THE TWO JOURNALS

The production of the *Annals of the INA* in El Harrach is more or less constant for the period 2000-2008, with the exception of the year 2002 and 2005. During the period 2000-2009 the INRAA magazine shows as well a certain balance in productivity, except for the first and the last two years. Here the editors policy might come into effect.

The breakdown by geographical area and by research laboratories shows the predominance of research laboratories at the national level. Though there is an international cooperation with France in the *Annals of the INA* of El Harrach, it remains absolutely very low. Also, only a mediocre interest is focused on the Maghrebian collaboration for the two magazines.

Most articles published in the two reviews are written on Crop Production and Animal Production. Thus not only a coherence exists between the two journals, but also a sufficient adequacy is revealed with the priorities of the economic development programs. The same holds for the lack of certain special fields. Moreover the most productive authors are characterized by shared interests, such as Animal Production, Crop Production and Food Technology.

According to bibliometric laws, very few researchers are characterized by frequent productivity in the two journals studied. The analysis of the cooperations shows a weak tendency towards a networked scientific activity by connecting with national or international bodies. But curiously, the cooperation between ENSA and INRAA is more apparent in the articles of the INRAA Review than in the *Annales de l'INA* of El Harrach. The contribution of the USTHB is also more present in the *Annals of the INA* than in the INRAA *Agronomic Research Review*. Each journal seems to refocus on itself, either because it does not succeed in encouraging its collaborators, or because its diffusion does not allow a regular contribution of occasional writers.

5. CONCLUSION AND RECOMMENDATIONS

This comparative analysis of the articles of the two journals focuses only on a few well-defined themes. But even if the represented research fronts convene with the main expectations of the different primary economic development programs, they remain in a framework of unfavorable institutional compartmentalization between national and international research institutions. In effect the researchers work is only acting within a restricting scheme instead within a one of co-production.

A major problem arises from the impossibility of Algerian researchers to integrate into international exchange networks. Okubo (1996b) states that "scientific creation is still a mostly national act. The co-authorships show that the researchers first count on national knowledge within or between the same laboratories".

We believe that Algerian magazines must react on two major points so that they can fit into the ranks of prestigious and academic journals:

- Internationalization of science through co-publications

Algerian editorial journal boards must imperatively encourage co-publications with foreign bodies and laboratories. Esterle & Filliatreau (2014) confirm the importance of collaborative work when they speak about science and technology indicators: "International co-publications are one of the indicators of the internationalization of science and provide a good measure of the scope and quality of a country's research activity at the international level".

- Reputated journals

According to Durand-Barthez (2005) "the current profusion of scientific publications, which is constantly growing, necessitates qualitative filtering that is increasingly reflected in the quotations. Requiring a researcher to be quoted in high-impact journals is a common institutional behavior". Hence, Algerian journals must reconsider the criteria of scientific production in order to integrate their journals in a much more internationalized science sphere without frontiers.

REFERENCES

Durand-Barthez, Manuel (2005). L'évaluation des publications scientifiques: Du facteur d'impact à l'indice de notoriété. Retrieved 10/03/2014 from <http://archivesic.ccsd.cnrs.fr/docs/00/08/38/70/HTML/index.html>

Esterle, Laurence, & Filliatreau, Ghislaine (2014). *Indicateurs de sciences et de technologies. Rapport de l'Observatoire des sciences et des technologies. Edition 2014*. Paris: Economica.

Godin, Benoît (2006). On the origins of bibliometrics. *Scientometrics*, 68, 109–133.

Jovanović, Miloš (2012). Eine kleine Frühgeschichte der Bibliometrie. *Information - Wissenschaft & Praxis*, 63 (2), 71–80.

Larédo, Philippe, & Vinck, Dominique (1991). *Evaluer: Médiation et préparation des actions futures*. Bruxelles: De Boeck Professional Publishing.

McKeen Cattell, James (1903). Statistics of American Psychologists. *American Journal of Psychology* 14, 310–328.

Merton, Robert K. (1973). Changing Foci of Interest in the Sciences and Technology. In Merton, R. K., *The Sociology of Science. Theoretical and Empirical Investigations.*. Chicago: University of Chicago Press, 173-203

Okubo, Yoshiko (1996a). Les chercheurs peuvent-ils s'intéresser à une mesure de la science? Qu'est-ce que la bibliométrie? In *L'information scientifique et technique. Nouveaux enjeux documentaires et éditoriaux. Colloques INRA, 21-23 octobre 1996*, Tours.

Okubo, Yoshiko (1996b). *L'internationalisation de la science. Une analyse bibliométrique*. Paris: INIST-CNRS.

PNDA. (1997). Ministère de l'agriculture et de la pêche. *Données macro-économiques sur l'agriculture algérienne : perspectives 1996-2000*. Mars, 1997, 47 p.

PNR. (2008). Loi no. 98-11 du 29 Rabie Ethani 1419 correspondant au 22 août 1998 portant loi d'orientation et de programme à projection quinquennale sur la recherche scientifique et le développement technologique 1998-2002. In *Journal officiel*, no. 10, 2008

Rostaing, Hervé (1993). *Veille technologique et bibliométrie: concepts, outils, applications. Thèse de doctorat en Sciences de l'Information et de la Communication*. Marseille: Université de Droit et des Sciences d'Aix-Marseille.

Salaün, Jean-Michel, & Van Cuyck, Alain (1991). *Les usages et les besoins des documents numériques dans l'enseignement supérieur et la recherche. Rapport final. Programme Numérisation pour l'Enseignement supérieur et la Recherche. Octobre*. Villeurbanne: ENSIB.

Slimani Rahima, Khaled. L'utilisation des indicateurs d'activités pour l'évaluation de la revue « les annales de l'Institut National Agronomique d'El-Harrach » de 1939 à 1999. In *Séminaire national le RADA à l'ère du numérique*. INA El-Harrach, 7-8 novembre 2005

Tunger, Dirk (2017). Bibliometrie: Quo vadis? In Babik, W. et al. (Eds.), *Theory, Semantics and Organization of Knowledge. Fortschritte der Wissenorganisation 13*, Würzburg: Ergon, 388-398

van Raan, Anthony F. J. (2003). The use of bibliometric analysis in research performance assessment and monitoring of interdisciplinary scientific developments. *Technikfolgenabschätzung* 12, 20-29.



CARTAS COMO GÊNERO DO DISCURSO CIENTÍFICO NA ÁREA DE MEDICINA TROPICAL: ANÁLISE BIBLIOMÉTRICA E MAPEAMENTO DE REDE TERMINOLÓGICA

Maria José Veloso da Costa Santos¹, Vania Lisboa da Silveira Guedes²

¹Universidade Federal do Rio de Janeiro, 0000-0003-0473-5680, msantos1402@gmail.com

²Universidade Federal do Rio de Janeiro, 0000-0001-5854-5677, vanialisboa@facc.ufrj.br

RESUMO A presente pesquisa desenvolve análise quantitativa e qualitativa em documentos de arquivo do cientista Adolpho Lutz (1855-1940), particularmente, utilizando cartas de sua correspondência científica passiva, sobre febre amarela e malária, que serviu como canal de comunicação para a troca de conhecimentos com seus pares, na área de Medicina Tropical, no final do século XIX e primeiras quatro décadas do século XX. O estudo se inscreve na fronteira entre a História da Ciência e a Ciência da Informação, porque analisa a trajetória da Medicina Tropical no Brasil à luz de pressupostos teóricos da Ciência da Informação. Objetiva investigar a viabilidade de aplicação da fórmula do Ponto de Transição de Goffman enunciada como contribuição à lei de Zipf em cartas com conteúdo científico, para reconhecimento da terminologia que compõe o campo semântico da área, visando à modelagem da rede terminológica da Medicina Tropical à época de Lutz. Do ponto de vista teórico-metodológico, para a composição do *corpus*, foram selecionadas 99 cartas do acervo da Biblioteca Virtual em Saúde Adolpho Lutz. O *corpus* foi processado à luz da Lei de Zipf e do Ponto de Transição de Goffman e os resultados confirmam a premissa de que a carta é um tipo de gênero do discurso científico com alto grau de informatividade em relação aos temas estudados, de tal modo que, termos com alto conteúdo semântico situam-se na Região de Concentração de Goffman, resultados semelhante aos atingidos por pesquisas realizadas tendo como *corpus* artigos de periódicos científicos.

PALAVRAS-CHAVE Cartas científicas. Lutz, Adolpho. Lei de Zipf. Ponto de Transição de Goffman. Bibliometria. Medicina Tropical.

ABSTRACT The present research develops quantitative and qualitative analysis in archival documents of the scientist Adolpho Lutz (1855-1940) using letters from his passive scientific correspondence on yellow fever and malaria that served as a communication channel for the exchange of knowledge with his peers in the area of Tropical Medicine in the late nineteenth and first four decades of the twentieth century. The study is part of the frontier between the History of Science and Information Science, because it analyzes the trajectory of Tropical Medicine in Brazil in the light of theoretical assumptions of Information Science. It aims to investigate the feasibility of applying Goffman's Transition Point formula as a contribution to Zipf's law in letters with scientific content, to recognize the terminology that composes the semantic field of the area, aiming at the modeling of the terminology network of Tropical Medicine at the time of Lutz. From the theoretical-methodological point of view, for the composition of the corpus, 99 letters were selected from the collection of the Virtual Health Library Adolpho Lutz. The corpus was processed in light of Zipf's Law and Goffman's Transition Point, and the results confirm the premise that the letter is a kind of genre of scientific discourse with a high degree of informativeness in relation to the subjects studied, so that, Terms with high semantic content are located in the Goffman Concentration Region, results similar to those reached by research carried out using corpus articles from scientific journals.

KEY WORDS *Scientific letters. Lutz, Adolpho. Zipf's Law. Goffman Transition Point. Bibliometrics. Tropical Medicine.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A documentação produzida por cientistas e por instituições retrata o cotidiano no contexto político, econômico e social aos quais estão inseridos, bem como revela as relações entre seus pares e instituições congêneres, em níveis nacional e internacional. São fontes documentais que estão intimamente relacionadas à formação do campo de investigação da História da Ciência e se constituem matéria-prima importante para os pesquisadores que atuam nesse campo. Welfelé (2004, p. 64) ao se referir às questões enfrentadas em relação a documentos de arquivo produzidos no *Centre National de la Recherche Scientifique* (CNRS), na França, observa que “a produção de documentos da ciência [...] estão nos laboratórios, lá onde a ciência se elabora, se manifesta e é trabalhada [...]”, isto porque proporcionam o entendimento das práticas científicas em uma época. Para Braga (1974, p. 160) o documento “não é uma unidade estática de comunicação, a ser armazenado e recuperado; é uma parte mutável da estrutura social da Ciência, produzido em um conjunto de circunstâncias, utilizado em outro”.

Adolpho Lutz nasceu na cidade do Rio de Janeiro em 1855 e faleceu no ano de 1940 nessa mesma cidade. É considerado por muitos como “o mais completo, versátil cientista que o Instituto Oswaldo Cruz já possuiu”. Atuou nas áreas de Clínica Médica, Helminologia, Bacteriologia, Terapêutica, Veterinária, Dermatologia, Protozoologia, Malacologia, Micologia, Entomologia (BENCHIMOL; SÁ, 2004), integrando esses campos de pesquisa à Medicina Tropical.

A Medicina Tropical é o ramo da Medicina que desenvolve pesquisas básicas e aplicadas na área de doenças infecciosas e parasitárias, geralmente associadas à descrição daquelas transmitidas por vetores intermediários, tais como os insetos, os moluscos ou os vermes. Esse ramo da Medicina engloba pesquisas nas áreas de Microbiologia, de Parasitologia e de Entomologia, aliando-as às pesquisas de campo concentradas nas interações dos organismos, em seus ambientes naturais, mais frequentes nas regiões tropicais e subtropicais. (SOCIEDADE BRASILEIRA DE MEDICINA TROPICAL, 2015; LOWY, 2006). De acordo com Lowy (2006, p. 37, 38) a Medicina Tropical “nasceu [...] da adaptação das “ciências pasteurianas” às doenças dos países do Sul”, ou seja, do Hemisfério Sul. Portanto, está estreitamente ligada à expansão colonial, daí o interesse de seu fundador em concentrar seus estudos nas doenças das colônias inglesas de além-mar, localizadas em regiões tropicais e subtropicais.

O presente trabalho desenvolve uma análise bibliométrica na correspondência científica passiva do médico brasileiro Adolpho Lutz (1855-1940), que serviu como canal de comunicação com seus pares, para a troca de informações na área de Medicina Tropical, sobre duas doenças tropicais: a febre amarela e a malária, no período do final do século XIX, até a década de 1940 do século XX. Nesse sentido, a pesquisa encontra-se na fronteira entre a História da Ciência e a Ciência da Informação, porque analisa a trajetória da Medicina Tropical no Brasil à luz de pressupostos teóricos da Ciência da Informação, como, a Comunicação Científica e estudos bibliométricos. A pesquisa visa a atestar a viabilidade de utilização da lei de Zipf e aplicação da fórmula do Ponto de Transição de Goffman, para determinar o

campo semântico da área em questão, bem como a modelagem de rede terminológica utilizada à época do cientista, que, segundo Kobashi e Santos (2006, p. 31) vêm a ser a “cartografia temática” dessa área.

FUNDAMENTAÇÃO TEÓRICA

As cartas, desde os tempos mais remotos, são canais informais de comunicação entre cientistas e eram utilizadas para garantir a primazia da pesquisa, daí serem também consideradas como os primeiros registros de texto científico. Bazerman (2006) acredita que o primeiro artigo científico emergiu da correspondência de Henry Oldenburg, secretário da Royal Society of London, com outros estudiosos. Oldenburg foi o primeiro editor do *Philosophical Transactions of the Royal Society of London* que, junto com o *Journal des Savants* publicado na França, são os primeiros periódicos científicos que surgiram, ambos publicados no ano de 1665. Os primeiros números do *Philosophical* foram editados sob a forma de resumos dessa correspondência e das reuniões da *Royal Society*.

Silva (1988, p. 75) define o gênero carta como “modalidades culturais de comunicação escrita” e acrescenta que a carta “apresenta semelhanças com uma conversa casual”. São unidades que integram um conjunto maior denominado correspondência, onde as impressões, de quem escreve, ficam registradas. Para existir uma troca de correspondência é necessário que tanto o destinatário como o autor da carta se reconheçam enquanto agentes ativos de uma determinada discussão. Para Lima (1997, p. xxv) a correspondência “é a câmara subjetiva, através da qual percebemos o ponto de vista do indivíduo, permitindo-nos, assim, estabelecer as conexões entre a sua trajetória pessoal e o momento histórico em que foi a um só tempo espectador e ator”.

O gênero carta circula em domínios discursivos variados e em diferentes esferas sociais (pública, pessoal, profissional etc), de acordo com o propósito comunicativo e para os mais variados fins. Sendo assim, as cartas analisadas, no contexto dessa pesquisa, são cartas científicas, porque se situam no domínio discursivo científico.

A teoria dos gêneros revela que as produções discursivas são importantes fatores de organização da comunicação do conhecimento em contextos específicos porque estruturam e sustentam significados institucionalizados na sociedade (ANDERSEN, 2008). As práticas discursivas, do ponto de vista da teoria de gêneros, definem o campo semântico de uma especialidade e de que maneira o conhecimento é agregado e codificado. Swales (1990) recorre à análise de gêneros para esclarecer a natureza do uso da linguagem em uma comunidade discursiva, que utiliza gêneros particulares de discurso, familiares a seus integrantes, formando o que o autor denomina de redes sociais e retóricas. Conceitua ainda, análise de gêneros como a análise de qualquer tipo de discurso, falado ou escrito, de diferentes comunidades discursivas. Nesse sentido, cada comunidade define quais gêneros são relevantes para atingir seus objetivos particulares.

O paradigma da disciplina análise de domínio no campo da Organização do Conhecimento estabelece que o melhor meio para o entendimento da informação é o estudo de um domínio do conhecimento como comunidade de discurso ou comunidade de pensamento (HJØRLAND, 2002). Nesse sentido, Hyland (2009) esclarece que, é no uso da linguagem, em contextos específicos, que as necessidades enunciativas da especialidade e o significado se estabelecem. É por meio de determinados gêneros discursivos que se tipificam a partir de aspectos históricos, sociais, temáticos e ideológicos que a

comunidade científica se comunica. (BAKHTIN, 1981). Portanto, o gênero carta que circula no domínio da ciência, em diferentes épocas, é um tipo de gênero científico.

A terminologia utilizada em uma área especializada tem a função, segundo Café e Brashner (2011), de dar nomes aos conceitos adotados na comunicação científica daquela área. É o que Dahlberg (1993, p. 211) denomina de “unidades do conhecimento e suas possíveis combinações em palavras/termos ou afirmações”. Para Benvenites (2006, p. 252), a história de uma especialidade se resume no conhecimento de seus termos/conceitos que vêm constituir uma terminologia própria. Assim, é a terminologia que garante a consolidação de uma especialidade na ciência. (KOBASHI; SMIT; TÁLAMO, 2001). Moreira González (2004) observa que no texto especializado os conceitos significantes formam uma rede integrada por palavras-chave (termos), mediante as quais, os autores expressam suas ideias principais, o que vem a contribuir para a inexistência de barreiras na comunicação científica especializada. (KRIEGER; FINATTO, 2004).

Visto pela ótica desses autores, pode-se compreender que o monitoramento da terminologia utilizada em um domínio especializado, em determinada área do conhecimento, permite descobrir de que modo as teorias e os conceitos representativos desse domínio eram configurados em uma determinada época.

Recentemente, existem publicações de resultados de pesquisa que apontam para a aplicação bem sucedida de Análise de Redes Sociais (ARS) para o mapeamento de redes terminológicas na Organização do Conhecimento. Segundo Fóris (2013), em vários estudos sobre frequência de palavras, as Leis de Zipf vêm sendo aplicada para a análise da terminologia, de várias áreas do conhecimento, e provam indiretamente a existência de redes de linguagem. Esses estudos aproximam abordagens teóricas e metodológicas da Terminologia com a Bibliometria, da Cientometria e de Análise de Redes Sociais.

Nesse estudo, propõe-se a aproximação de algumas abordagens de redes sociais, sobretudo no que tange a suas características terminológicas e comunicativas, com a análise bibliométrica a partir de Zipf e Ponto de Transição de Goffman, como mecanismos de identificação de níveis distintos do conteúdo informativo das cartas científicas em análise.

METODOLOGIA

A presente pesquisa caracteriza-se como quali-quantitativa que segundo Goldemberg (2000, p. 62), seu uso simultâneo reúne um “conjunto de diferentes pontos de vista e diferentes maneiras de coletar e analisar os dados que permite uma ideia mais ampla e inteligível da complexidade do problema”. Apresenta como proposta teórico-metodológica a lei bibliométrica de Zipf e a aplicação da fórmula do Ponto de Transição de Goffman à luz da metodologia utilizada por Miranda Pao (1978), com vistas a reconhecer temas mais frequentes tratados na construção do conhecimento na área da Medicina Tropical, particularmente sobre duas doenças importantes no Brasil: a febre amarela e a malária.

A atividade humana é comandada pelo princípio do menor esforço, fato que chamou atenção de Zipf (1902-1950), em 1949, para a observação de textos que o autor julgou suficientemente longo. Notou que existe um padrão de comportamento de distribuição de palavras nesse texto, ao transformá-lo em unidades lexicais e ordená-las de acordo com sua frequência de ocorrência decrescente. É a partir de Zipf, que se inaugura o tratamento estatístico de textos escritos. Zipf enunciou uma segunda lei revisada e modificada por Booth, em 1967, que se focou principalmente nas palavras de baixa frequência de

ocorrência em um texto. Segundo Guedes (2010), Goffman, ao analisar a Lei Zipf/Booth verificou dois comportamentos distintos nas extremidades da lista: grande número de palavras que se repetem apenas uma única vez e número reduzido de palavras que se repetem com alta frequência de ocorrência. Nessa linha de raciocínio, o autor observou que deveria haver uma região na listagem, entre essas duas extremidades, que segundo Pao (1978), Goffman chamou de Região de Transição, onde estariam as palavras de maior conteúdo semântico de um texto.

Das observações de Goffman, pode-se depreender que o ponto crucial da análise bibliométrica de palavras de um determinado texto seria então, a identificação da região entre as extremidades da listagem de frequência, a partir da determinação do Ponto de Transição (*Transition Point*), conhecido como Ponto T de Goffman, calculado a partir da modificação da fórmula Lei de Zipf/Booth. Segundo Guedes (2013, p.57), Pao (1978), testou e comprovou a hipótese de Goffmann sobre a Região de Transição, onde estariam “as palavras de maior conteúdo semântico de um determinado texto”.

Para a composição da amostra selecionou-se na Biblioteca Virtual Adolpho Lutz o segmento correspondência e os temas **febre amarela** e **malária**. Assim, o *corpus* da presente pesquisa é composto de **99** cartas, das quais **62** são sobre febre amarela e **37** sobre malária.

RESULTADOS

O *corpus* foi composto pelo texto de 99 cartas que foi transformado em unidades lexicais empregando-se o software *RankWords* que apresentou uma lista de **16.717** palavras. Desse total, observou-se que **3.795 (22,7%)** são palavras distintas. Essas palavras apresentaram a frequência de ocorrência entre **uma** vez a **748** vezes. Uma única palavra foi repetida **748 (4,5%)** vezes (a preposição **de**), enquanto que **1.506 (39,6%)** palavras distintas aparecem uma única vez no *corpus*, ocupando a frequência **1**. O índice médio de repetição de cada palavra no *corpus* foi de **4** vezes (4,4).

Importa observar que essa distribuição encontra-se de acordo com o princípio de Pareto “muitos com pouco, poucos com muito” (Lei dos 80/20) e, também, com o Efeito Matheus na ciência (MERTON, 1968), comum aos indicadores produzidos pelas análises bibliométricas. O *corpus* também foi considerado **suficientemente** longo, de acordo com as premissas estabelecidas por Zipf.

O cálculo do Ponto T foi realizado a partir da fórmula Goffman, apresentando os seguintes resultados: o Ponto T localiza-se na frequência **62,4** da listagem de frequências de ocorrência. A frequência 62 corresponde à palavra **Senhor**. Em seguida, localizou-se o *rank* referente a essa palavra situada no *rank* **32**.

Ao testar a hipótese de que a Região de Transição de Goffman produz as palavras de maior carga semântica de um determinado texto, Pao (1978), delimitou a região, projetando para o final da listagem de palavras, o mesmo número de palavras que ocorreu acima da frequência que corresponde ao Ponto T. Assim, a palavra ligada à frequência, que indica o Ponto T (62,4), encontra-se no *rank* **32**. Seguindo os passos de Pao (1968), projetou-se o número de *ranks* acima do Ponto T para a parte inferior da listagem, chegando-se ao *rank* **64**. Com isso, foi determinada a Região de Concentração de Goffman que situa-se entre o *rank* **1** e o *rank* **64**. Eliminando-se as palavras que exercem função sintática no texto como: preposições, artigos, conjunções, pronomes e verbos auxiliares, identificou-se, as palavras com alta carga semântica, conforme o quadro 1 a seguir.

Quadro 1. Palavras com alta carga semântica

Palavras	Frequência	Rank
Lutz	54	36
Trabalho	48	42
Carta	46	44
Espécies	44	48
Mosquitos	39	51
Colega	38	55
Tempo	35	56
Amigo	33	61
Febre	31	63

Fonte: Dados da pesquisa.

O quadro 1, apresenta as **10** palavras com alto conteúdo semântico no *corpus* analisado e a figura 1 refere-se à nuvem de palavras da Região de Concentração de Goffman no domínio da febre amarela e malária, na Medicina Tropical, no período analisado. A figura 1, a seguir refere-se à nuvem de palavras da Região de Concentração de Goffman.



Figura 1- Nuvem de Palavras da Região de Concentração de Goffman

Para avaliar a coesão da redeterminológica, bem como para mostrar o emprego de estruturas linguísticas e o foco temático dos pesquisadores da época de Lutz foi realizado o levantamento de um subgrupo dessa rede referente também, à Região de Concentração de Goffman, que por sua vez, gerou o grafo de duas cliques. A figura 2, a seguir, mostra as duas cliques juntas.

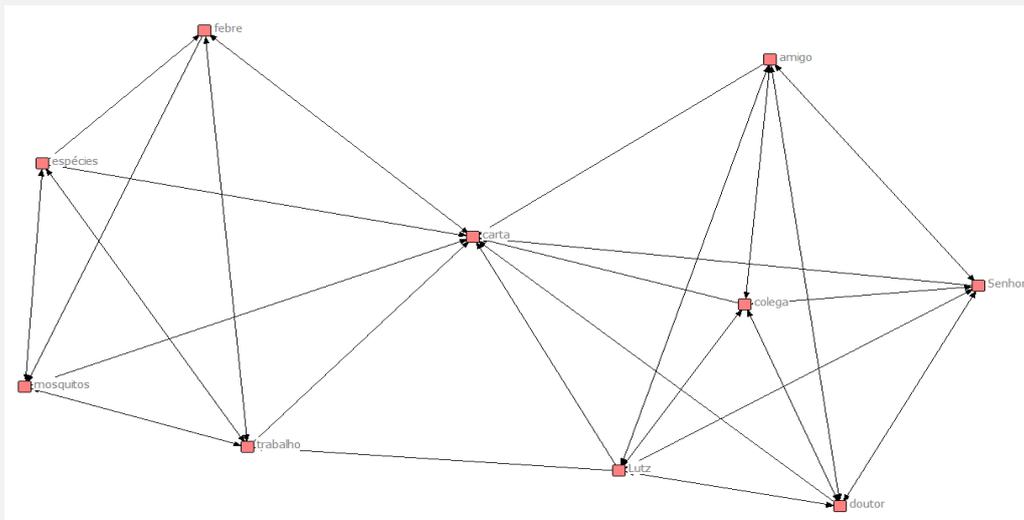


Figura 2 - Grafo das cliques da Rede Terminológica

A **clique da direita** contém estruturas reconhecidas na Linguística como estratégias de polidez na comunicação científica (senhor e doutor), outras estruturas revelam a relação informal entre Lutz e seus missivistas (Lutz, amigo e colega), ao mesmo tempo que denotam o grau de intimidade, amistosidade e interação entre os componentes de seu Colégio Invisível. A **clique da esquerda** revela o foco temático das pesquisas e dos conhecimentos que os pesquisadores compartilhavam (trabalho, espécies, mosquitos e febre), o que está de acordo com o que Braga (1974, p. 161) defende: “80% da comunicação entre os pesquisadores para a troca de experiências na geração de conhecimentos são provenientes do Colégio Invisível”. Percebe-se, ainda, nessa clique, que é notória a intensa atividade de pesquisa, visando a combater o vetor (mosquito) e, assim, evitar a contaminação de outros indivíduos com a febre amarela e a malária, conforme é mostrado no grafo acima.

Observa-se também, que a palavra **carta**, embora não seja o conceito central no discurso dos atores, com alto valor semântico, ela conecta as duas cliques estruturalmente. É a palavra carta que apresenta maior centralidade de intermediação, o que significa que a carta foi o instrumento mediador na construção do discurso científico dos atores mediante o intercâmbio de ideias e a colaboração.

CONCLUSÕES

O objetivo da presente tese foi analisar e descrever os canais informais de comunicação científica, no caso em estudo, as cartas trocadas entre Adolpho Lutz e outros cientistas, em nível nacional e internacional, com vistas a identificar cientistas, temas e subtemas ligados à produção de conhecimento sobre febre amarela e malária na Medicina Tropical, no final do século XIX e início do século XX, no Brasil. Em que pese as dificuldades encontradas, considera-se que o objetivo estabelecido para a pesquisa foi atingido.

Existe a tendência na área de Ciência da Informação, de que estudos quantitativos, em um domínio científico, devem ser acompanhados de estudos qualitativos. Dessa forma, percebe-se, no estudo em questão, que a Sociologia da Ciência e a História da Ciência aproximam-se da Ciência da Informação, aliando às análises quantitativas às análises qualitativas, assegurando assim o entendimento do campo científico e da comunidade científica em análise.

Portanto, pode-se reiterar a opinião de autores da área de CI, considerando que as pesquisas bibliométricas analisam continuamente a informação, em quaisquer áreas do conhecimento, tanto em nível micro de pesquisa, ou seja, em domínios especializados do conhecimento, como em nível macro, em estados, países e até internacionalmente.

Os resultados produzidos com as análises bibliométricas do *corpus* mostraram-se satisfatórios, uma vez que a aplicação da fórmula do Ponto de T e do mecanismo de delimitação da Região de Transição de Goffman apontaram palavras com alto teor semântico, em relação ao domínio científico analisado. Observou-se, ainda, nessa região, a recorrência de palavras de baixo conteúdo semântico como artigos, preposições, conjunções, verbos auxiliares etc que desempenham função sintática na comunicação.

No que se refere ao procedimento de delimitação da Região de Transição de Goffman, foi verificada ainda a aplicabilidade desse mecanismo para a análise qualitativa de conteúdo, empregando-se a técnica de análise categorial, que permitiu a contextualização, no *corpus*, de termos relevantes identificados na região.

Nessa perspectiva, conclui-se que esse estudo reitera o fato de que pesquisas que utilizam o Ponto T de Goffman, possibilitam a identificação de temas relevantes, em uma área do conhecimento, mostrando, nesse estudo, resultados coerentes com o estado da arte do domínio do conhecimento analisado, à época de Lutz.

Quanto aos resultados do mapeamento das redes terminológica, estes atenderam plenamente à investigação. O grafo da rede terminológica representou as palavras que compõem a terminologia das subáreas analisadas, com destaque para palavras com alto conteúdo semântico identificadas a partir do mecanismo de delimitação da Região de Transição de Goffman.

Cabe ressaltar, que o fluxo da rede é dinâmico, está sempre em transição e mostra a situação, em um dado momento, o que significa que a realização de futuras pesquisas, em outros períodos, possivelmente, dará origem a novos trabalhos, na área.

Deve-se, também, enfatizar que nesse estudo, os resultados revelaram nível satisfatório de coerência com temas importantes, do ponto de vista social, discutidos pela comunidade científica, à época, com a intenção de avançar no conhecimento sobre a febre amarela e a malária, visando a criar estratégias de prevenção e controle dessas doenças, fenômeno que é mostrado no subgrupo da rede terminológica.

Em relação à premissa de que a carta é um tipo de gênero do discurso científico, apesar da utilização de estruturas e linguagem diferenciadas, como por exemplo, estratégias de aproximação (você, amigo, colega etc) e de polidez (prezado, doutor, senhor etc), os resultados confirmaram essa premissa. Nesse sentido, a pesquisa revelou o alto grau de informatividade do texto das cartas, em relação aos temas estudados, de tal modo que termos com alto conteúdo semântico situam-se na Região de Concentração de Goffman, de forma semelhante aos resultados atingidos por pesquisas realizadas tendo como *corpus* artigos de periódicos científicos.

REFERÊNCIAS BIBLIOGRÁFICAS

Andersen, J. (2008) The concept of genre in information studies. *Annual Review of Information Science and Technology*, 42, 339-367.

Bakhtin, M.M. (2004). Os gêneros do discurso. In M. M. Bakhtin. *Estética da criação verbal*. São Paulo: Martins Fontes.

Bazerman, C. (2006) Cartas e a base social de gêneros diferenciados. In C. Bazerman. *Gêneros textuais, tipificação e interação*. 2. ed. São Paulo: Cortez.

Benchimol, J. L.; SÁ, M. R. (Eds.) (2004). *Adolpho Lutz: obra completa*. Rio de Janeiro: Editora FIOCRUZ.

Biblioteca Virtual em Saúde Adolpho Lutz. Recuperado em: <http://www.bvsalutz.coc.fiocruz.br> Acesso: mar. 2017.

Braga, G. (1974) Informação, Ciência, política científica: o pensamento de Derek de Solla Price. *Ciência da Informação*, 3 (2), 155-177.

Café, L.; Brascher, M. (2011) Organização do conhecimento: teorias semânticas como base para estudo e representação de conceitos. *Inf. Inf.*, Londrina, v. 16, n. 3, p. 25-51, jan./ jun.

Dahlberg, I. (1993) Knowledge organization: its scopes and possibilities. *Knowledge Organization*, 20, 211-222.

Fóris, A. (2013) Network theory and terminology. *Knowledge Organization*, v. 40, n. 6, p.422 – 429.

Goldemberg, M. (2000) *A arte de pesquisar: como fazer pesquisa qualitativa em Ciências Sociais*. 4. ed. Rio de Janeiro: Record.

Guedes, V.L.S (2010). *Nominalizações deverbais em artigos científicos: uma contribuição para a análise e a indexação temática da informação*. (Tese de Doutorado não publicada, Universidade Federal do Rio de Janeiro).

Hjørland, B. (2003) Fundamentals of knowledge organization. In J.A. Frías, C. Travieso (Eds.) *Trends in Knowledge Organization Research* (p.83-116). Salamanca: Ediciones Universidad de Salamanca.

Hyland, K. (2009) *Academic discourse: english in a global context*. New York: Continuum International Publishing Group.

Kobashi, N.Y.; Santos, R.N.M.S. (2006) Institucionalização da pesquisa científica no Brasil: cartografia temática. *TransInformação*, 18 (1), 27-36, jan./abr.

Krieger, M.G.; Finatto, M.J.B. (2004). *Introdução à terminologia: teoria e prática*. São Paulo: Contexto.

Lima, A.L.G.S. (1997). *Meu caro Oswaldo: a história da saúde vista através da correspondência de Oswaldo Cruz (1899-1907)*. (Dissertação de Doutorado não publicada, Instituto de Filosofia e Ciências Sociais, Universidade Federal do Rio de Janeiro).

Löwy, I. (2006). *Vírus, mosquitos e modernidade: a febre amarela no Brasil entre ciência e política*. Rio de Janeiro: Editora FIOCRUZ.

Moreiro González, J.A. (2004). *El contenido de los documentos textuales su análisis y representación mediante el lenguaje natural*. Gijón: Ediciones TREA.

Pao, M. L.(1978). Automatic text analysis based on transition phenomena of word. *Journal of American Society for Information Science*, 29 (3), 121-124.

Silva, V.L.P.(1988). *Cartas cariocas: a variação do sujeito na escrita informal*. (Tese de Doutorado, Universidade Federal do Rio de Janeiro).

Swales, J.M. (1990). *Genre analysis: english in academic and research settings*. Cambridge: Cambridge University Press.

Welfelê, O. (2004). A proveta arquivada: reflexões sobre os arquivos e os documentos oriundos da prática científica contemporânea. Tradução de Maria Celina de Melo e Silva. *Revista da Sociedade Brasileira de História da Ciência*, 2 (1), 65-72, jan./ jun.

Zipf, G.K. (1949) The meaning frequency relationship of words. *Journal of general Psychology*, 33, 251-256.

Zipf, G.K. (1949). *Human behavior and the Principle of Least Effort*. Cambridge, MA: Addison-Wesley.



ESTRUTURAS SEMÂNTICAS DA REPRESENTAÇÃO DO CONHECIMENTO ARQUIVÍSTICO: ANALISANDO ELEMENTOS ESTRUTURANTES E DE CONTEÚDO

Wanessa Rodrigues Martins¹, Thiago Henrique Bragato Barros²

¹Bacharel em Arquivologia pela Universidade Federal do Pará, Mestranda em Ciência da Informação-UFPA, 0000-0002-9648-8121, wanessa060734@gmail.com

²Professor Adjunto na Faculdade de Arquivologia e do Programa de Pós-Graduação em Ciência da Informação Universidade Federal do Pará. 0000-0001-7439-5779, thiagobarros@ufpa.br

RESUMO A representação arquivística faz parte de um desdobramento teórico-conceitual recente no âmbito da disciplina dos arquivos que visa integrar as atividades de classificação e descrição enquanto parte de um processo de representação e acesso aos documentos de arquivo. Visa-se com esta pesquisa contribuir de maneira significativa para a área de organização da informação, especialmente no que se relaciona a sistematização de metodologias. Tem-se como objetivo estabelecer parâmetros metodológicos para a construção textual de sentido nas descrições arquivísticas por meio da Semântica do Texto. Para este fim, tem-se como metodologia de pesquisa a semântica textual, visando a análise profunda de produções textuais técnicas advindas da representação em arquivos. Fixa-se na tradição brasileira visando contribuir para elaboração de normativas relacionada a área, analisa-se o Arquivo Público do Pará e o Arquivo Nacional, no que se refere as suas políticas em representação de arquivos, por meio de análise de catálogos de descrição de ambas as instituições tendo a semântica textual como método de análise. Busca-se contribuir de maneira significativa para o amadurecimento metodológico da arquivística e auxiliar na consolidação da representação arquivística enquanto uma abordagem metodológica na tradição iberoamericana. As análises realizadas demonstraram a aplicação de elementos da semântica textual aos instrumentos de descrição arquivística. Percebeu-se que a existência de vários elementos pertinentes à semântica do texto encontrados no texto como coesão, coerência, situacionalidade, informatividade, intencionalidade, fatores de contextualização e consistência e relevância, sendo comuns nos catálogos.

PALAVRAS-CHAVE *Representação Arquivística, Descrição, Semântica Textual, Instrumentos de Pesquisa; Análise Semântica.*

ABSTRACT Archival representation is part of a recent theoretical-conceptual development within the archival science that aims to integrate classification and description activities as part of a process of representation and access to archival documents. This research aimed to contribute significantly to the area of information organization, especially in what relates to the systematization of methodologies strategies in the scope of this process. The objective is to establish methodological parameters for the textual construction of meaning in the archival descriptions through semantics. For this purpose, we have as a research methodology the textual semantics, aiming at the deep analysis of textual productions techniques derived from the representation in archives. It is established in the Brazilian tradition in order to contribute to the elaboration of regulations related to the area; the Public Archive of Para and the National Archive are analyzed, with regard to their policies in archival representation, through analysis of catalogs describing both institutions having textual semantics as a method of analysis. It seeks to contribute significantly to the methodological maturation of archival science and to assist in the consolidation of archival representation as a methodological approach in the iberoamerican tradition. Carried out analyses demonstrated the application

of elements of textual semantics to the instruments of archival description. We noticed that the existence of several elements pertinent to the semantics of the text found in the text as cohesion, coherence, situationally, intentionality, contextualization factors, consistency, and relevance, being common in all catalogs.

KEY-WORDS *Archival Representation; Description; Textual semantics; Finding aids; Semantic analyses*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A Descrição e a Classificação foram as primeiras atividades Arquivísticas a trabalhar as questões técnicas e práticas em manuais. Neste sentido, estas funções marcam a inauguração da Arquivística enquanto uma disciplina técnico-científica a partir do final do século XIX, e, além disso, inauguram uma nova perceptiva no tratamento relacionado aos documentos de arquivos enquanto instituições socioculturais e o papel dos arquivistas, agentes responsáveis pela memória contida nos arquivos e pela imagem destas instituições perante a sociedade à qual estão inseridas.

A atividade de descrever além de proporcionar acesso, garante a preservação dos documentos. O arquivista para desenvolver as atividades de descrição fazia uso de conhecimentos oriundos, basicamente, de duas áreas correlatas do conhecimento, a paleografia e a diplomática, as quais o auxiliavam na leitura e transcrição dos documentos. Descrever significava representar, identificar e organizar, após esses procedimentos registrava-se as informações em instrumentos de pesquisa.

No decorrer da década de 1980 as atividades de descrição e arranjo passam a ser compreendidas de forma associada, começando então um processo de representação, como ressaltam autores como Lopes (2013), Tognoli (2012) e Barros (2014 & 2016). Agregando à descrição características de uma atividade não apenas de controle, mas de acesso. Quando trabalhos em prol da normalização começarão a surgir. Atribuindo, deste modo, mais parâmetros metodológicos à Arquivística.

Busca-se analisar a classificação e a descrição como parte constituinte do processo de Representação e acesso aos documentos de arquivo. Para tanto, tem em vista que a Representação Arquivística é resultado de um ciclo de atividades conjuntas inerentes a estas atividades. Esta visão integradora é recente e especialmente no contexto iberoamericano encontra-se em estado inicial no que diz respeito a pesquisa e produções metodológicas e sistematizadoras.

Fixa-se na tradição brasileira visando contribuir para elaboração de normativas relacionada a área, analisa-se o Arquivo Público do Pará e o Arquivo Nacional, no que se refere à produção de instrumentos de pesquisa, especificamente catálogos de descrição. Tendo como objetivos comparar as diferentes abordagens em representação nas diferentes instituições selecionadas, demonstrar como os arquivistas das duas instituições trabalham no que se refere à construção dos instrumentos de pesquisa, analisar as políticas existentes no que se refere a Representação Arquivística, investigar os instrumentos de pesquisa por meio da semântica textual e propor por meio da semântica textual princípios visando a sistematização do conteúdo descritos nos instrumentos de pesquisa.

Esta análise se deu em dois principais momentos. Em um primeiro momento será apresentado um levantamento bibliográfico a respeito da Representação Arquivística e Semântica Textual, seus

princípios teóricos, seu uso na atualidade por parte da Ciência da Informação, aspectos de sua constituição histórica e como efetuar a análise semântica dentro da Ciência da Informação.

Depois de discutirmos os preceitos da Semântica Textual, fez-se um levantamento predominantemente bibliográfico da teoria e da prática em relação à Representação Arquivística, visando compreender como se constituíram e se constituem a teoria e ao final efetuam-se as análises.

Trabalhou-se aspectos gerais da linguística enquanto um campo em desenvolvimento, buscou-se os marcos disciplinares na área da linguística quanto ao assunto em questão. Estudou-se de forma mais exaustiva as autoras Leonor Lopes Fávero e Ingedore G. Villaça Koch por representarem o pioneirismo e também a atualidade do tema em língua portuguesa.

A linguística textual é uma área da linguística recém-criada, inicialmente desenvolvida na Europa, especificamente na Alemanha, nos anos 1960. Ela concentra seus estudos, como o próprio nome remete, ao texto propriamente dito (Fávero & Koch, 1997).

As causas que levaram os linguistas a desenvolverem as gramáticas textuais justificam-se, segundo as autoras, devido “a ordem das palavras no enunciado, a relação tópico-comentário, a entoação, as relações entre sentenças não ligadas por conjunções, a concordância dos tempos verbais [...] com referência a um contexto situacional.” (Fávero & Koch, 1997, p. 16).

A inserção destas novas perspectivas com relação ao texto pode ser compreendida como o atendimento e a resposta a uma série de necessidades latentes ao âmbito textual, como as ressaltadas acima, que transpõem o nível da palavra e da frase tal análise precisa, a partir de então, ser mais abrangente e completa, e isto seria alcançado por meio do contexto textual.

As autoras destacam três períodos que foram essenciais na “passagem da teoria da frase à teoria do texto”, para isso apresentam as considerações de Conte (1977, apud, Fávero & Koch, 2012) em que o autor os elenca respectivamente como o da análise transfrástica, das gramáticas textuais e das teorias do texto. Vale ressaltar que esses momentos não são considerados como um percurso exatamente cronológico, mas “sim tipológico, por não haver, entre eles, uma sucessão temporal, constituindo-se cada um deles em um tipo diferente de desenvolvimento teórico” (Fávero & Koch, 2012, p. 17).

Na Linguística textual e na Teoria do Texto, o “texto” é o conceito principal, seja de cunho oral ou escrito expressos em pelo menos dois signos, precisa ser delimitado para poder ser trabalhado pela Linguística Textual, portanto, seu início e fim deverão ser previamente delimitados de maneira explícita. São exemplos diálogos, sermões, livros dentre outros que representam um aspecto delimitado em si. (Fávero & Koch, 2015).

Por este intermédio se traçará como se deu a constituição e o percurso teórico e conceitual da análise do texto e seus elementos essenciais para as análises que serão feitas dos planos de classificação e dos instrumentos de descrição das instituições selecionadas.

A construção do sentido textual engloba diversos elementos essenciais, necessários à compreensão semântica do texto. Com relação a isso as autoras elencam os elementos enunciados por Beugrande & Dressler (1981 apud Fávero & Koch, 2015) e por meio destes autores construiu-se as análises dos instrumentos de pesquisa de ambas as instituições.

O método da pesquisa caracteriza-se por ser um estudo exploratório, teórico e documental, tendo a Semântica Textual como uma metodologia de análise para a compreensão dos caminhos da Representação Arquivística por meio dos planos de classificação e instrumentos de pesquisa das instituições analisadas.

A aplicação da Semântica aos produtos da representação, a saber, os catálogos das instituições estudadas, visa subsidiar metodologicamente a discussão a respeito da criação de descritores com conteúdos normalizados no contexto da representação em arquivos.

Como procedimentos adotados para analisar os catálogos de descrição do Arquivo Público do Estado do Pará e do Arquivo Nacional serão aplicados os elementos definidos no tópico anterior.

A análise se deu de maneira prática e sucinta, desse modo, cumprindo o objetivo para o qual se propôs esta pesquisa. Destacando os critérios de contextualização dos instrumentos, de consistência, relevância, focalização, ou seja, todos os requisitos destacados, buscando a existência, ou não, destes elementos.

METODOLOGIA

2 ABORDAGENS LINGUÍSTICAS EM CIÊNCIA DA INFORMAÇÃO: LINGUÍSTICA TEXTUAL ENQUANTO UMA ABORDAGEM DE PESQUISA

A inserção destas novas perspectivas, conforme destacado na introdução, com relação ao texto pode ser compreendida como o atendimento e a resposta a uma série de necessidades latentes ao âmbito textual, como as ressaltadas acima, que transpõem o nível da palavra e da frase tal análise precisa, a partir de então, ser mais abrangente e completa, e isto seria alcançado por meio do contexto textual.

As autoras destacam três períodos que foram essenciais na “passagem da teoria da frase à teoria do texto”, para isso apresentam as considerações de Conte (1977, apud, Fávero & Koch, 2012) em que o autor os elenca respectivamente como o da análise transfrástica, das gramáticas textuais e das teorias do texto. Vale ressaltar que esses momentos não são considerados como um percurso exatamente cronológico, mas “sim tipológico, por não haver, entre eles, uma sucessão temporal, constituindo-se cada um deles em um tipo diferente de desenvolvimento teórico” (Fávero & Koch, 2012, p. 17).

No primeiro momento, o da análise transfrástica, a pesquisa ainda estava inclinada à análise do enunciado ou da sequência de enunciados e o seu percurso até o texto. Esse estudo objetivava inferir quais relações poderiam existir entre os enunciados.

Na visão das autoras, esse primeiro momento, ainda que significativo por transpor o nível da frase, não configurou autonomia ao tratamento do texto e nem pode construir um modelo teórico para servir como guia às pesquisas relacionadas. (Fávero & Koch, 2012).

O segundo momento, das gramáticas textuais, visava estudar fenômenos linguísticos dos quais a gramática do enunciado não explicava. A gramática textual considerava que entre o texto e o enunciado havia uma diferença qualitativa e que o texto era “mais que uma sequência de enunciados” (p.19), sendo, portanto, a sua compreensão e produção resultadas da “competência textual”. Esta competência elencada pelas autoras consiste em:

- a) verificar o que faz com que um texto seja um texto, isto é, determinar seus *princípios de constituição*, os fatores responsáveis pela sua *coerência*, as condições em que se manifestam a textualidade (Texthaftigkeit);
- b) levantar critérios para a delimitação de textos, já que a completude é uma das características essenciais do texto;
- c) diferenciar as várias espécies de textos. (Fávero & Koch, 2012, p. 19).

As autoras destacam o modelo de gramática textual de Petöfi, pois é o que mais se destacou nesse segundo momento. Petöfi acreditava em um modelo de gramática textual de base não linear, isto quer dizer, que não seguia um padrão pré-estabelecido, e nem formal, pois esta base textual constaria de “uma *representação semântica*, indeterminada com respeito às manifestações lineares das sequências dos enunciados” (Fávero & Koch, 2012, p. 20. Grifo nosso).

Petöfi postula ser este modelo de gramática apto a tornar possível:

- a) a análise de textos, isto é, a atribuição a uma manifestação linear de todas as bases textuais possíveis; b) a síntese de textos, ou seja, a geração de todas as bases possíveis textuais; c) a comparação de textos. Neste modelo, o léxico, com suas representações semânticas intencionais, assume função relevante. (Fávero & Koch, 2012, p. 20).

O terceiro e último momento é o das teorias do texto. Desse momento em diante recebem maior destaque o contexto pragmático do texto.

O âmbito de investigação se estende do texto ao contexto, entendido, em geral, como conjunto de condições – externas ao texto – da produção, recepção e da interpretação do texto. (Fávero & Koch, 2012, p. 20).

Sobre este aspecto Linguístico – textual, vários autores fizeram considerações a exemplo de Dressler, Schmidt e Oller. A pragmática do texto foi abordada de diversas maneiras. A seguir observam-se estas definições por meio de um quadro conceitual para melhor visualização e compreensão.

Infere-se, a partir do percurso traçado, que a Linguística Textual contemporânea estende sua análise do texto ao contexto, isto é, aos fatores que se encontram para além do texto. Esta forma de analisar não só é mais abrangente como muito mais profunda em termos semânticos do que até então se havia proposto em relação ao texto. É, portanto, esse viés mais prático e palpável, isto é, pragmático, que vai trazer uma renovação identitária à Linguística Textual a partir do século XX.

DESENVOLVIMENTO

3 REPRESENTAÇÃO ARQUIVÍSTICA

A descrição esteve presente sob diversas formas as quais consistiam ou na íntegra transcrição do documento (cópias), a fim de se preservar o original, ou na elaboração de inventários, guias e catálogos elementos que representavam de forma sucinta o acervo e a documentação.

Estes instrumentos passaram a ser orientados, a partir do século XIX, por manuais e princípios, a exemplo do Manual de Arranjo e descrição de arquivos cunhado pelos arquivistas holandeses em 1889,

do princípio de respeito aos fundos atribuído ao historiador francês Natalis de Wally instaurado por meio de uma circular em 1841, e no século seguinte por meio do livro de Hillary Jenkison sob o título: *A manual of archive administration: including the problems of war archives and archive making* (Barros, 2014).

Arguin, destaca que “a publicação de manuais representa uma etapa na constituição de uma disciplina científica” *et al* (Rousseau & Couture, 1998, p. 53). Com isto pode-se inferir que com a aquisição do status de disciplina os preceitos práticos, teóricos e metodológicos da Arquivística difundiram-se e adquiriram consistência no ambiente acadêmico proporcionada pela pesquisa, em especial, e consequentemente por meio de suas publicações.

As concepções e transformações ocorridas na Arquivística em grande parte são reflexos do contexto social e político em que se viveu (e vive). Por exemplo, as legislações criadas a partir dos governos democráticos instauram a transparência das atividades administrativas, judiciais e legislativas e, por conseguinte, possibilitam o acesso aos registros. (Rousseau & Couture, 1998).

Na Arquivística estas mudanças são perceptíveis por caracterizarem períodos considerados renovadores, seja na prática ou teoria arquivística, como a revisão de princípios, a concepção do objeto de estudo e a adaptação de seus métodos às tecnologias emergentes e a abertura dos espaços de arquivos de modo mais abrangente à sociedade.

Sobre estas mudanças Tognoli (2012, p.82) destaca que “uma mudança de paradigma é anunciada na Arquivística quando Hugh Taylor admite a obsolescência dos princípios e métodos arquivísticos promulgados no século anterior, caracterizando o final do século XX como um período de revolução científica...” isto significa uma profunda reformulação na estrutura teórica e metodológica na Arquivística, que para o contexto já não correspondia de forma satisfatória.

Inclusive quando se percebe a aproximação dos arquivistas e da Arquivística às relações interdisciplinares com a Ciência da Informação e Teoria da Informação, como no caso dos trabalhos de Fonseca (2005) e Taylor (1984).

Deste modo as normas de descrição arquivística são um dos possíveis resultados destas transformações à medida foram pensadas para refletirem todo o processo da produção documental, deste a criação à guarda permanente. A partir de quando discutidas, em meados da década de 1980 e promulgadas nos anos de 1990-2000, passam a ser um dos principais meios de representação em arquivos seja em âmbito analógico ou digital, especialmente nos países que não possuíam iniciativas nacionais de normalização.

No contexto brasileiro as normas de descrição chegaram um pouco depois. Apenas após a participação do Brasil na revisão da 2ª edição da ISAD(G). Na qual tinha como tarefa primordial “traduzir a norma e divulgá-la o mais amplamente possível. O Arquivo Nacional assumiu essa tarefa e no início de 1998 publicou a primeira edição brasileira das normas internacionais ISAD(G) e ISAAR (CPF)” (Conarq, 2006, p. 8).

Sobre isto Lopes elenca uma série de fatores que considera terem sido influenciados positivamente em virtude de se ter um parâmetro, em âmbito internacional, pelo qual se guiar, como se vê “Os arquivistas ganharam, com estas normas, um importante instrumento para o seu trabalho. [...]O avanço das normalizações, quiçá em outras áreas, contribuirá para quem deseja uma arquivística teórica e prática, uma ciência social aplicada, como qualquer outra” (Lopes, 2013, p. 299).

Esta consolidação metodológica da Arquivística em grande parte é fruto da corrente integradora enunciada nos anos 1980 por Rousseau e Couture onde as funções arquivísticas: classificação e descrição passam a ser compreendidas como fases de um processo, uma associação, um trabalho contínuo e interligado. Isto fica entendido quando os autores refletem sobre a informação e seu contexto de produção.

“ A produção de informações orgânicas registradas dá origem aos arquivos do organismo. Sob esta designação são agrupados todos os documentos, seja qual for o seu suporte e idade, produzidos e recebidos pelo organismo no exercício das suas funções.” (Rousseau, J., & Couture, C. 1998, p. 65)

Rousseau e Couture indicam que focar no documento é algo muito generalizado (constatação bastante presente na Arquivística dos manuais) e que, portanto, esse valor - informação orgânica - seria obtido através da informação que aquele documento tivesse, daí então aparecem termos como informação orgânica e não orgânica.

A representação arquivística, conforme Lopes (2013) visa dentre outras coisas olhar a informação desde suas primeiras fases perpassando pela avaliação, classificação, e finalmente e de modo mais profundo e específico nos arquivos permanentes com a descrição.

A dinamicidade do processo informacional vai além da simples análise do documento pressupõe um estudo, uma pesquisa. “A conexão entre o documento e seu contexto é um processo intelectual que é materializado por meio dos instrumentos de pesquisa resultados dos processos de representação do conhecimento arquivístico e entendidos como uma ponte entre o usuário e o documento” (Tognoli, 2012, p.89).

Portanto, o processo de representação é muito mais complexo e amplo do que faz-se parecer, especialmente quando se reflete além da visão unilateral dos produtos do processo, no caso, planos de classificação e instrumentos de pesquisa. Nesta perceptiva integradora se reconhece a representação descritiva como um processo contínuo e mutável, passível de mudanças, acréscimos e alterações especialmente porque a informação é dinâmica, em relação a sua natureza e seu uso, por ser reflexo do contexto social em que se insere e expressa o qual vive em constante ressignificação.

4 ANÁLISE DAS PRÁTICAS EM REPRESENTAÇÃO ARQUIVÍSTICA NO CONTEXTO BRASILEIRO: ARQUIVO NACIONAL E ARQUIVO PÚBLICO DO ESTADO DO PARÁ

Para fins desta análise e aplicação prática da teoria linguística e arquivística apresentadas, selecionou-se estas duas instituições arquivísticas por se tratarem das principais referenciais em âmbito nacional e estadual de prática arquivística, fundamentais em seus cenários de atuação, do fato de tratarem-se de instituições em instâncias diferentes, estabelecendo assim um universo representativo de análise.

Foram selecionados dois catálogos de ambas as instituições Do Arquivo Nacional os catálogos referentes aos seguintes fundos:

- Instituto de Pesquisas e Estudos Sociais – IPES (Catálogos de Filmes)

No Arquivo Público do Estado do Pará as correspondentes à;

- Secretaria da Capitania (Catálogo de Documentos Manuscritos – Período Colonial 1649 – 1823), e, ainda o:

Os catálogos são considerados um dos principais e mais antigos instrumentos de pesquisa no âmbito dos arquivos. Ainda hoje são produzidos e em muitos deles já se encontra em seu conteúdo estrutura baseada no formato das normalizações arquivísticas, especialmente ISAD (G) e NOBRADE (Norma Brasileira de descrição Arquivística) como é o caso dos catálogos encontrados no Arquivo Nacional.

Para tanto faz-se necessário buscar a definição de catálogo. De acordo com o Dicionário Brasileiro de Terminologia Arquivística consiste em “Instrumento de pesquisa organizado segundo critérios temáticos, cronológicos, onomásticos ou toponímicos, reunindo a descrição individualizada de documentos pertencentes a um ou mais fundos, de forma sumária ou analítica”. (ARQUIVO NACIONAL, 2005, p.45)

Os catálogos e a descrição das séries são desde os primórdios da organização de arquivos os instrumentos mais utilizado nas instituições arquivística. Assim, sua análise torna-se fundamental e representativa no estudo proposto. Os catálogos referentes a esta análise são organizados temática e cronologicamente cada um corresponde a um fundo e específico, sendo um destes de natureza pessoal, o de Apolônio de Carvalho, e, os demais de instituições públicas brasileiras que se encontram custodiados pelos respectivos arquivos, Nacional e Estadual do Pará

4.1 Análise dos Catálogos do Arquivo Nacional: Instituto de Pesquisas e Estudos Sociais – IPES (Catálogos de Filmes)

INSTITUTO DE PESQUISAS E ESTUDOS SOCIAIS – IPES (CATÁLOGOS DE FILMES)	
Coesão Textual	Há remissão e referencia quanto ao assunto tratado o que garante a sequência do sentido no texto.
Coerência	As ideias estão estruturadas de maneira lógica no corpo do texto.
Situacionalidade	<p>Da Situação para o texto: a produção do acervo é diretamente influenciada pela situação contemporânea à época e se faz relevante a medida que faz ser conhecida a partir da representação descritiva, apresentada no catálogo, o conhecimento social-político e cultural de uma época, e, se trata da representação textual de acervos filmográficos e iconográficos que registra a sua principal função que “era integrar os diversos movimentos sociais de direita para criar as bases de uma oposição que pudesse deter o “avanço do comunismo soviético no Ocidente”. Dessa forma, o Instituto promoveu intensa campanha antigovernamental. Associando as propostas do governo ao comunismo, o IPES utilizou os mais diversos meios de comunicação na defesa da democracia e da livre iniciativa. Publicou artigos nos principais jornais do país; produziu uma série de 14 filmes de "doutrinação democrática", apresentados em todo o país; financiou cursos, seminários, conferências públicas; publicou e distribuiu inúmeros livros, folhetos e panfletos anticomunistas, dentre os quais UNE, instrumento de subversão, de Sônia Seganfredo, dirigido aos estudantes universitários, então considerados como um dos pilares da infiltração comunista[...]” (Arquivo Nacional, 2012, p.6).</p> <p>É um texto extremamente formal e com uma variedade linguística não variada por conta de seu contexto mais especificamente político e suas ações em torno disso.</p>

	Do texto para a situação: levando em consideração os aspectos mediadores presentes tanto na produção quanto na inferência de um texto, percebeu-se que a descrição procurou manter-se isenta de aspectos pessoais embora não represente de modo tão real quanto o foi no período em que ocorreu. Procura mediar assim as informações sendo concisa especialmente nos campos “conteúdo” e “indexação”.
Informatividade	Distribuição da informação: Apresenta informações novas referente a IPES e o conteúdo do acervo associando e contextualizando com informações referentes a acontecimentos contemporâneos à época tanto em âmbito nacional quanto internacional. Combina em sua estrutura os movimentos de retroação e progressão, fazendo referencia à informações já dadas para a progressão do sentido textual. Grau de Previsibilidade/Expectabilidade da Informação: sendo este elemento mais comum na linguagem poética e metafórica, pela imprevisibilidade do enredo e por procurar superar ou ao menos preencher as expectativas, não se espera que um Instrumento de Descrição Arquivística não seja previsível ou não supra a expectativa do usuário, em verdade, esse é o objetivo. Nesse sentido não se pode considerá-lo menos informativo porque, simplesmente, sua forma não se enquadra na poesia e nem na metáfora.
Intertextualidade	É possível a inferir o sentido do texto sem a necessidade de leitura prévia de outros textos. O texto conseguiu fazer compreensível seu sentido. Este aspecto não é comum neste tipo de texto, embora o seja nos demais textos descritivos que não os arquivísticos. É o que se observa nos instrumentos e normas.
Intencionalidade	Um texto de descrição documental é extremamente intencional. Sua principal função é realizar a ação de informar/comunicar. Este aspecto, portanto, é encontrado neste texto.
Aceitabilidade	O sentido do texto justifica sua aceitabilidade. Compreende-se que se não houvesse coerência e coesão seria inaplicável a aceitabilidade por se tratar de um instrumento informativo, sendo assim, o Princípio de Cooperação de Grice (1975) é inviável de postular entre o texto e o leitor/usuário em caso de incoerências. Portanto, não aplicável em textos descritivos arquivísticos.
Fatores de Contextualização	Fator evidenciado em cada elemento de descrição do Catálogo. A exemplo da datação, tema, local, proveniência, dentre outros.
Consistência e Relevância	Considera-o consistente por não haver contradição no texto, e, desse modo, se faz verdadeiro. A relevância se dá por nas descrições dos vídeos se fazerem presentes a temática relacionada às funções da IPES.
Focalização	Não se aplica.

Fonte: Elaborado pelos autores. Dados da pesquisa 2016.

4.2 Análise dos catálogos do Arquivo Público do Estado do Pará: Catálogo de Documentos Manuscritos da Secretaria da Capitania (Período Colonial 1649 – 1823)

Quadro 2 – Análise textual do Catálogo de Documentos Manuscritos da Secretaria da Capitania (Período Colonial 1649 – 1823)

SECRETARIA DA CAPITANIA (CATÁLOGO DE DOCUMENTOS MANUSCRITOS – PERÍODO COLONIAL 1649 – 1823)	
Coesão Textual	Remissiva/ referencial: presente pela tomada de dados anteriores. Sequencial: o sentido texto é continuado. Percebe-se que estes elementos e complementam e interconectam no texto e são essenciais a este.
Coerência	Lógica estruturada nos elementos de constituição do texto apresentado.
Situacionalidade	Da situação para o texto: As situações contemporâneas ao texto delimitam-se ao governo do antigo Estado do Grão-Pará, Maranhão e Rio Negro, reunindo documentos datados dos séculos XVII, XVIII e XIX até o ano de 1840., representam as relações deste com demais secretarias, governos e particulares, sendo, portanto, a documentação, a expressão exata destas situações. Do texto para a situação: a narração descritiva procura ser fiel ao conteúdo documental sem expressar expressões pessoais, mas permitir que esta possa ser recebida de modo mais semelhante ao que está documentado, ainda que se em totalidade não o for.
Informatividade	Distribuição da informação: informações são retomadas no decorrer da descrição para que o sentido do texto e o assunto tenha continuidade dessa maneira e progrida coesa e coerentemente. Grau de Previsibilidade/Expectabilidade da Informação: Como já citado, os textos descritivos arquivísticos possuem caráter objetivo e específico se propõem a representar em maior grau, isto é, de maneira mais minuciosa e aprofundada a informação, mas também de força concisa. Portanto, possui um alto grau e função informativa, mesmo que preveja seu conteúdo.
Intertextualidade	Não aplicável.
Intencionalidade	Todo texto descritivo arquivístico é intencionalmente produzido. Intencionalidade nesse contexto refere-se tanto à organização da informação, como também, à intenção de comunicar.
Aceitabilidade	Não se aplica.
Fatores de Contextualização	Da “Área de Identificação” à “Área de Notas” deste Catálogo encontram-se os exemplos de fatores de contextualização.
Consistência e Relevância	Consistente pela verdade traçada nos tópicos e itens do texto. Desse modo um tópico atribui relevância ao seguinte por dar continuidade ao sentido textual.
Focalização	Não se aplica.

Fonte: Elaborado pelos autores. Dados da pesquisa 2016.

As análises realizadas demonstraram a aplicação de elementos da semântica textual aos instrumentos de descrição arquivística. Percebeu-se que a existência de vários elementos pertinentes à semântica do texto encontrados no texto como coesão, coerência, situacionalidade, informatividade, intencionalidade, fatores de contextualização e consistência e relevância, sendo comuns nos quatro catálogos.

Constatou-se ainda a inviabilidade de alguns elementos na análise dos textos descritivos arquivísticos, como a aceitabilidade é justificada, como se pode ver, no corpo da análise acima, ambos estão relacionados ao sentido do texto.

A aceitabilidade prevê que se aceite o texto ainda que este não possua sentido ou que este esteja explícito, todavia, como já dito, num texto descritivo arquivístico, os elementos que constroem o sentido textual precisam estar bem evidentes de modo a oferecer maior compreensão acerca da documentação e do arquivo.

A focalização se refere às diversas formas de interpretação que um texto possui. A depender do foco que se dá à interpretação de um texto ele pode assumir vários sentidos, na medida em que seus usuários fizerem suas buscas e interpretações. Vê-se a focalização como elemento de possível identificação nos textos de descrição arquivística por se compreender que a visão do leitor e a contextualização podem influenciar na interpretação deste, especialmente, lhe proporcionando ressignificação.

Todavia, por se referirem a ações interpretativas dos usuários e os vários sentidos que um único texto pode ter a partir das diferentes focalizações estabelecidas e buscadas pelos usuários, ressalta-se a grande relevância desses aspectos, mas considera-se que estes são elementos destinados ao uso do texto e ao usuário, portanto, se dá em um momento posterior ao buscado na análise deste trabalho, que é a construção e produção do texto, momento anterior a este.

Destaca-se a intertextualidade, por sua vez, por proporcionar a conexão com outros textos. Entende-se que seria necessário e importante traçar na descrição relações que os documentos pudessem ter com outros fundos arquivísticos (pessoas físicas ou jurídicas). Nos fundos analisados a intertextualidade não foi identificada.

Outro destaque na análise é a empregabilidade da Norma Brasileira de Descrição Arquivística - NOBRADE – aos catálogos. Em três destes a descrição esta normalizada, e, esta constatação faz uma grande diferença em profundidade e contextualização da representação do conteúdo dos acervos.

A ausência de normalização arquivística é sentida no catálogo do Ministério da Marinha, em que, embora seja referente a apenas uma parcela da documentação, as informações apresentadas no catálogo são plenamente compreensíveis, mas uma contextualização mais aplicada se faz importante e necessária para dimensão das informações que se supõe ser produzida em um Ministério.

Portanto, a construção dos instrumentos, quando desvendados por meio da semântica textual, fica clara a sua intencionalidade, subjetividade, intertextualidade e referência. Assim, a Representação Arquivística, é encarada neste trabalho de relacionar as instâncias em que o texto se inscreve. Visa-se assim demonstrar aspectos constituintes e fundamentais destes produtos.

CONSIDERAÇÕES FINAIS

É preciso que os arquivistas, cientistas da informação, bibliotecários, museólogos, admitam acima de qualquer suspeita a falta de parâmetros claros quando se fala, da elaboração precisa das ações de síntese e análise. Em arquivística, especialmente, existem sim a formalização e a normalização, porém, essas ações não aprofundaram-se da maneira devida na questão fundamental dos descritores: seu conteúdo e melhor forma de apresentação.

Este trabalho visou, ao logo de seu desenvolvimento, destacar a importância de relações interdisciplinares entre as áreas para solução de problemas, quando busca na Linguística formas de análise e compreensão dos fenômenos de nossa área.

Destacou-se também a relação existente entre os processos de descrição e classificação, sempre compreendidos da prática de forma integrada porém tradicionalmente pouco discutido na teoria da área.

Por fim, cumpriu-se de maneira plena com os objetivos que se propôs no início de não tão simples trajetória de pesquisa. Sugere-se novos estudos e em nível maior de aprofundamento relacionado arquivística, linguística e ciência da informação.

Afinal, é só por meio de trabalhos dessa natureza que nossa área irá construir-se de forma plena e que responda a questões postas aos arquivistas pela sociedade.

É necessário, então, uma revisão de nossas normas para padronizar não só suas tabelas, linguagem de máquina e apresentação. É fundamental também, padronizar conteúdos e seus respectivos contextos afim de facilitar a busca por parte do usuário.

O trabalho apresentou amplamente os conceitos de texto e a semântica do texto. Proporcionou, com isso, subsídio para a posterior análise dos catálogos descritivos. Abordou a temática da representação no contexto da Arquivística e da Ciência da Informação e também demonstrou que a representação Arquivística é uma construção teórica dos últimos trinta anos.

Constatou na análise a inexistência de políticas de descrição arquivística institucionalizada e de parâmetros claros para a construção dos descritores e dos instrumentos de pesquisa de modo geral.

Embora exista os meios formulares e normalizados de descrição na atualidade, quando passamos para uma análise profunda dos mesmos, percebe-se acima de tudo a subjetividade, intencionalidade e outras ações destacadas ao longo da análise, existentes nestes textos assim como na maioria dos textos, ainda que se esteja falando de textos técnico-científicos.

REFERÊNCIAS BIBLIOGRÁFICAS

ARQUIVO NACIONAL (Brasil).(2000). ISAD(G): *Norma geral internacional de descrição arquivística*. Rio de Janeiro: Arquivo Nacional.

ARQUIVO NACIONAL (Brasil).(2005). *Dicionário brasileiro de terminologia arquivística*. Rio de Janeiro: Arquivo Nacional.

BARROS, T. H. B.(2014) *A Representação da informação Arquivística: Uma Análise do discurso teórico e institucional a partir dos contextos Espanhol, Canadense e Brasileiro*. Tese de Doutorado, Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, São Paulo, Brasil.

BARROS, T. H. B.(2016) A Indexação e a Arquivística: aproximações iniciais no universo teórico da organização e representação do conhecimento. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, 21:46, 33-44.

KOCH, I. G. V. (1997) *Linguística textual: retrospecto e perspectivas*. Alfa: São Paulo, 41: 67-78.

KOCH, I. V.(2015) *Introdução à Linguística Textual: trajetória e grandes temas.* – 2. ed. – São Paulo : Contexto.

LOPES, Luís Carlos.(2013) *A nova arquivística na modernização administrativa.* 3. ed. Brasília: Annabel Lee.

TOGNOLI, N. B. (2012). *A representação na arquivística contemporânea.* Ci. Inf., ISSN 1983-5213, Brasília, 5: 2, 79-92.

ROUSSEAU, J. ; COUTURE, C. (1998). *Os fundamentos da disciplina arquivística.* Varona - Salamanca: Publicações Dom Quixote.



AS PESQUISAS SOBRE PERIÓDICO CIENTÍFICO PUBLICADAS NO BRASIL: UMA REVISÃO SISTEMÁTICA DE LITERATURA

Jorge Santa Anna¹, Maria Elizabeth de Oliveira Costa²

¹Universidade Federal de Minas Gerais. 0000-0002-0709-3639, professorjorgeufes@gmail.com

²Universidade Federal de Minas Gerais, mabethcosta@gmail.com

RESUMO Constitui uma revisão sistemática da literatura sobre o tema periódicos científicos realizada na Base de Dados em Ciência da Informação. Objetiva identificar os artigos que versam sobre essa temática, de modo a apresentar uma lista de referenciais teóricos e refletir sobre os principais assuntos contemplados nessas pesquisas. Por meio dos descritores periódico científico, periódico eletrônico e portais de periódicos, selecionou-se um total de 52 trabalhos publicados no período de 1972 a 2017. Foi possível constatar como assuntos recorrentes na literatura, alguns aspectos relativos aos periódicos eletrônicos e portais, tais como: o processo de migração do impresso ao eletrônico, a adesão ao acesso aberto e o uso de softwares livres, como o Sistema Eletrônico de Editoração de Revistas, além dos procedimentos de avaliação, uso e satisfação do usuário quanto da busca e recuperação da informação nos ambientes digitais. Sugere-se nova revisão sistemática em base de dados internacional como também o mapeamento da produção científica considerando ano de publicação, autorias, periódicos utilizados para publicação e métodos de pesquisa, no conjunto dos referenciais teóricos analisados.

PALAVRAS-CHAVE *Periódico científico, Periódico científico eletrônico, Portais de periódicos, Acesso aberto, Sistema eletrônico de editoração de revistas.*

ABSTRACT It is a systematic review of the literature on the topic of scientific journals held in the Database in Information Science. It aims to identify the articles that deal with this topic, in order to present a list of theoretical references and to reflect on the main subjects contemplated in these researches. A total of 52 papers published in the period from 1972 to 2017 were selected through the periodic descriptors, scientific periodicals, electronic journals and periodicals. It was possible to verify as recurrent subjects in the literature some aspects related to electronic journals and portals, such as : The process of migration from print to electronic, adherence to open access and the use of free software, such as the Electronic Publishing System, as well as user evaluation, use and satisfaction procedures as well as search and retrieval of information in Digital environments. It is suggested a new systematic review in international database as well as the mapping of scientific production considering year of publication, authors, periodicals used for publication and research methods, in the set of theoretical references analyzed.

KEYWORDS *Scientific journal, Electronic scientific periodical, Journal portals, Open access, Electronic publishing system of journals.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

No campo da Ciência da Informação, nos últimos anos, a revisão sistemática da literatura vem ocupando espaço, sendo publicada, normalmente, em anais de eventos, como também em periódicos especializados da área, garantindo, nesse contexto, para o aprofundamento dos estudos aplicados realizados nas instituições de ensino e de pesquisa. Presumimos a importância do emprego desse tipo de estudo na Ciência da Informação, principalmente, pela escassez de procedimentos metodológicos específicos consolidados pela literatura dessa área do conhecimento, conforme proclamado por Mueller (2007) e Valentim (2005).

Como exemplos de alguns estudos recentes realizados na Ciência da Informação, cujo método de investigação foi a revisão sistemática, pode-se citar a pesquisa de Autran *et al.* (2016), sobre o tema gestão do conhecimento, publicada em periódico científico e a pesquisa de Santa Anna (2016), a respeito da temática Gestão de Coleções Bibliográficas, apresentada em evento da área. Salienta-se que primeiro trabalho foi publicado na Revista Biblionline, em João Pessoa – Paraíba, Brasil. Já o segundo estudo foi apresentado no X Encontro da Associação de Educação e Investigação em Ciência da Informação na Iberoamérica e no Caribe, evento internacional ocorrido de 21 a 23 de setembro de 2016, na cidade de Belo Horizonte, Brasil. De modo geral, esses estudos disponibilizaram informações sobre as descobertas alcançadas e forneceram embasamento teórico e metodológico, proporcionando a formação de listas de referenciais teóricos para sustentação de estudos futuros relacionados à temática de pesquisa.

Dentre os diversos temas que instigam a realização de pesquisas mais profundas no campo da Ciência da Informação, no Brasil, cita-se a questão dos periódicos científicos. Em linhas gerais, essas publicações, embora sejam consideradas como o principal veículo de comunicação científica e disseminação do conhecimento (Targino, 2001, Fachin & Hillesheim, 2006), atravessam momentos de redefinição, sobretudo com o surgimento dos periódicos eletrônicos (Mueller, 2003). Com a incorporação das tecnologias digitais, alguns desafios são impostos, tais como: a necessidade de agrupamento dessas publicações em portais institucionais (Garrido & Rodrigues, 2010), o tratamento conferido pelas bibliotecas digitais no que tange à busca e recuperação da informação nos portais (Cendón, Souza & Ribeiro, 2011) e, por fim, questões relacionadas ao acesso aberto, direitos autorais e preservação digital (Mueller, 2006 & Sayão, 2010), dentre inúmeras outras questões a serem resolvidas.

Diante da importância atribuída ao método da revisão sistemática e, considerando a relevância do tema relativo aos periódicos científicos e os diversos desafios inseridos no âmbito desse tema, o artigo que ora se apresenta objetiva apresentar o que tem sido publicado na literatura da área de Ciência da Informação no Brasil sobre o tema periódicos científicos, considerando como fonte de pesquisa, os artigos de periódicos divulgados nas principais revistas dessa área de conhecimento.

A partir dos resultados alcançados com esta revisão, será possível mapear as pesquisas que foram e estão sendo realizadas, em face do recorte temporal considerado para seleção das fontes de investigação. Portanto, serão apontados quais assuntos e métodos de investigação permeiam a literatura até então publicada sobre o tema periódicos científicos no contexto brasileiro. A partir dessas constatações, espera-se que os estudos mapeados possam servir como referenciais teóricos para embasar o trabalho dos pesquisadores, como também contribua como sugestões de bibliografias direcionadas à prática docente.

METODOLOGIA

Em obediência aos rigores metodológicos estabelecidos na revisão sistemática, a fim de se obter a lista completa dos trabalhos para análise de seus assuntos, procedeu-se às seguintes etapas iniciais: 1 - escolha do ambiente para seleção das fontes; 2 - processo de seleção das fontes; e, por fim, 3 - exposição da amostra de pesquisa, conforme recomendado nos estudos de Rother (2007) e Galvão e Pereira (2014) sobre os caminhos que norteiam esse tipo de revisão.

Portanto, para início de conversa, optou-se em analisar apenas os trabalhos publicados na modalidade artigo de periódico. Essa escolha se deve ao fato dos periódicos constituírem como os principais meios de comunicação utilizados por cientistas de todo o mundo. A fim de se obter os artigos, recorreu-se à consulta ao catálogo da Base de Dados em Ciência da Informação (Brapci), haja vista a importância dessa base para a Ciência da Informação e por sua especialidade na indexação de periódicos dessa área de conhecimento no âmbito brasileiro.

Sobre essa base de dados, importante mencionar que ela manifesta-se como um produto gerenciado pela Universidade Federal do Paraná, cujo objetivo é subsidiar estudos e propostas na área de Ciência da Informação, fundamentando-se em atividades planejadas institucionalmente. Para elaboração da referida base de dados, foram identificados os títulos de periódicos da área de Ciência da Informação e indexados seus artigos, constituindo-se a base de dados referenciais. Sendo assim, confirma-se que a Brapci amplia o espaço documentário permitido ao pesquisador, facilita a visão de conjunto da produção na área, ao mesmo tempo, que revela especificidades do domínio científico (Brapci, 2017).

No que tange ao atual acervo dessa base, destaca-se que existem 16.555 trabalhos distribuídos em 35 revistas cadastradas, de abrangência nacional, 2.244 trabalhos em eventos, além de uma tese e de um livro indexado. Ademais, ressalta-se que, dos periódicos disponíveis, 27 estão ativos e 8 históricos (descontinuados) (Brapci, 2017).

No que se refere ao intervalo de tempo considerado para seleção dos artigos, essa questão fidelizou a delimitação da busca inserida no próprio catálogo da Brapci, ou seja, recuperaram-se, tão somente, os estudos publicados entre o ano de 1972 a 2017, não tendo nenhuma restrição quanto ao idioma da publicação. Ressalta-se que todo o acervo da Brapci é formado por estudos publicados entre esses anos, incluindo-se as publicações impressas, as quais foram devidamente digitalizadas a fim de compor o acervo digital.

Quanto às estratégias para busca e recuperação das fontes indexadas na Brapci, utilizaram-se os seguintes descritores no catálogo eletrônico, a saber: 1 - periódico científico, 2 - periódico eletrônico, e 3 - portais de periódicos. Esses descritores foram escolhidos em face da representatividade conceitual que contemplam no âmbito da temática periódico, e por eles estarem inseridos em diversos vocabulários controlados para indexação. Por conseguinte, para seleção da amostra de pesquisa, procedeu-se à prática da leitura ao título, resumo, introdução e conclusão dos trabalhos recuperados.

Os procedimentos criteriosos para formação da amostra de pesquisa podem ser visualizados na tabela 1.

Tabela 1. Rigores metodológicos para seleção da amostra de pesquisa

Do tipo de publicação a ser analisada	Artigos de periódicos
Do ambiente de investigação	BRAPCI
Da delimitação do tempo	De 1972 a 2017
Dos descritores utilizados na busca dos artigos	1 - periódico científico, 2 - periódico eletrônico, 3 – portais de periódicos
Das estratégia para identificação dos assuntos dos artigos	Leitura ao título, resumo, introdução e conclusão dos artigos

Referência: dados da pesquisa (2017)

Especificamente, quanto ao processo de busca por meio dos descritores, e, posteriormente, após as atividades de leitura, muitos trabalhos foram descartados em face de não contemplarem, especificamente, assuntos relativos aos periódicos científicos como meios de comunicação científica. A tabela 2 demonstra as ocorrências após o processo de busca e posterior à atividade de leitura para seleção dos materiais.

Tabela 2. Comparativo entre a quantidade de artigos recuperados e os selecionados após a prática da leitura

Descritores	Artigos recuperados	Artigos selecionados após leitura
Periódico científico	40	25
Periódico eletrônico	34	22
Portais de periódicos	6	5

Referência: dados da pesquisa (2017)

Sendo assim, após os procedimentos de seleção das fontes de pesquisa, a amostra para análise foi composta por um total de 52 artigos, distribuídos em variadas revistas da Ciência da Informação no Brasil e que se encontram cadastradas na Brapci. Os dados completos desses trabalhos selecionados podem ser consultados na lista de referências, ao final deste texto.

APRESENTAÇÃO DE DADOS E RESULTADOS

Analisando o conteúdo abordado em cada um dos referenciais teóricos selecionados na amostragem de pesquisa, percebe-se que, o periódico científico constitui um dos mais importantes canais de comunicação científica (MEADOWS, 2001), o qual tem proporcionado novos rumos às práticas científicas e, por decorrência, aumentado o nível de produção em todas as áreas do conhecimento (BÉGAULT, 2009). Sua gênese está ligada à criação do *Journal des Sçavans*, em 5 de janeiro de 1665, na cidade de Paris, o qual se apresentava como uma publicação de pesquisas relativas a diversas áreas, além de divulgar informações não científicas (ALVES, 2011). No decorrer dos tempos, com a criação dos periódicos especializados e com o aparecimento do periódico eletrônico (OLIVEIRA, 2008), intensificam-se novas formas de divulgar os resultados de pesquisa (MUELLER, 2003), contribuindo, excessivamente para a democratização do conhecimento, conforme discorrido no estudo de Corrêa (2012).

Mesmo desencadeando contribuições valiosas para a ciência, conforme relatado por Dias (2002a), é apenas com o aparecimento dos periódicos científicos eletrônicos que o interesse em se estudar os periódicos se intensificou, como nos ensina Targino (2001). Considerando o período de tempo determinado para esta revisão, revela-se que já nas últimas décadas do século XX, estudos foram realizados no sentido de investigar as características intrínsecas dos periódicos, considerando a

realidade onde estavam inseridos, como discorrido por Foresti (1986), como também realizaram-se projetos de elaboração das coleções de periódicos nas bibliotecas digitais, enfatizando a importância dos consórcios entre bibliotecas (KRZYZANOWSKI; TARUHN, 1998) e a migração dos periódicos do formato impresso para o digital, como apontou Simeão (2001) e Benchimol, Arruda e Silva (2016). Os resultados apontados nos estudos sobre migração são similares no que se refere às vantagens em se adotar o sistema eletrônico, principalmente quando o sistema é sustentado por softwares livres, proporcionando otimização de tempo e custos.

Os estudos são unânimes ao confirmar a importância da migração, fato esse que justifica o aparecimento do periódico misto (impresso e eletrônico) nas primeiras décadas do século XX, e, no decorrer dos anos, o desaparecimento da forma impressa, como reforçado, também, nos estudos de Sayão (2010). Dessa forma, segundo pesquisa de Bertin e Fortaleza (2007), o periódico científico, desde sua criação no século XVII (MEADOWS, 2001; MUELLER, 2003; OLIVEIRA, 2008; BÉGAULT, 2009; ALVES, 2011), insere-se em um progresso rumo à sua legitimação para a disseminação da informação pela comunidade científica. Contudo, mesmo com a incorporação das tecnologias digitais e, por conseguinte, a redefinição dos processos editoriais, a essência do fazer científico presente nos periódicos permaneceu inalterada, como apontado nas pesquisas de Mueller (2003, 2006).

A literatura demonstra a popularização dos periódicos científicos a partir do uso dos recursos digitais. Maior impulso é viabilizado com o desenvolvimento de estudos para garantir o acesso livre a essas publicações o que desencadeia a necessidade constante de métodos avaliativos dos periódicos, sejam eles de natureza quantitativa quanto qualitativa (DIAS, 2002a). Como consequência, assistiu-se, no decorrer das primeiras décadas do século XXI, o interesse em investigar a satisfação dos usuários das plataformas digitais dos periódicos (DIAS, 2002b). Por outro lado, é também dessa popularização, segundo o estudo de Tenopir (2001), que o artigo de periódico tem sido considerado como um dos principais elementos utilizados para elaboração dos trabalhos científicos.

Semelhante aos resultados oriundos da pesquisa de Tenopir (2001), Chinnasamy (2012) e Waete, Moura e Mangue (2013) confirmam o aumento do uso dos periódicos, haja vista promover a atualização do conhecimento por parte dos pesquisadores ao acessarem o conteúdo dos artigos. Segundo o primeiro estudo, a maioria dos sujeitos investigados na pesquisa acessa os periódicos eletrônicos semanalmente (44%). Já no segundo estudo, tanto os professores quanto pesquisadores, utilizam de forma rotineira periódicos eletrônicos em suas atividades de ensino e pesquisa.

O processo de informatização das unidades de informação, certamente, é visto como um fator essencial para estimular o uso das coleções de periódicos eletrônicos, desencadeando esforços por parte de bibliotecários e demais envolvidos com o mercado da informação, na elaboração das bibliotecas digitais (OHIRA et al., 2003; CRUZ, 2003). O estudo de Neves (2004), por exemplo, representa uma brilhante iniciativa nesse sentido, uma vez que constatou-se, no âmbito das bibliotecas, a viabilidade em se construir acervos digitais com acesso livre a texto completo dos artigos de periódicos, seja no que tange às facilidades e rapidez no acesso, quanto à redução de custos envolvidos com a editoração e tratamento dessas publicações.

Sendo assim, mudanças de hábitos e práticas têm permeado o processo de editoração de periódicos quanto os fazeres dos profissionais que lidam com o tratamento e disseminação das publicações. A migração dos formatos impressos dos periódicos para o eletrônico, segundo estudo de Bomfá e Castro (2004), proporcionou vantagens semelhantes aos resultados obtidos no estudo de Neves (2004), tais como rapidez, integração, fluxo da comunicação e redução de custos. Os periódicos em plataformas

digitais despertaram a necessidade de um rigoroso processo de normalização, no entanto, a inexistência de políticas institucionais de periódicos e o não cumprimento aos padrões de normalização da Associação Brasileira de Normas Técnicas (ABNT) representa um entrave, conforme relatado no estudo de Mota e Autran (2005). Essa mesma ocorrência foi identificada na pesquisa desenvolvida por Mendonça, Fachin e Varvakis (2006), a qual presenciou discrepância na normalização dos periódicos na área de Ciência da Informação, tendo em vista as normas internacionais, como a International Organization for Standardization (ISO). Semelhante a esse estudo, Medeiros, Fachin e Rados (2008) constataram a necessidade de se reestruturar e avaliar os modelos de avaliação para periódicos, principalmente no que tange à normalização, uma vez que, na Ciência da Informação, os periódicos se encontram parcialmente normalizados. Com efeito, mesmo com esses desafios, observa-se uma adesão contínua ao processo de editoração eletrônica (CRUZ, 2003; BOMFÁ; CASTRO, 2004; NEVES, 2005) a partir do ano 2000 e preferência de uso por artigos eletrônicos por parte dos usuários (SOUZA; ALBUQUERQUE, 2005; COSTA, 2007).

O movimento de acesso livre ao conhecimento instigou o desenvolvimento de softwares livres para gerenciamento do processo editorial. Portanto, muitos relatos de pesquisa são encontrados na literatura sobre implantação e avaliação desses sistemas, sobretudo o Sistema Eletrônico de Editoração de Revistas (SEER), o qual tem evidenciado vantagens em todo o ciclo editorial das publicações periódicas, como nos relatou Arellano, Santos e Fonseca (2005) e Meirelles (2006). Para Souza (2010), o SEER representou um avanço na editoração eletrônica no Brasil, sobretudo por ser um software de acesso gratuito, como também, por possuir uma interface amigável, dentre outras vantagens. No entanto, o referido estudo menciona a necessidade primordial de se construir uma política nacional de acesso aberto, para, posteriormente, implementar ações, estratégias e instrumentos para garantir a democratização ao conhecimento. Ainda sobre a plataforma SEER, a pesquisa de Lima e Miranda (2011), realizada no contexto do estado do Rio Grande do Sul, acerca de critérios de qualidade dessa plataforma, constatou que, de um modo geral, os periódicos no SEER necessitam de um cuidado maior por parte dos editores e atores que participam desse processo, visando aumentar a visibilidade e a credibilidade dos trabalhos publicados nos periódicos utilizadores do SEER e das Instituições de Ensino Superior brasileiras.

Nesse âmbito, considerando os problemas identificados com o uso da plataforma SEER e as disparidades do acesso livre ao conhecimento, estratégias são pensadas no sentido de garantir a funcionalidade de um sistema de editoração que redefina os papéis dos agentes envolvidos e clarifique as funções para cada um (MUELLER, 2006). Portanto, o estudo de uso e usuários dos acervos eletrônicos, permeado por técnicas de avaliação da usabilidade e dos níveis de satisfação torna-se um assunto recorrente na literatura (MARCONDES; MENDONÇA, 2006; ODDONE; MEIRELLES, 2006). Contudo, a pesquisa de Oddone e Meirelles (2006) identificou, no Portal da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), algumas problemáticas nos estudos de usabilidade, tais como: a falta de padronização dos termos técnicos e dos indicadores de desempenho para aferir a usabilidade e satisfação dos usuários do referido portal.

Em linhas gerais, a importância atribuída aos estudos do uso e de usuários, sobretudo no que se refere à estruturação dos portais de periódicos, deve-se ao fato desses estudos viabilizarem a identificação de problemas e permitirem a realização de modificações na ergonomia dos produtos e serviços oferecidos nos ambientes digitais (ROZADOS; ALVAREZ, 2013). Com base nos resultados desses estudos, é possível realizar a avaliação do que é oferecido, o que acarreta o aumento da qualidade (facilidade de uso) na interação humano-computador (ALVAREZ; ROZADOS, 2015).

Além das investigações voltadas ao fator humano (uso e satisfação), os sistemas gerenciais viabilizaram outras variáveis a serem analisadas e avaliadas. Encontram-se diversos estudos na literatura acerca do levantamento das características intrínsecas dos periódicos, como também, o nível de desempenho dos sistemas informatizados. Assim, segundo Pinto, Rodrigues Barquin e Moreira-González (2006), a análise de citação, por exemplo, pode ser um instrumento eficaz na aferição de qualidade e impacto dos periódicos. Corrobora com esse pensamento, as propostas de Mattos e Dias (2010), para quem, a análise de citações permite a identificação ou não do periódico em portais ou em outros locais de indexação e estoques. Da mesma forma que o uso das citações, outro fator a ser investigado nas avaliações, diz respeito ao Qualis/CAPES, o qual pode consolidar-se como um fator para qualificar o foco, objetivos e práticas editoriais dos periódicos, como nos relata Jacon (2007).

O estudo de Job e Ferreira (2012) também confirma a importância e necessidade da avaliação dos periódicos, sendo que o processo editorial eletrônico contribui para otimizar os procedimentos avaliativos, como também consolida a formalização de diversos instrumentos, métodos e técnicas de avaliação, sobretudo no que tange a indicadores de desempenho. Como exemplo dessas estratégias para avaliação, Vilan Filho, Arruda e Perucchi (2012) mencionam a participação da bibliometria, da cientometria, além da análise de citações, as quais contribuem efetivamente na identificação das citações de um campo de conhecimento no âmbito de outros campos ou em áreas científicas amplas ou específicas.

No âmbito da editoração de periódicos eletrônicos e da disponibilização dessas publicações no ambiente digital, as bibliotecas estabelecem novos métodos de tratamento e gestão das coleções (FACHIN; HILLESHEIM, 2006), consolidando, nesse âmbito, o surgimento de bases de dados, como a Scielo e de bibliotecas digitais, como o Portal CAPES. Porém, ainda existem resistência quanto à aceitação do periódico eletrônico e desses instrumentos de controle da produção científica por alguns usuários, sobretudo no que tange aos problemas relativos à preservação e arquivamento da informação, garantia de acesso ao longo do tempo, desenvolvimento de interfaces confiáveis e disponibilização de coleções retrospectivas (OLIVEIRA, 2007).

As bibliotecas em comunhão com cientistas de diversas áreas devem ocupar lugar especial na gestão dos periódicos eletrônicos, atuando tanto na organização, disseminação quanto na avaliação dessas publicações. Contudo, além dessas funções, é aceitável o estabelecimento de parâmetros para medir a produção científica. Diversos estudos encontrados na amostra desta pesquisa evidenciam a prática de mapeamento e/ou levantamento dos periódicos eletrônicos (VALÉRIO; PINHEIRO, 2008; MARTINS; BRAILLE, 2009), como também o uso de técnicas especializadas, como as análises estatísticas e bibliométricas (COSTA, GUIMARÃES, 2010; FERREIRA, 2010). Em linhas gerais, essas estratégias, conforme refletido por Valério e Pinheiro (2008), têm como objetivo principal assegurar a divulgação e comunicação do conhecimento científico, em processos de compartilhamento de interesses.

Com o surgimento dos periódicos científicos eletrônicos, observam-se vantagens e melhoria na produção científica. Todavia, não resta dúvida, em meio a essas melhorias, o surgimento de diversos desafios, os quais requerem estudos mais aprofundados, a fim de serem solucionados. Dentre alguns desses desafios considerados nas obras de Café e Fachin (2007) e Sayão (2010), cita-se a preservação digital. Esses autores sugerem como solução desse problema, a elaboração de repositórios de dados científicos, além do estabelecimento - de acordo com resultados do segundo estudo - de políticas específicas para nortear os trabalhos das bibliotecas, dos editores e demais envolvidos; todavia, na conclusão do primeiro estudo mencionado, foi constatada a necessidade de uma estrutura material e

humana mais estável para garantir a gestão dos periódicos eletrônicos e o acesso livre ao conhecimento científico.

Além dos problemas supracitados, o estudo de Mueller (2006) identifica a questão do acesso livre, o qual reconfigura o fluxo da comunicação científica e redefine o papel desempenhado por muitas editoras. Mesmo constituindo um assunto controverso, o acesso livre foi bem aceito no âmbito da produção científica, a partir do uso das bibliotecas digitais, as quais promovem a satisfação do usuário, por meio dos estudos de usabilidade, haja vista adequar-se à realidade e necessidades específicas (COSTA; RAMALHO, 2010; CUNHA; CENDÓN, 2010).

Nos últimos anos, as bibliotecas digitais têm assumido diversos papéis no âmbito da comunicação e divulgação da produção científica, como também têm se manifestado com terminologias diferenciadas. Observa-se a elaboração e o uso, cada vez mais frequente, dos portais de periódicos, também denominados de estoques de periódicos científicos (COSTA, 2007), os quais se configuram como indicadores de excelência das instituições mantenedoras, em função do possível aumento do número de acessos (GARRIDO; RODRIGUES, 2010). Em concordância com essas autoras, o estudo de Ferreira e Cunha (2008) também descreve a visibilidade dos periódicos como maior característica dos portais. Tal estudo confirmou-se que o fato das revistas participarem de um portal único, que congrega várias outras revistas da mesma área, potencializa sua visibilidade e acessibilidade às diversas publicações da referida área.

Os benefícios da visibilidade de um periódico em portais são discutidos no estudo de Bentancourt e Rocha (2012), para quem os metadados das publicações, ao serem bem tratados, viabilizam melhorias na visibilidade do objeto digital. Também discutem essa mesma questão, Blattmann e Santos (2013), ao garantirem as vantagens dos portais de periódicos, os quais possibilitam reduzir os custos, resgatam a memória cultural, facilitam a manutenção e a segurança do sistema operacional e ampliam a visibilidade das revistas hospedadas, e, com o acesso aberto, a tendência é aumentar o acesso e uso da informação publicada. Além dos portais condicionarem a visibilidade, o estudo de Shintaku, Brito e Carvalho Neto (2015) destaca o potencial das tecnologias, ao realizarem o cruzamento dos dados e permitir a visibilidade do periódico em diferentes locais a que esteja armazenado. Dessa forma, esses autores investigam os critérios para inserção dos periódicos em dois importantes indexadores, a saber: o Latindex e SciELO e concluem que o primeiro indexador apresenta-se como mais flexível, haja vista a quantidade maior de revistas em seu catálogo.

Além dos assuntos relacionados ao crescimento dos periódicos científicos e eletrônicos, como também as tecnologias utilizadas no processo editorial, os métodos de avaliação, as estratégias de busca e recuperação da informação - visando o uso efetivo desses meios de comunicação, por conseguinte, a satisfação dos usuários - importante enfatizar outras questões abordadas na literatura, tais como: o trabalho multiprofissional, o planejamento e gestão dos periódicos e portais, o trabalho em equipe, os gastos envolvidos, os modelos teórico-metodológicos, e, por fim, as políticas públicas para apoiar as ações desempenhadas na construção desses estoques de informação, como abordado nas pesquisas de Fachin e Hillesheim (2006), Café e Fachin (2007), Garrido e Rodrigues (2010) e Sayão (2010).

Evidencia-se que o sucesso de qualquer empreendimento no âmbito da criação, gestão e organização dos periódicos eletrônicos depende da somatória de conhecimento entre diferentes profissionais envolvidos com a comunicação científica, sobretudo aqueles encarregados, por excelência, com as técnicas de tratamento dessas publicações, como os bibliotecários (FACHIN; HILLESHEIM, 2006). Todavia, conforme descoberto por Santana e Francelin (2016), as potencialidades desses profissionais

extravasam os fazeres apenas técnicos – normalmente realizados nas bibliotecas junto aos processos de organização da informação - mas eles também podem atuar em conjunto com as equipes de periódicos científicos, de modo a contribuir com o fluxo do processo editorial nas plataformas dos periódicos.

CONCLUSÕES

A partir dos artigos analisados nesta pesquisa, constata-se, em linhas gerais, o crescimento do número de pesquisa ao longo dos anos, além de diversificada utilização de métodos de investigação. Também se percebe a presença de diversos assuntos pertencentes à temática periódicos científicos, sobretudo no que se refere aos periódicos no âmbito das novas tecnologias (periódicos eletrônicos), em que são apresentados estudos sobre vantagens e desafios desse novo formato de publicação periódica. De modo geral, evidencia-se a relevância da temática periódicos científicos, no âmago da literatura em Ciência da Informação publicada no Brasil.

REFERÊNCIAS BIBLIOGRÁFICAS

Autran, M. M. M. *et al.* (2016). Revisão sistemática: desvelando a gestão do conhecimento nos anais do Enancib. *Biblionline*, 12(2), 84-100. Recuperado em 20 março, 2017, de <http://periodicos.ufpb.br/ojs2/index.php/biblio/article/view/28818/15518>.

Alvarez, G. R. N., & Rozados, H. B. F. (2015). Usabilidade da seção de submissão do SEER: o caso da revista Em Questão. *Informação & Informação*, 20(1), 92-109. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/17762>.

Alves, L. C. (2011). Informação e os sistemas de comunicação científica na ciência da informação. *DataGramaZero*, 12(3), 1-13. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/10229>.

Arellano, M. N. M. R., Santos, R., & Fonseca, R. N. (2005). SEER: disseminação de um sistema eletrônico para editoração de revistas científicas no Brasil. *Arquivística.net*, 1(2), 75-82. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/3949>.

Bégault, B. A. (2009). The scientific journal, a role for mediation of information between researchers: what future in the digital environment? *Revista Eletrônica de Comunicação, Informação & Inovação em Saúde*, 3(3), 91-96. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/18215>.

Benchimol, A., Arruda, M. I. M. & Silva, T. C. (2016). Boletim do Museu Paraense Emílio Goeldi: do impresso ao eletrônico. *Informação e Sociedade: Estudos*, 26(3), 81-93. Recuperado em 20 março, 2017, de <http://www.ies.ufpb.br/ojs/index.php/ies/article/view/31121/16996>.

Bertin, P. C. R., Fortaleza, J. M. & suhet, A. R. (2007). Paradigma atual da comunicação científica e introdução da Revista Pesquisa Agropecuária Brasileira (PAB) no canal eletrônico. *Perspectivas em Ciência da Informação*, 12(3), 83-95. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/4799>.

Bentancourt, S. M. P., & rocha, R. P. (2012). Metadados de qualidade e visibilidade na comunicação científica. *Encontros Bibli*, 17(n. espec.), 82-101. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/12636>.

Blattmann, U., & Santos, R. N. M. (2014). Revistas científicas brasileiras e sua visibilidade no acesso aberto. *Informação & Sociedade: Estudos*, 24(3), 99-106. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/18410>.

Bomfá, C. R. Z. & Castro, J. O. E. E. (2004). Desenvolvimento de revistas científicas em mídia digital: o caso da revista produção online. *Ciência da Informação*, 32(2), 39-48. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/1148>.

Brapci (2017). *Sobre a Brapci*. Recuperado em 20 março, 2017, de www.brapci.ufpr.br/index.php.

Cafe, L. G. M. A., & Fachin, G. R. B. R. (2007). Data providers, service providers and scientific journals available in information science and related areas. *Encontros Bibli*, núm. Espec., 1-18. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/16490>.

Cendón, B. V., Souza, J. L., & Ribeiro, N. (2011). A. Satisfação dos usuários do Portal de Periódicos da Capes: um estudo sobre a obtenção de sucesso no uso do sistema. *Perspectivas em Ciência da Informação*, 16(2), 67-100. Recuperado em 20 março, 2017, de <http://www.scielo.br/pdf/pci/v16n2/06.pdf>.

Chinnasamy, B. (2012). Acesso a periódicos eletrônicos através do consórcio Ugc-Infonet por docentes e acadêmicos da Alagappa University, Índia. *Brazilian Journal of Information Science*, 6(1), 39-53. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/13192>.

Corrêa, E. C. D. (2012). Os artigos científicos em tempos de web 2.0: uma reflexão teórica. *Revista ACB*, 17(1), 42-58. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/11756>.

Costa, R. O. (2007). A comunicação eletrônica e a alteração de tempo e espaço na produção do conhecimento científico. *Ciência da Informação*, 36(2), 7-15. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/5714>.

Costa, S. M. S., & Guimarães, L. V. S. Qualidade de periódicos científicos eletrônicos brasileiros que utilizam o sistema eletrônico de editoração de revistas (SEER). *Informação & Informação*, 15(n. espec.), 75-92. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/9524>.

Costa, L. F., & Ramalho, F. A. (2010). A "cosmopédia" da CAPES: contexto de uso do portal de periódicos. *Tendências da Pesquisa Brasileira em Ciência da Informação*, 3(1), 138-154, Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/9331>.

Cruz, A. A. A. C. et al. (2003). Impacto dos periódicos eletrônicos em bibliotecas universitárias. *Ciência da Informação*, 32(2), 47-53. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/1101>.

Cunha, A. U. L., & Cendón, B. V. (2010). Uso de bibliotecas digitais de periódicos: um estudo comparativo do uso do portal de periódicos CAPES entre áreas do conhecimento. *Perspectivas em Ciência da Informação*, 15(1), 70-91. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/8780>.

Dias, G. A. (2002a). Avaliação do acesso a periódicos eletrônicos na web pela análise do arquivo de log de acesso. *Ciência da Informação*, 31(1), 7-12. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/5793>.

Dias, G. A. (2002b). Periódicos eletrônicos: considerações relativas à aceitação deste recurso pelos usuários. *Ciência da Informação*, 31(3), 18-25. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/1096>.

Fachin, G. R. B., & Hillesheim, A. I. A. (2006). *Periódico científico: padronização e organização*. Florianópolis: Editora da UFSC.

Ferreira, S. M. S. P., & Cunha, A. S. (2008). Portal revcom & google analytics: acessando a caixa-preta da informação. *Em Questão*, 14(1), 41-61. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/6767>.

Ferreira, A. G. C. (2010). Bibliometria na avaliação de periódicos científicos. *DataGramaZero*, 11(3), 5. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/8927>.

Foresti, N. R. A. B. (1986). A revista ciência da informação no contexto de sua instituição: algumas considerações. *Ciência da Informação*, 15(2), 143-150. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/2216>.

Galvão, T. F., & Pereira, M. G. (2014). Revisões sistemáticas da literatura: passos para sua elaboração. *Epidemiologia Serviço Saúde*, 23(1), 183-184. Recuperado em 20 março, 2017, de <http://www.scielo.br/pdf/ress/v23n1/2237-9622-ress-23-01-00183.pdf>.

Garrido, I. S., & Rodrigues, R. S. (2010). Portais de periódicos científicos online: organização institucional das publicações. *Perspectivas em Ciência da Informação*, 15(2), 56-72. Recuperado em 20 março, 2017, de <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/943/732>.

Jacon, M. C. M. (2007). Base qualis e a indução do uso de periódicos da área de psicologia. *Transinformação*, 19(2), 189-197. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/467>.

Job, I., Mattos, A. M. & Ferreira, A. G. C. (2012). They are on the podium: analysis of brazilian journals in physical education and sports. *Em Questão*, 18(3), 161-178. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/11920>.

Krzyzanowski, R. F., & Taruhn, R. (1998). Biblioteca eletrônica de revistas científicas internacionais: projeto de consórcio. *Ciência da Informação*, 27(2), 193-197. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/3070>.

Lima, J. L. B., & Miranda, A. L. C. O. D. (2011). Periódicos eletrônicos com o uso do SEER no Rio Grande do Sul: critérios de qualidade. *Perspectivas em Gestão & Conhecimento*, 1(1), 163-179. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/10255>.

Marcondes, C. H., & Mendonça, M. L. A. R. (2006). Avaliação de periódicos eletrônicos acadêmicos brasileiros: uma proposta de método baseado na análise de links para o site do periódico. *Transinformação*, 18(2), 123-130. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/437>.

- Martins, C. U. A., & Braile, D. M. (2009). Análise cientométrica dos periódicos em ciência da saúde e áreas correlatas disponíveis no portal de periódicos da CAPES. *Perspectivas em Ciência da Informação*, 14(3), 75-93. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/7939>.
- Mattos, A. M., & Dias, E. W. (2010). Periódicos eletrônicos sobre administração disponíveis no portal de periódicos da coordenação de aperfeiçoamento de pessoal de nível superior: uma avaliação. *Ciência da Informação*, 39(1), 51-66. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/9448>.
- Medeiros, G. M., Fachin, G. R. B. R., & Rados, G. R. J. V. (2008). Padronização de periódicos científicos on-line da área de biblioteconomia e ciência da informação: adequação às normas ISO. *Revista ACB*, 13(2), 415-438. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/7996>.
- Meirelles, R. F. A. (2006). O sistema eletrônico de editoração de revistas (SEER) e sua adoção em periódicos brasileiros. *Arquivística.net*, 2(2), 197-207. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/4453>.
- Meadows, J. (2001). Os periódicos científicos e a transição do meio impresso para o eletrônico. *Revista de Biblioteconomia de Brasília*, 25(1), 5-14. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/635>.
- Mendonça, T. C., Fachin, G. R. B. R., & Varvakis, G. R. (2006). Padronização de periódicos científicos on-line: estudo aplicado na área de biblioteconomia e ciência da informação. *Informação & Sociedade: Estudos*, 16(1), 179-191. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/4172>.
- Mota, A. R. S., & Autran, M. M. M. (2005). Periódicos científicos: a experiência nos programas de pós-graduação da Universidade Federal da Paraíba – Campus I. *Biblionline*, 1(1), 1-17. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/9278>.
- Mueller, S. P. M. (2003). O periódico científico. In: CAMPELLO, B. S.; CENDÓN, B. V.; KREMER, J. M. (Org.). *Fontes de informação para pesquisadores e profissionais*. Belo Horizonte: Editora UFMG.
- Mueller, S. P. M. (2006). A comunicação científica e o movimento de acesso livre ao conhecimento. *Ciência da Informação*, 35(2), 27-38. Recuperado em 20 março, 2017, de <http://www.scielo.br/pdf/ci/v35n2/a04v35n2.pdf>.
- Mueller, S. P. M. (2007). *Métodos para a pesquisa em Ciência da Informação*. Brasília: Thesaurus.
- Neves, T. M. G. (2004). Livre acesso à publicação acadêmica. *Ciência da Informação*, 33(3), 116-121. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/3049>.
- Oddone, N. E., & Meirelles, R. F. A. (2006). O portal de periódicos da Capes e os indicadores de desempenho da informação eletrônica. *DataGramaZero*, 7(3), 1-8. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/3730>.
- Ohira, M. L. B. et al (2003). Análise dos periódicos eletrônicos (full text) em ciência da informação: América Latina, Caribe, Portugal e Espanha. *Informação & Informação*, 8(1), 1-24. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/4345>.

Oliveira, R. B. P. M. (2007). Uso de periódicos científicos eletrônicos por docentes e pós-graduandos do Instituto de Geociências da USP. *Ciência da Informação*, 36(3), 59-66. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/5137>.

Oliveira, R. B. P. M. (2008). Scientific electronic journals: definitions and history. *Informação & Sociedade: Estudos*, 18(2), 69-77. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/16664>.

Pinto, A. L., Barquín, B. A. R. G., & Moreira-González, J. A. Análisis de citación de la revista ciência da informação del IBICT. *Ciência da Informação*, 35(3), 153-165. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/4506>.

Rother, E. T. (2007). Revisão sistemática x Revisão Narrativa. *Acta Paulista Enfermagem*, 20(2), 1-2. Recuperado em 20 março, 2017, de <http://www.scielo.br/pdf/ape/v20n2/a01v20n2.pdf>.

Rozados, H. B. F., & Alvarez, G. R. N. (2013). Electronic systems for Scientific Journals Publishing: the issue of usability. *Revista ACB*, 18(2), 860-883. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/13581>.

Santana, S. A., & Francelin, M. M. (2016). O bibliotecário e a editoração de periódicos científicos. *Revista Brasileira de Biblioteconomia e Documentação*, 12 (1), 2-24. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/20851>.

Santa Anna, J. (2016). Formação e desenvolvimento de coleções em acervos bibliográficos: o estado da arte no Brasil. In: ENCONTRO DA ASSOCIAÇÃO DE EDUCAÇÃO E INVESTIGAÇÃO EM CIÊNCIA DA INFORMAÇÃO NA IBEROAMÉRICA E NO CARIBE, 10., 2016, Belo Horizonte, Escola de Ciência da Informação, Universidade Federal de Minas Gerais. Recuperado em 20 março, 2017, de <http://edicic2016.eci.ufmg.br/wp-content/uploads/2016/09/EDICIC-2016-PROGRAMA%C3%87%C3%83O-GERAL1509.pdf>.

Sayão, L. F. (2010). Repositórios digitais confiáveis para a preservação de periódicos eletrônicos científicos. *PontodeAcesso*, 4(3), 68-94. Recuperado em 20 março, 2017, de <https://portalseer.ufba.br/index.php/revistaici/article/view/4709/3565>.

Shintaku, M., Brito, R. F., & Carvalho Neto, S. L. (2014). A avaliação dos portais de revistas brasileiros implementados com o SEER/OJS por meio do levantamento da indexação pelo Latindex e Scielo. *Informação & Sociedade: Estudos*, 24(2), 139-148. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/18392>.

Simeão, E. L. M. S. (2001). A experiência da revista de biblioteconomia de Brasília na internet. *Revista de Biblioteconomia de Brasília*, 25(1), 127-140. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/8511>.

Souza, J. L. A. (2010). Revistas eletrônicas com uso de software livre. *DataGramaZero*, 11(4), 1-9. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/9008>.

Souza, T. E. R., & Albuquerque, M. E. B. C. (2005). Periódicos científicos em biblioteconomia e ciência da informação: consulta por alunos concluintes do curso de biblioteconomia da UFPB. *Biblionline*, 1(2), 1-23. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/9268>.

Targino, M. G. (2001). Comunicação científica na sociedade tecnológica: Periódicos eletrônicos em discussão. *Comunicação e Sociedade*, 3(1), 1-2.

Tenopir, C., & King, D. (2001). A importância dos periódicos para o trabalho científico. *Revista de Biblioteconomia de Brasília*, 25(1), 15-26. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/1169>.

Valentim, M. L. P. (2005). *Métodos qualitativos de pesquisa em ciência da informação*. São Paulo: Polis, 2005.

Valerio, P. M., & Pinheiro, L. N. V. R. (2008). Da comunicação científica à divulgação. *Transinformação*, 20(2), 159-169, 2008. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/486>.

Vilan Filho, J. L., Arruda, R. Z. V., & Perucchi, V. (2012). Análise das citações aos periódicos científicos brasileiros das áreas de informação. *Em Questão*, 18(3), 115-127. Recuperado em 20 março, de <http://www.brapci.ufpr.br/brapci/v/a/11919>.

Waete, R. Z., Moura, M. A., & Mangué, M. V. (2013). Use of electronic journals in institutions of higher education and research centers in mozambique. *InCID*, 4(2), 165-184. Recuperado em 20 março, 2017, de <http://www.brapci.ufpr.br/brapci/v/a/16827>



ONTOLOGÍA EN GESTIÓN DE PROYECTOS. UN ENFOQUE PRÁCTICO

Anisleiby Fernández Hernández¹, María José López Huertas²

¹Universidad de las Ciencias Informáticas, ani@uci.cu

²Universidad de Granada, mjlopez@ugr.es

RESUMEN La Gestión de Proyectos que se realiza en algunas organizaciones no se realiza de manera efectiva. Existe falta de conocimiento en el control y seguimiento de proyectos, expresadas en carencias en las áreas de conocimiento e insuficiencias para el tratamiento de la incertidumbre de los datos y la ambigüedad en los conceptos. Esto trae consigo duplicidad y heterogeneidad de la información, por la falta de integración de la misma, además de las diferencias en los conceptos utilizados para el intercambio de información, causando incongruencias en los resultados, falta de estandarización de estructuras, calidad, consistencia y disponibilidad en los datos. En este sentido las Ontologías juegan un rol fundamental, al proveer un vocabulario común y sin ambigüedades para referirse a los términos en el área aplicada, pudiéndose compartir o reutilizar éstos entre diferentes aplicaciones que hagan uso de la Ontología. En esta investigación se propone la creación de una Ontología en Gestión de Proyectos desde un enfoque práctico. Se utiliza como metodología para el diseño y construcción de la ontología Methontology, la cual permitió que el proceso de desarrollo fuera organizado y evolutivo. Como herramienta para desarrollar la ontología se utiliza el Protégé 4.3, permitiendo construir ontologías completamente compatibles con el lenguaje escogido: OWL; además cuenta con un entorno abierto y fácil de entender.

PALABRAS CLAVE *Ontologías, Gestión de Proyectos, Dominio del Conocimiento, Gestión del Conocimiento.*

ABSTRACT The Project Management that is done in some organizations is not done in an effective way. There is a lack of knowledge in the control and monitoring of projects, expressed in shortcomings in the areas of knowledge and insufficiencies for the treatment of data uncertainty and ambiguity in concepts. This leads to duplication and heterogeneity of the information, due to the lack of integration of the information, in addition to the differences in the concepts used for the exchange of information, causing inconsistencies in results, lack of standardization of structures, quality, consistency and availability in the data. In this sense the Ontologies play a fundamental role, by providing a common and unambiguous vocabulary to refer to the terms in the applied area, being able to share or reuse these between different applications that make use of the Ontology. This research proposes the creation of an Ontology in Project Management from a practical perspective. It is used as a methodology for the design and construction of the ontology Methontology, which allowed the development process to be organized and evolutionary. As a tool to develop the ontology, Protégé 4.3 is used, allowing the construction of ontologies completely compatible with the chosen language: OWL ; also has an open and easy to understand environment.

KEYWORDS *Ontologies, Project Management, Knowledge Domain, Knowledge Management.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUCCIÓN

En los últimos años se ha venido experimentando un acelerado avance de las tecnologías, donde el incremento de la información en la red de redes y el gran cúmulo de datos en un área del conocimiento es tan amplio, así como difícil su acceso. Por tanto, a pesar de que la web actual, ha alcanzado un gran éxito, presenta como principales problemas la sobrecarga y heterogeneidad de la información existentes en ella.

La Web Semántica, por extensión la Web de los metadatos en este sentido cobra vital importancia, siendo la respuesta ideal para lograr un ciberespacio más accesible y utilizable a nivel mundial, permitiendo que la información se encuentre más estructurada y favoreciendo el procesamiento estructurado de la misma y por ende de fácil acceso para los usuarios que acceden a ella.

Las Ontologías están ocupando un lugar importante en la Ingeniería del Conocimiento. Muchas soluciones, en distintas áreas, aprovechan las bondades que éstas brindan para organizar el conocimiento en dominios específicos.

En la actualidad muchas de las aplicaciones relacionadas con el Comercio Electrónico, Bioinformática, Recuperación de la Información, Educación, Gestión del Conocimiento, Procesamiento en Lenguaje Natural y en la Web Semántica están soportadas en Ontologías.

De manera más específica, las ontologías son el estudio de las categorías de las cosas que existen o podrían existir en cierto dominio. En la última década, este término ha ganado relevancia entre los Ingenieros de Conocimiento tomando una interpretación particular; y es por ello que, en 1995, Guarino y Giarretta proponen utilizar la palabra “Ontología” (con O mayúscula) para referirse a ella en el contexto de la Ingeniería del Conocimiento.

Es común que cada comunidad que desarrolla Ontologías adopte una definición propia dependiendo de sus necesidades. Entre las tantas definiciones que se pueden encontrar, la más aceptada es la propuesta por Gruber en 1993: *“una Ontología es una especificación formal y explícita de una conceptualización compartida”* (p13). Los términos utilizados en esta definición se basan principalmente en lo siguiente:

- Conceptualización: Modelo abstracto de un fenómeno, que puede ser visto como un conjunto de reglas informales que restringen su estructura. Por lo general se expresa como un conjunto de conceptos (entidades, atributos, procesos), sus definiciones e interrelaciones.
- Formal: Organización teórica de términos y relaciones usados como herramienta para el análisis de los conceptos de un dominio.
- Compartida: Se refiere a la captura del conocimiento consensual que es aceptado por una comunidad.
- Explícita: Concierne a la especificación de los conceptos y a las restricciones sobre éstos Uschold y Gruninger.

Las Ontologías exhiben características especiales para la representación del conocimiento y el procesamiento de éste en sistemas inteligentes. Según Chandrasekaran (1999); Gruber (1993); Guarino (1995); McGuinness (1995); Schreiber (2000) se destacan las siguientes:

- Las Ontologías proveen un vocabulario común y sin ambigüedades para referirse a los términos en el área aplicada, pudiéndose compartir o reutilizar éstos entre diferentes aplicaciones que hagan uso de la Ontología.
- Además de un vocabulario común, especifican una taxonomía o herencia de conceptos que establecen una categorización o clasificación de las entidades del dominio. Una buena taxonomía es simple y fácil de recordar, separa sus entidades de forma mutuamente excluyente, y define grupos y subgrupos sin ambigüedad.
- El vocabulario y la taxonomía representan un marco de trabajo conceptual para el análisis, discusión o consulta de información de un dominio.
- Una Ontología incluye una completa generalización/especificación de sus clases y subclases, las cuales están formalmente especificadas (incluyendo sus relaciones e instancias) asegurando la consistencia en los procesos deductivos.
- Las Ontologías son implementadas en un lenguaje específico de representación Ontológica (ontology representation languages) de manera que la especificación de sus clases, relaciones entre éstas y sus restricciones dependerán de las características de dicho lenguaje.

Por tanto, el dominio representado en una ontología, como sistemas basados en conocimiento, es un lenguaje controlado y un conocimiento que puede utilizarse como una forma de preservación del mismo por parte de los expertos, o no, en cualquier campo de aplicación, permitiendo una gestión rápida, eficaz de la información para su recuperación, uso y reutilización para la toma de decisiones. Siendo los sistemas basados en el conocimiento una de las técnicas de Inteligencia Artificial más empleadas en la actualidad en el proceso de toma de decisiones.

Una organización inteligente fundamenta el éxito en el logro de sus objetivos, en la definición de estrategias para conservar y desarrollar su conocimiento, pero inevitablemente existen muchas organizaciones que tienen diversos problemas para gestionarlo de manera adecuada.

En este caso en muchas organizaciones las insuficiencias en la Gestión de Proyectos es causa de innumerables pérdidas económicas y atrasos en proyectos con elevado impacto social. Proyectos con dificultades en la planificación, el control y su seguimiento son evidencias de ello. Asociado además a la Gestión de los Proyectos de inversión se identifican problemas de seguridad que pueden comprometer el desarrollo y la seguridad, tanto de empresas como de estados y gobiernos (Piñero, 2013).

Entre las principales causas de estos efectos está la falta de conocimiento en el control y seguimiento de proyectos, así como las debilidades de las herramientas para la toma de decisiones en la Gestión de Proyectos, expresadas en carencias en las áreas de conocimiento e insuficiencias para el tratamiento de la incertidumbre de los datos y la ambigüedad en los conceptos. Lo cual trae consigo duplicidad y heterogeneidad de la información, por la falta de integración de la misma, además de las diferencias en los conceptos utilizados para el intercambio de información causando incongruencias en los resultados, falta de estandarización de estructuras, calidad, consistencia y disponibilidad en los datos.

DOMINIO GESTIÓN DE PROYECTOS

Kerzner define la Gestión de Proyectos como: *"the planning, organizing, directing, and controlling of company resources for a relatively short-term objective that has been established to complete specific goals and objectives"* (p20).

Kirsch la considera como: *"Application of formal and informal techniques, tools, methods, and heuristics [...], which are used by the project manager to motivate and guide a team to carry out a project within a given set of constraints"* (p18).

El Project Management Institute se refiere a la Gestión de Proyectos como: *"la aplicación de conocimiento, habilidades, herramientas y técnicas a las actividades de un proyecto para satisfacer los requisitos del proyecto. Se aplican e integran los procesos de dirección de proyectos de inicio, planificación, ejecución, seguimiento y control y cierre. En este proceso el director del proyecto es la persona responsable de alcanzar los objetivos del proyecto"* (p20).

La Gestión de Proyectos tiene como finalidad principal la planificación, el seguimiento y control de las actividades y de los recursos humanos y materiales que intervienen en el desarrollo de un Sistema de Información. Como consecuencia de este control es posible conocer en todo momento qué problemas se producen y resolverlos o paliarlos de manera inmediata.

Cada proyecto tiene un comienzo definido y un final definido. El final se alcanza cuando se han logrado los objetivos del proyecto o cuando queda claro que los objetivos del proyecto no serán o no podrán ser alcanzados, o cuando la necesidad del proyecto ya no exista y el proyecto sea cancelado. En cada caso, sin embargo, la duración de un proyecto es limitada.

Los proyectos se desarrollan en cada organización con un fin único: alcanzar beneficios.

En cuanto a la tipología de proyectos que existen, cada autor se refiere a una clasificación en dependencia de un parámetro concreto por el cual se rige. Muchos lo definen de diferentes maneras.

No existe una única manera, que sea la mejor, para definir el ciclo de vida ideal de un proyecto. Algunas organizaciones han establecido políticas que estandarizan todos los proyectos con un ciclo de vida único, mientras que otras permiten al equipo de dirección del proyecto elegir el ciclo de vida más apropiado para el proyecto del equipo. Asimismo, las prácticas comunes de la industria a menudo conducen a usar un ciclo de vida preferido dentro de dicha industria.

Entre las principales áreas de conocimiento dentro de la Gestión de Proyectos el Project Management Institute (2009) señala 9 grandes áreas:

- 1) **Gestión de la Integración del Proyecto:** Incluye los procesos y actividades necesarios para identificar, definir, combinar, unificar y coordinar los distintos procesos y actividades de dirección de proyectos dentro de los Grupos de Procesos de Dirección de Proyectos.
- 2) **Gestión del Alcance del Proyecto:** Incluye los procesos necesarios para asegurarse que el proyecto incluya todo el trabajo requerido, y sólo el trabajo requerido, para completar el proyecto satisfactoriamente. La gestión del alcance del proyecto se relaciona principalmente con la definición y el control de lo que está y no está incluido en el proyecto.

- 3) **Gestión del Tiempo del Proyecto:** incluye los procesos necesarios para lograr la conclusión del proyecto a tiempo.
- 4) **Gestión de los Costes del Proyecto:** Incluye los procesos involucrados en la planificación, estimación, preparación del presupuesto y control de costes de forma que el proyecto se pueda completar dentro del presupuesto aprobado.
- 5) **Gestión de la Calidad del Proyecto:** Incluyen todas las actividades de la organización ejecutante que determinan las políticas, los objetivos y las responsabilidades relativas a la calidad de modo que el proyecto satisfaga las necesidades por las cuales se emprendió.
- 6) **Gestión de los Recursos Humanos del Proyecto:** Incluye los procesos que organizan y dirigen el equipo del proyecto. El equipo del proyecto está compuesto por las personas a quienes se les han asignado roles y responsabilidades para concluir el proyecto. Si bien es común hablar de asignación de roles y responsabilidades, los miembros del equipo deberían participar en gran parte de la planificación y toma de decisiones del proyecto. La participación temprana de los miembros del equipo aporta experiencia durante el proceso de planificación y fortalece el compromiso con el proyecto.
- 7) **Gestión de las Comunicaciones del Proyecto:** Incluye los procesos que organizan y dirigen el equipo del proyecto. El equipo del proyecto está compuesto por las personas a quienes se les han asignado roles y responsabilidades para concluir el proyecto. Si bien es común hablar de asignación de roles y responsabilidades, los miembros del equipo deberían participar en gran parte de la planificación y toma de decisiones del proyecto. La participación temprana de los miembros del equipo aporta experiencia durante el proceso de planificación y fortalece el compromiso con el proyecto.
- 8) **Gestión de los Riesgos del Proyecto:** Incluye los procesos relacionados con la planificación de la gestión de riesgos, la identificación y el análisis de riesgos, las respuestas a los riesgos, y el seguimiento y control de riesgos de un proyecto; la mayoría de estos procesos se actualizan durante el proyecto. Los objetivos de la Gestión de los Riesgos del Proyecto son aumentar la probabilidad y el impacto de los eventos positivos, y disminuir la probabilidad y el impacto de los eventos adversos para el proyecto.
- 9) **Gestión de las Adquisiciones del Proyecto:** Incluye los procesos para comprar o adquirir los productos, servicios o resultados necesarios fuera del equipo del proyecto para realizar el trabajo.

METODOLOGÍA

Las herramientas de desarrollo de ontologías, han mejorado enormemente desde la creación de los primeros entornos. Se pueden distinguir dos grupos de herramientas, teniendo en cuenta su evolución desde que aparecieron a mediados de la década de 1990, Farquhar (1997):

- Herramientas cuyo modelo de conocimientos se corresponde directamente con el de un lenguaje de ontologías.
- Plataformas integradas. Cuya característica principal es que tienen una arquitectura flexible y extensible, y cuyo modelo de conocimientos es normalmente independiente de lenguajes de ontologías existentes.

A consideración de los autores en esta investigación se usa el Protegé, el cual es una herramienta a través de la cual el usuario puede construir ontologías de dominio, generar usuarios de entrada de datos y efectuar la propia entrada de datos. Aplicación amigable de modelado de conocimiento. Es una herramienta que permite acceso a aplicaciones externas basadas en conocimiento. Además es una biblioteca a la que otras aplicaciones pueden acceder, fundamentalmente a las bases de conocimiento de las cuales se dispone.

Usando las vistas de la interface de usuario de Protegé, los diseñadores de ontologías básicamente crean clases y les asignan propiedades. Usando la función *resulting ontologies*, Protegé es capaz de generar automáticamente interfaces de usuario que soporten la creación de instancias. En cada clase de la ontología el sistema crea un formulario con componentes de edición para cada propiedad de la clase. Protegé actualmente puede ser usado para cargar, editar y salvar ontologías en varios formatos como CLIPS, RDF, XML, UML, OWL.

La investigación que se realiza es de tipo exploratoria y se usan métodos empíricos tales como el análisis documental clásico para abordar las cuestiones referentes a la Gestión de Proyectos y las Ontologías. A partir de la revisión de la documentación y la literatura especializada, se obtuvieron los referentes teóricos y conceptuales. Se abordan las características principales, métodos, cuestiones del ciclo de vida, entre otros elementos, de la Gestión de Proyectos, para ello se tomó en cuenta los criterios expuestos en la Guía del PMBoK, el cual es un estándar desarrollado por el Project Management Institute, en sus diferentes versiones.

Este documento es reconocido en la literatura como una de las principales herramientas para la profesión y una referencia esencial para la biblioteca de cada proyecto de gestión profesional. Es un estándar reconocido internacionalmente por la (IEEE Std 1490-2009) que provee los fundamentos de la Gestión de Proyectos, los cuales son aplicables a un amplio rango de proyectos, incluyendo construcción, software, ingeniería. También ofrece un conjunto de procesos, generalmente reconocidas como buenas prácticas y que permite obtener resultados a través de industrias y organizaciones, con prácticas de base que impulsan los resultados del negocio para cualquier organización local, regional y mundial.

Para el diseño de la Ontología se ha seguido la metodología Methontology (la cual fue propuesta por el grupo de Ingeniería Ontológica de la Universidad Politécnica de Madrid) (2004, 1997, 2000). Debido a que Methontology propone un ciclo de vida basado en la evolución de prototipos que permite añadir, cambiar y eliminar términos en cada nueva versión (prototipo) de la Ontología, de acuerdo a las necesidades que se tengan, se propone para este estudio, además el esquema de plantillas, diagramas y

tablas utilizados en las tareas de conceptualización facilitan la integración y cooperación de desarrolladores y expertos del dominio. Esta metodología también sigue los pasos definidos en el estándar IEEE 1074 de desarrollo de software y sus actividades más significativas a groso modo son las siguientes:

La actividad de especificación permite determinar por qué se construye la Ontología, cuál será su uso, y quiénes serán sus usuarios finales.

La actividad de conceptualización merece una especial atención, ya que determina el resto de la creación de la Ontología.

Tarea 1: Construir el glosario de términos. El glosario de términos debe incluir todos los términos relevantes del dominio (conceptos, instancias, atributos, relaciones entre conceptos, etc.), sus descripciones en lenguaje natural.

Tarea 2: Construir la taxonomía de conceptos. Cuando el glosario de términos tenga una cantidad importante de elementos, se debe construir una taxonomía que defina la jerarquía entre los conceptos. Se debe evaluar que la taxonomía creada no contenga errores.

Tarea 3: Construir un diagrama de relaciones binarias. El objetivo de este diagrama es establecer las relaciones entre los conceptos.

Tarea 4: Construir el diccionario de conceptos Tarea 5: Definir las relaciones binarias en detalle. Se crea la tabla de relaciones binarias en la que se describe detalladamente todas las relaciones binarias incluidas en el diccionario de conceptos.

En lo adelante se refiere a construir, reglas, axiomas e instancias.

RESULTADOS

Es imposible representar el mundo real, o alguna parte de él, con todos los detalles. Para reproducir algún fenómeno o parte del mundo, llamado dominio, es necesario focalizar o limitar el número

Las ontologías para representar el conocimiento precisan los siguientes componentes : conceptos, relaciones, funciones, instancias y axiomas, siendo el conjunto de relaciones conceptuales un indicador complejo de una ontología.

De la revisión del PMBok se realizaron varias propuestas de representación de las clases del dominio de Gestión de Proyectos, como un primer paso para la construcción de la Ontología, las cuales se muestran a continuación:

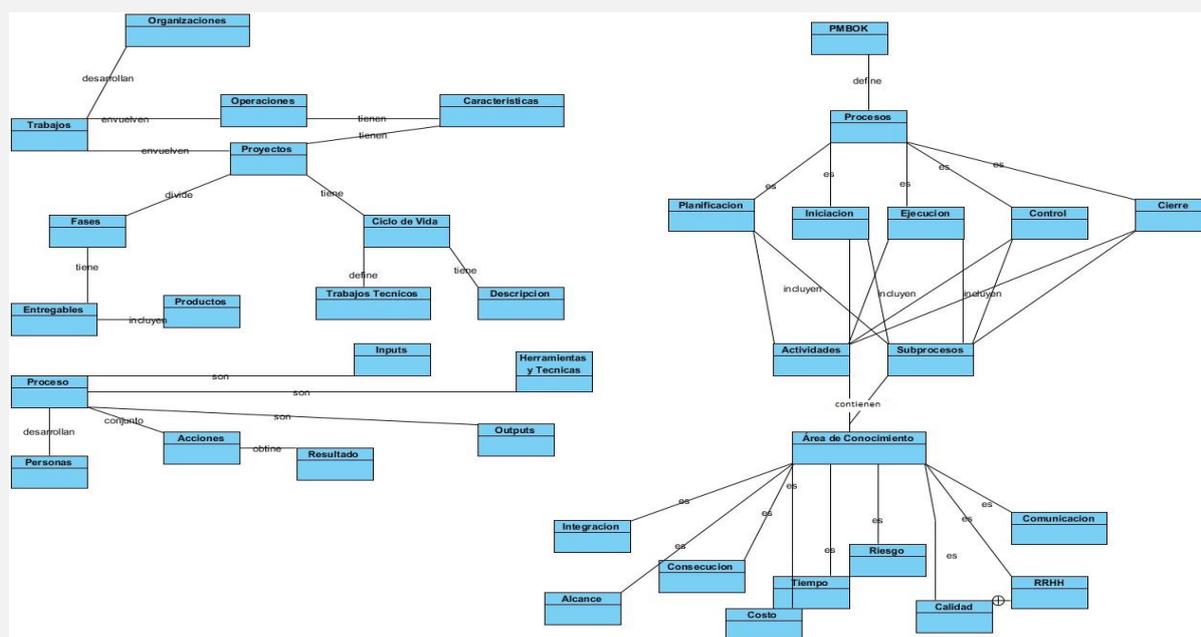


Figura 1. Diagrama de clases desde el PMBoK

Seguidamente fue construido un glosario de términos como segundo paso, el cual se muestra en la tabla a continuación:

Tabla 1. Glosario de términos

No	Término	Descripción	Tipo
1	Gestión	Proceso orientado a obtener resultados eficientes y eficaces, a partir de una utilización adecuada de las capacidades institucionales	Relación
2	Organización	Como parte del ciclo de dirección, se le vincula con las actividades de combinación de la distribución de los recursos humanos y materiales en función de las complejidades y magnitud de las tareas a cumplir.	Concepto
3	Proceso	Secuencia ordenada y lógica de actividades repetitivas que se realizan por una persona o grupo con la capacidad de transformar una entrada (inputs) en salidas o resultados o impactos programados (outputs) para un destinatario (dentro o fuera de la institución), añadiendo valor con determinada relevancia social.	Concepto
4	Proyecto	Conjunto de actividades coordinadas y controladas con fechas de inicio y de finalización, llevadas a cabo para lograr un objetivo, incluyendo las limitaciones de tiempo, costo y recursos.	Concepto
5	Calidad	Grado en el que unos conjuntos de elementos inherentes a una organización cumplen con una necesidad o estándar establecido. Es acción (Proceso de gestión) y efecto (Resultado)	Concepto
6	Comunicación	Acción y efecto de comunicar o comunicarse ya sea de forma verbal, escrita, por medio de señas. Conjunto de técnicas que permiten la difusión de mensajes escritos o audiovisuales a una audiencia vasta y heterogénea	Concepto

7	Control	Proceso o función de la dirección que tiene como objetivo dar seguimiento y evaluar el cumplimiento de una meta, objetivo, estrategia y otros elementos vinculados con los procesos universitarios	Concepto
8	Trabajo en equipo	Es una de las condiciones de trabajo de tipo psicológico que más influye en los trabajadores de forma positiva, es aquella que permite que haya compañerismo y trabajo en equipo.	Relación
9	Desarrolla	Es la relación que existe entre Organizaciones y Trabajos	Relación
10	Envuelven	Es la relación que existe entre Trabajos y Operaciones o Trabajos y Proyectos	Relación
11	Divide	Es la relación que existe entre Proyecto y Fases	Relación
12	Tiene	Es la relación que existe entre Proyecto y Ciclo de Vida o Fases y Entregable	Relación
13	Incluyen	Es la relación que existe entre Entregable y Producto	Relación
14	Conjunto	Es la relación que existe entre Proceso y Acciones	Relación
15	Son	Es la relación que existe entre Proceso e Inputs o Proceso y Herramientas-Técnicas o Proceso y Outputs	Relación
16	Obtiene	Es la relación que existe entre Accione y Resultados	Relación
17	Contiene	Es la relación que existe entre todos los procesos y sus actividades y sus subprocesos	Relación

Cuando el glosario de términos ya está confeccionado, el desarrollador de la Ontología construye las taxonomías de conceptos que definen su jerarquía. Definiéndose cuál es la clase padre y cuál la hija. En la tabla a continuación se muestra cómo quedarían confeccionadas las dependencias.

Tabla 2. Representación de conceptos según niveles de jerarquía

Clase Padre	Clase Hija
<i>Organización</i>	Trabajos
<i>Trabajos</i>	Operaciones
<i>Trabajos</i>	Proyecto
<i>Proyecto</i>	Fases
<i>proyecto</i>	Ciclo de Vida
<i>Fases</i>	Entregable
<i>Entregable</i>	Producto
<i>Procesos</i>	Input
<i>Proceso</i>	Herramienta-Técnicas
<i>Proceso</i>	Outputs
<i>Área-Conocimiento</i>	Alcance
<i>Área-Conocimiento</i>	Tiempo
<i>Área-Conocimiento</i>	Costo
<i>Área-Conocimiento</i>	Riesgo
<i>Área-Conocimiento</i>	Calidad
<i>Área-Conocimiento</i>	RRHH
<i>Área-Conocimiento</i>	Comunicación
<i>Área-Conocimiento</i>	Integración
<i>Área-Conocimiento</i>	Consecución
<i>Procesos-PMBOK</i>	Iniciación
<i>Procesos-PMBOK</i>	Planificación
<i>Procesos-PMBOK</i>	Ejecución
<i>Procesos-PMBOK</i>	Control

Procesos-PMBOK

Cierre

Para llevar la ontología del modelo conceptual al modelo formal se usa el editor Protégé 4.3, donde se muestran los conceptos, relaciones entre conceptos, instancias y atributos de la ontología.

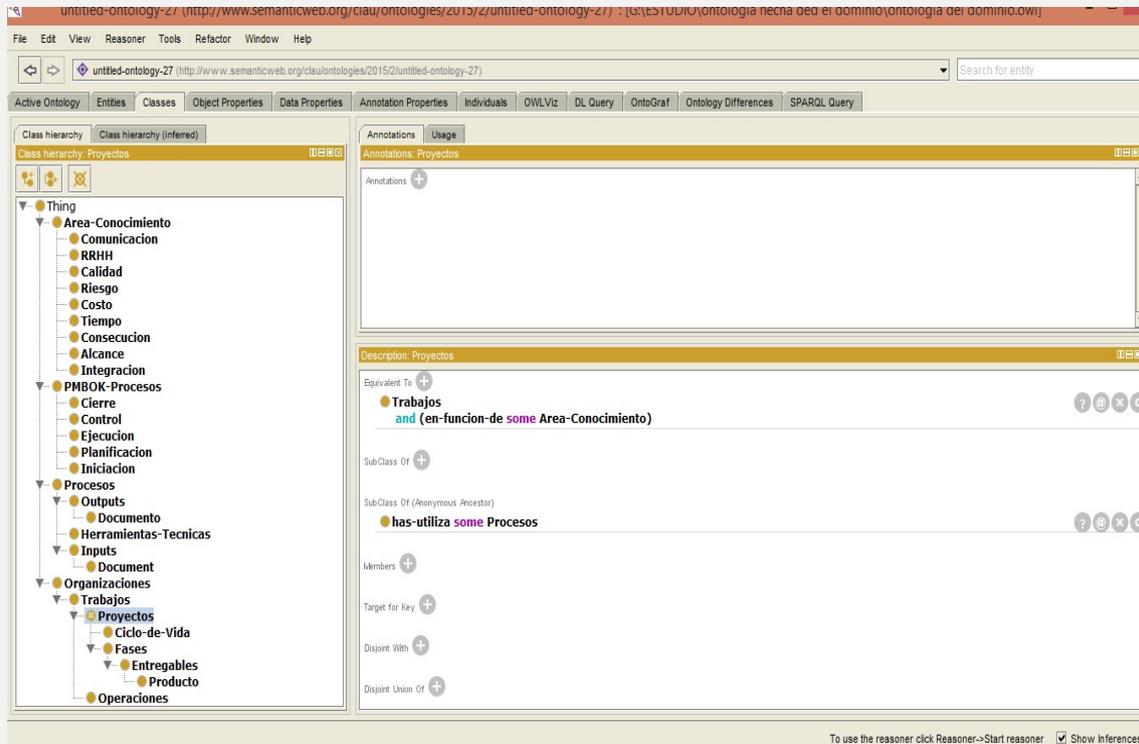


Figura 2. Creación de clases en el Protégé

Cuando se inicia el editor Protégé se trabaja primeramente en la pestaña Classes, todas las clases dependen de owl: Thing, que es una palabra clave, las pestañas más importantes son, Classes (Conceptos), Object Properties (Relaciones) y OWL Viz.

Las restricciones se usan para limitar las relaciones válidas entre individuos, pueden ser necesarias y suficiente o necesarias, adicionalmente se tienen cuantificadores universales o existenciales y se muestra la taxonomía de las clases en Protégé.



Figura 3. Clases en el protégé del dominio

Properties: las relaciones pueden ser de 2 tipos, “Object Properties” las cuales se establecen entre individuos y “Datatype Properties” la cual se establece entre individuos y esquemas xml, se recomienda nombrar las propiedades con la primera letra en minúscula. Las propiedades tienen unas características que se pueden asignar por ejemplo: Funcional, Funcional inversa, Simétrica y Transitiva, también es posible asignar los dominios y los rangos a cada propiedad, estos son axiomas que se utilizan para que el razonador haga ciertas inferencias, violar una restricción de dominio o de rango no significa necesariamente que la ontología sea inconsistente o que contenga errores.

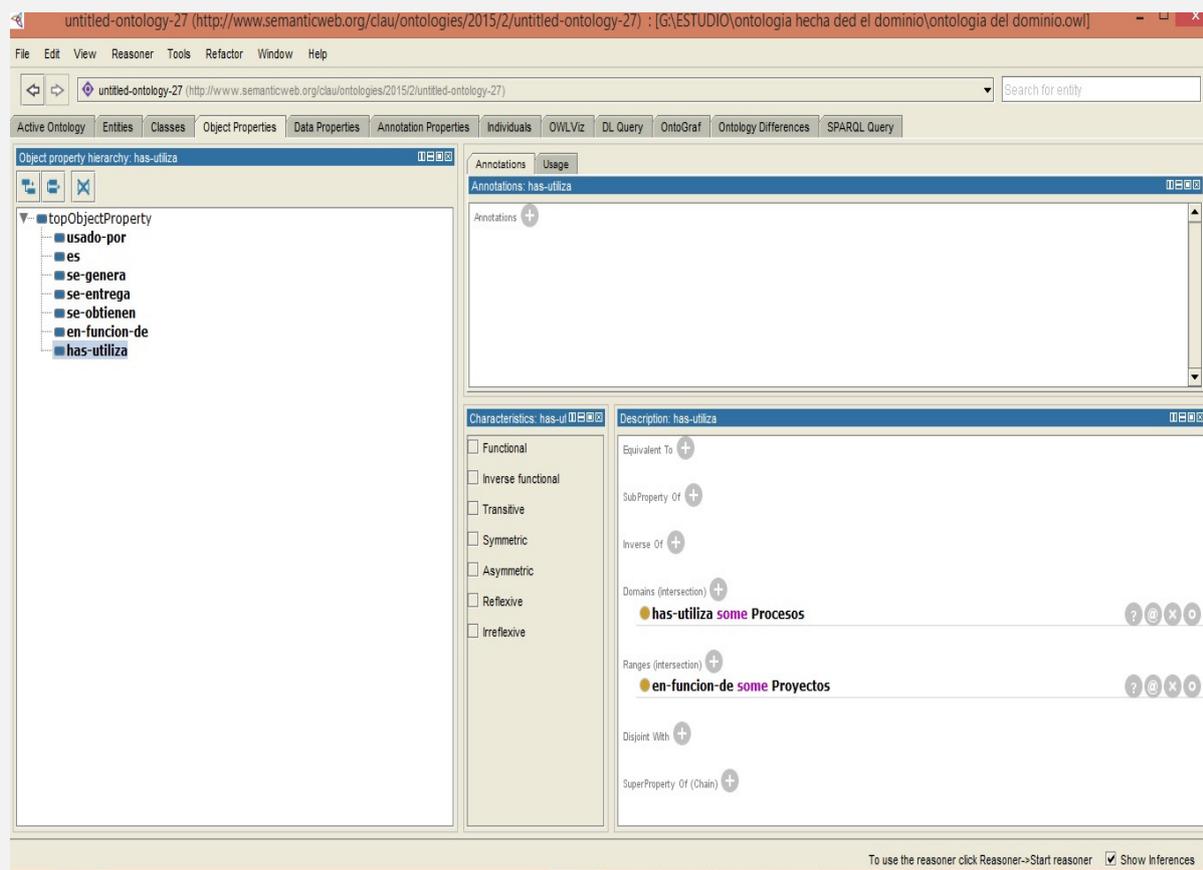


Figura 4. Creación de propiedades del dominio

OWL Viz: es un plugin para Protégé que permite visualizar con grafos los conceptos y las relaciones que tienen creadas, adicional a este plugin se debe tener instalado el GraphViz, las principales funciones de esta pantalla son:

- Mostrar clases
- Mostrar subclases
- Mostrar superclases
- Ocultar clases
- Ocultar subclases
- Ocultar superclases
- Ocultar clases especificando el radio
- Ocultar todas las clases

A continuación se muestra la (figura 5) de cómo quedaría la ontología Procedimiento de Pruebas de Software en el editor Protégé.



Figura 5. Modelo de la ontología mediante owlviz

También se muestra cómo quedaría la ontología en la pestaña OntoGraf del editor Protégé, aquí las clases están relacionadas unas con otras, donde cada color indica las diferentes relaciones que existen entre una clase y otra.

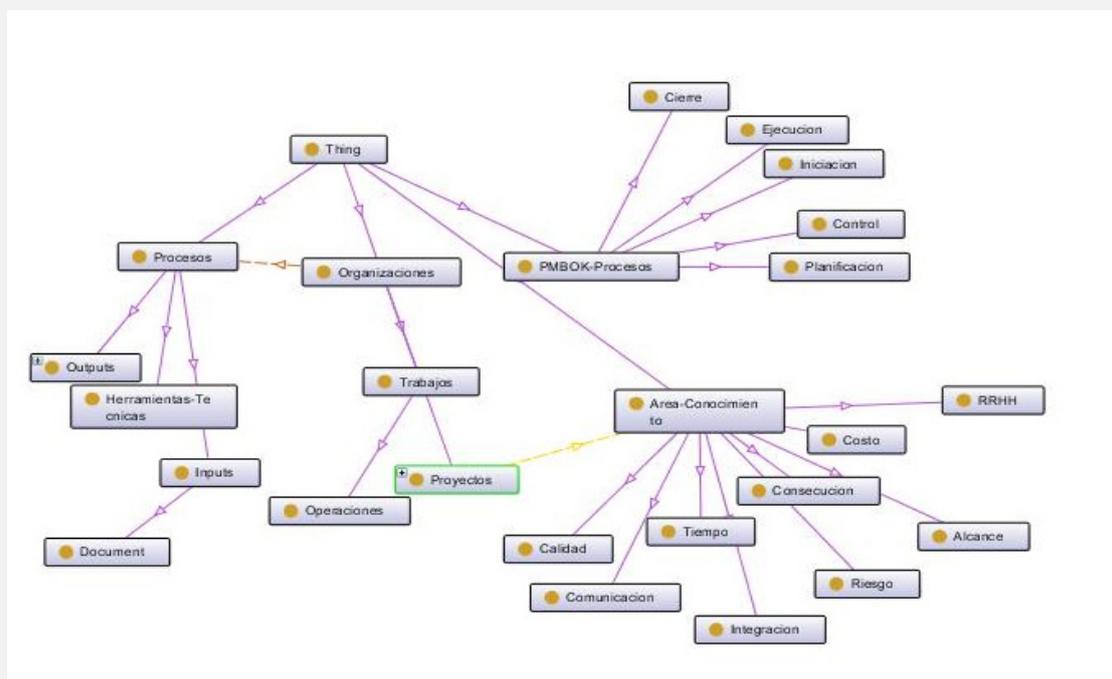


Figura 6. Modelo de clases mediante Ontograp

CONCLUSIONES

La realización de una ontología para el sistema de gestión de proyectos ofrece a la comunidad interesada una propuesta de organización y estructuración del conocimiento en el dominio de gestión de proyectos, lo cual trae consigo que solucione varios de los problemas planteados, como las búsquedas y obtención de información correcta o acorde a lo que se pregunta.

El método de Methontology permitió estructurar el conocimiento adquirido mediante técnicas, a través de un conjunto de tablas y diagramas que pueden ser entendidos por los expertos del dominio.

El diseño y desarrollo de la ontología utilizando el Protégé 4.3, así como la capacidad de este de permitir la recuperación de información, la convierte en una herramienta alternativa a considerar para consultar información durante el proceso de toma de decisiones en el dominio de gestión de proyectos.

Al utilizar protégé, permite la integración de varios razonadores en el mismo, e intentar garantizar un diseño ontológico semánticamente robusto para que las ontologías a realizar puedan responder de forma concreta en el proceso de recuperación de información.

RECOMENDACIONES

- Continuar enriqueciendo la ontología con condiciones que permitan agregar valor semántico a la misma.

REFERENCIAS BIBLIOGRÁFICAS

- Ajenjo, A. D. Dirección y Gestión de Proyectos. Un enfoque práctico (Segunda ed.), 2005.
- Barba-Romero, S. Multicriterion decision in management: principles and practice: Kluwer Academic Publishers, 2000.
- Barchini y Álvarez-Herrera, (2006). Sistemas de Información: Nuevos escenarios basados en Ontologías. Revista de Gestão da Tecnologia e Sistemas de Informação 2006, 3(1):3-18 ISSN online: 1807-1775
- Bechhofer, S. OWL: Web ontology language. In Encyclopedia of Database Systems (pp. 2008-2009). Springer US.2009
- Corcho, O., Fernández, M., Gómez-Pérez, A., y López, A. Building legal ontologies with METHONTOLOGY and WebODE. Law and the Semantic Web. Legal Ontologies, Methodologies, Legal Information Retrieval, and Applications. Springer-Verlag, LNAI 3369. Chandrasekaran B., Josephson J.R., Benjamins V.R., What are Ontologies, and Why Do We Need Them? IEEE Intelligent Systems, 1999, 14(1): 20–26.
- Duncan, W. PMBOK—A Guide to the Project Management Body of Knowledge. ZDA: Project Management Institute (PMI), 1996.
- Farquhar, A. Ontolingua Tutorial. University of Stanford. Knowledge Systems Lab Stanford University. California, Estados Unidos. [Internet] 1997 [citado 20 febrero 2016]. Recuperado de: <http://www.wksl.stanford.edu/people/axf/tutorial.pdf>
- Gómez-Pérez A., Fernández-López M. y Corcho M., Ontological Engineering, Springer Verlag London, Mayo, 2004.
- Guarino N., Giaretta P., Ontologies and Knowledge Bases: Towards a Terminological Clarification. In: Mars N (ed) Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing (KBKS'95), University of Twente, Enschede, The Netherlands. IOS Press, Amsterdam, the Netherlands, pp 25–32. 1995.
- Gruber, T. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. Available as Technical Report *KSL 93-04*, Knowledge Systems Laboratory, Stanford University. [Internet]. 1993 [citado abril 2016] Recuperado de: <http://citeseer.ist.psu.edu/gruber93toward.html>
- Gruninger, M.; Fox, M. Methodology for the Design and Evaluation of Ontologies. En: Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI-95, Montreal, 1998.
- Juristo, N. Methontology: from ontological art towards ontological engineering, 1997.
- Kerzner, H. Project Management (Third ed.). New York: Van Nostrand Reinhold, 1989.
- Kirsch, L. Software project management: An integrated perspective for an emerging paradigm. Framing the Domains of IT Management: Projecting the Future. Through the Past, 2000: 285-304.

McGuinness D.L., Ontologies Come of Age, in *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*, eds. D. Fensel, J. Hendler, H. Lieberman, & W. Wahlster, MIT Press, Boston, MA, 2002:1–18.

Noy 2000 Noy N. y McGuinness D. *Ontology development 101: A Guide to creating your first ontology*. Stanford University. Stanford knowledge Systems Laboratory. Technical Report KSL-01-05. 2001 [Internet] 2005 [citado 20 febrero 2016]. Recuperado de: http://www.cs.man.ac.uk/~ocorcho/documents/LawSemWeb2004_CorchoEtAl.pdf

PMI. *A Guide to the Project Management Body of Knowledge (PMBOK® Guide)*. Estados Unidos de América : Project Management Institute, 2009.

Piñero et al., Gespro. Paquete para la gestión de proyectos. Nueva Empresa. *Revista cubana de gestión empresarial*, 2013, 9(1):45-53, ISSN 1682-2455. La Habana, Cuba.

Pressman. *Ingeniería de Software (Sexta ed.)*. S.L: Mc Graw Hill, 2005.

Ramos E. y Nuñez H., *ONTOLOGÍAS: Componentes, Metodologías, Lenguajes, Herramientas y Aplicaciones*, *Lecturas en Ciencias de la Computación*, Universidad Central de Venezuela, ISSN 1316-6239, 2012.

Schreiber, G., Wielinga, B., de Hoog, R., Akkermans, H., y Van de Velde, W., *CommonKADS: A Comprehensive Methodology for KBS Development*, *IEEE Expert*, 1994, 9(6):28– 37.

Swartout B. et.al. *Toward Distributed Use of Large-Scale Ontologies*. In: Farquhar A, Gruninger M, Gómez-Pérez A, Uschold M, van der Vet P (eds) *AAAI'97 Spring Symposium on Ontological Engineering*. Stanford University, California, 1997: 138-148.

Sure, Y. *D1.2.3 Methods for Ontology Evaluation*. Knowledge Web Consortium. Project Number IST-2004-507-507-482. 2005.

Uschold M., Gruninger M., *Ontologies: Principles, Methods and Applications*. AIAI- TR-191. *Knowledge Engineering Review*. [Internet] 2007 [citado 3 abril 2016] 11(2). Recuperado de: <http://citeseer.ist.psu.edu/uschold96ontologie.html>, 2007.



REVISÃO SISTEMÁTICA BASEADA EM PESQUISA BIBLIOGRÁFICA ESTRUTURADA – PPBE: UM MAPEAMENTO SOBRE ANÁLISE FACETADA APLICADA À ARQUITETURA DA INFORMAÇÃO

Graciane Silva Bruzuinga Borges¹, Gercina Ângela de Lima²

¹Escola de Ciência da Informação, Universidade Federal de Minas Gerais, 0000-0002-6677-9702, gracianebruzuinga@gmail.com

²Escola de Ciência da Informação, Universidade Federal de Minas Gerais, 0000-0003-0735-3856, limagercina@gmail.com

RESUMO A gestão de conteúdos textuais dentro das organizações é vital para um fluxo informacional contínuo e eficiente, que irá impactar diretamente nos processos decisórios internos e no reuso de dados. Por meio dos recursos oriundos da Arquitetura da Informação – AI – é possível o desenvolvimento de ambientes informacionais digitais mais bem estruturados. Com o objetivo de analisar, a partir da literatura da área, as possibilidades metodológicas para modelagem desse tipo de ambiente, apresenta-se um modelo de revisão sistemática baseado em um procedimento de pesquisa bibliográfica estruturada. O método de trabalho é caracterizado pelo viés analítico-sintético, por meio do qual os documentos são identificados, selecionados e sintetizados de maneira sistematizada. Esse método prevê três módulos principais: 1) Processo de Pesquisa Bibliográfica Estruturada – PPBE; 2) Processo de Gestão de Resultados – PGR; e 3) Processo de Redação – PR, podendo ser utilizada para subsidiar os capítulos correspondentes à revisão de literatura; fundamentação histórica, teórica e metodológica, todos previstos em teses e dissertações. Como resultados, foram obtidos 68 documentos no período de 1985 a 2017, sendo: 15 documentos para subsidiar a revisão de literatura; 25 documentos para a fundamentação teórica e 28 documentos para a fundamentação metodológica. Conclui-se que a metodologia é consistente e será capaz de subsidiar o mapeamento das temáticas evidenciadas, sendo também aplicável a outras áreas do conhecimento.

PALAVRAS-CHAVE *Revisão Sistemática, Pesquisa Bibliográfica Estruturada, Arquitetura da Informação, Análise Facetada. Ambiente Informacional Digital.*

ABSTRACT The management textual content in the organizations is vital to a continuous and efficient information flow that will impact directly on the internal decision-making processes and data reuse. Through the resources provided by the Information Architecture - IA - it is possible to develop better structured informational digital environments. The object of this work is to analyse, from the literature of the area, the methodological possibilities for the modeling of this type of environment. So, a systematic review model is presented based on a structured bibliographic research procedure. The work method is characterized by the analytical-synthetic bias, through which documents are identified, selected and synthesized in a systematized way. This method provides three main steps: 1) Structured Bibliographic Research Process - SBRP; 2) Results Management Process - RMP; and 3) Writing Process - WP, it can be used to develop the chapters of the literature review; historical, theoretical and methodological foundation, all foreseen in theses and dissertations. As a result, 68 documents were obtained in the period from 1985 to 2017. Among these documents: 15 of them are to support the literature review; 25 for the theoretical basis and 28 for the

methodological basis. It is concluded that the methodology is consistent and that it will be able to support the mapping of the emphasized topics. Moreover, it is applicable to other areas of knowledge.

KEY WORDS *Systematic Review, Structured Bibliographic Research, Information Architecture, Faceted Analysis. Digital Information Environment.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Atualmente, observam-se investimentos expressivos das organizações no desenvolvimento de ambientes informacionais digitais, entendidos como sistemas de gestão de conteúdos. Estes visam armazenar, recuperar e compartilhar dados e informações estratégicos, possibilitando o reuso desses conteúdos e o favorecimento do processo decisório. Contudo, as chances de essas organizações obterem soluções efetivamente eficazes não são muito grandes, considerando o alto índice de projetos que são executados de maneira inconsistente. Projetos que levem em conta ferramentas oriundas da Arquitetura da Informação – AI – e fundamentados em uma metodologia consistente de modelagem terão mais chance de sucesso. Por sua característica de múltiplas possibilidades de representação da informação, acredita-se que a técnica da Análise Facetada – AF – poderá contribuir para o desenvolvimento de soluções mais bem modeladas às necessidades dos usuários destes ambientes.

A falta de sistemática e de controle dos fluxos informacionais nas organizações de médio e de grande porte prejudica a comunicação interna; aumenta ou mesmo impossibilita a recuperação eficiente de dados relevantes; causa retrabalho dispendioso e pode acarretar em perdas financeiras. De acordo com Davenport (2000), o principal motivo pelo qual os projetos de AI deveriam ser elaborados nas empresas deveria ser uma real mudança de comportamento dos colaboradores. Gerenciamento de informação pressupõe gestão de mudança, afinal, em um cenário de desenvolvimento de sistemas de gestão de conteúdos o uso efetivo desse sistema é, ou deveria ser, o principal objetivo da AI. Na perspectiva de Davenport (2000, p. 202), para a implementação de uma arquitetura eficiente é necessário considerar que ela mude a maneira como os colaboradores utilizam e lidam com as informações, do contrário não faria sentido propô-la.

O objetivo deste artigo é apresentar um modelo de revisão sistemática que tem como base metodológica a pesquisa bibliográfica do tipo estruturada. O intuito é identificar trabalhos oriundos da Arquitetura da Informação que abordem modelos aplicados ao desenvolvimento de ambientes informacionais digitais por meio do uso da abordagem da análise facetada.

ARQUITETURA DA INFORMAÇÃO E ANÁLISE FACETADA: INTERSEÇÕES

A AI serve para conduzir os usuários às informações de que necessitam da maneira mais direta e intuitiva possível, sendo responsável por, entre outras funções, modelar o ambiente de negócio e possibilitar a fluidez entre os setores geradores e consumidores de conteúdos. Davenport (2000, p. 205) aponta três motivos básicos de falha das arquiteturas quanto ao aspecto comportamental: 1) o arquiteto da informação não considerar a mudança de comportamento como objetivo principal do processo; 2) o conteúdo ser incompreensível para os usuários; e 3) o processo, em si, de desenvolvimento de AI inibir

as mudanças. Para o autor, uma ferramenta orientada ao cliente deverá ter princípios para definir a AI, que poderão ser a ponte entre a estratégia e a arquitetura.

A origem da AI remonta à década de 1970, e o termo *Arquitetura da Informação* foi cunhado por Richard Wurman (Wurman, 2000). Na década de 1990 e anos 2000, a área se desenvolveu de forma expressiva, impulsionada pelos trabalhos de Morville e Rosenfeld (2006), e se consolida no campo da *web*, influenciando em desenho estrutural de sistemas, usabilidade, recuperação da informação, *design* e interface.

Atualmente, a AI se apresenta como um campo voltado para a modelagem dos recursos disponíveis para as necessidades informacionais dos usuários. No entendimento de Davenport (2000, p. 200), a Arquitetura da Informação simplesmente se constitui de uma série de ferramentas que adaptam os recursos às necessidades de informação. Segundo o autor, a AI faz a ponte entre o comportamento, os processos e o pessoal especializado e outros âmbitos da empresa, como métodos administrativos, estrutura organizacional e espaço físico. Seguindo esta mesma abordagem, Lima (2016, p. 47) afirma que a AI surgiu como uma proposta para o desenho de estruturas informacionais em espaços digitais e, posteriormente, como uma alternativa de modelo de gestão do conhecimento nas organizações e na estruturação de *websites*.

Ao analisar e representar um documento, o ser humano tem como objetivo sua organização e, como consequência, a recuperação da informação nele contida, bem como sua utilização para geração de novos conhecimentos e solução de possíveis problemas. Para que o processo de recuperação da informação seja eficiente, é imprescindível que se estabeleça uma representação consistente, que partirá sempre do princípio da classificação. Para Speziali (1973, p. 462 como referido em Lima, 2004, p. 72), o objetivo de toda classificação é estabelecer uma ordem ou organização das coisas e dos pensamentos. Para a autora, a classificação facetada pode mostrar tanto um mapeamento do conhecimento científico de um determinado período, quando permitir a descoberta de conexões entre áreas, facilitando a recuperação da informação.

O princípio de Ranganathan sugere múltiplas dimensões de assuntos para análise e representação de determinado documento ou objeto. Originário da Biblioteconomia e idealizada por Shiyali Ramamrita Ranganathan (1892-1972) na década de 1930, seu *prolegomena* é composto por 46 cânones, 13 postulados e 22 princípios, registrada em cinco obras, dentre elas o *Prolegomena to Library Classification*, de 1933, publicada em 1967 (Lima, 2004; Maculan, 2014). De acordo com Maculan (2014, p. 114), este conceito da multidimensionalidade foi inspirado na ideia da “Árvore Baniana”, um tipo de figura indiana. A *Banyan Tree* se caracteriza por ocupar uma grande área e espalhar seus galhos pelo solo, onde estes desenvolvem raízes aéreas que nascem do tronco principal e criam novos troncos adicionais.

Depois que Ranganathan divulgou seu sistema de classificação chamado *Colon classification* (Classificação dos dois pontos), onde introduziu, entre outros novos termos, o termo *faceta* [...] ficou sendo, nos modernos estudos sobre teoria da classificação, o substituto de característica. Na realidade, os dois significam os diferentes aspectos ou pontos de vista dos assuntos. (Barbosa, 1969, p. 16)

O método analítico-sintético usado na construção da Classificação dos Dois Pontos é reconhecido como um marco na teoria da classificação bibliográfica, constituindo uma metodologia que influenciou, também, a elaboração de tesouros e mesmo a construção de ontologias, sendo comparado a um mapa do conhecimento (Souza, p. 84). Para Barbosa (1969), esse método corresponde ao uso de símbolos de

ligação que permitem as correlações dos assuntos, aspecto também observado na Classificação Decimal Universal – CDU (p. 25).

De acordo com Barbosa (1969) “depois da criação, em 1948, em Londres, do *Classification Research Group*, a classificação por facetas passou a ser uma realidade para campos especializados e, em muito menor grau, seu conceito filosófico, onde cada aspecto de um assunto é visto sob a manifestação das cinco categorias fundamentais.” (p. 166). Personalidade, Matéria, Energia, Espaço e Tempo – PMEST – correspondem às categorias mencionadas. De acordo com Lima (2002, p. 191), podem-se verificar muitos aspectos em um único assunto, e deve-se tentar sintetizá-los de maneira a descrevê-los da forma mais adequada possível.

A REVISÃO SISTEMÁTICA

Trabalhos científicos preveem uma consistente verificação e análise da literatura a respeito da temática abordada. Esse processo passa por etapas que vão desde a busca bibliográfica até a sua elaboração textual. Existem diferentes técnicas de revisão de literatura. Entre elas, está a revisão tradicional ou narrativa, e a sistemática. A revisão tradicional geralmente é mais ampla e é apresentada de forma mais geral. A revisão sistemática possui uma metodologia pré-determinada, com técnicas e procedimentos que podem ser reproduzidos. Segundo Greenhalgh (1997, p. 672 como referido em Botelho, Cunha, & Macedo, 2011), a revisão sistemática é considerada “como uma síntese de estudos primários que contém objetivos, materiais e métodos claramente explicitados e que foi conduzida de acordo com uma metodologia clara e reproduzível”. A colaboração Cochrane recomenda que a revisão bibliográfica sistemática seja efetuada em sete passos: 1) formulação da pergunta; 2) localização dos estudos; 3) avaliação crítica dos estudos; 4) coleta dos dados 5) análise e apresentação dos dados; 6) interpretação dos dados; 7) aprimoramento e atualização da revisão (Rother, 2007).

Para se observar e buscar uma solução para determinado problema complexo, a partir da literatura, é imprescindível aumentar a precisão das estimativas e resolver discrepâncias existentes entre os vários artigos que tratam do assunto. Para Testoni (2015), na revisão narrativa, o problema pode residir em não haver, de fato, um processo real no qual os artigos são selecionados, podendo o pesquisador escolher somente os artigos que estiverem de acordo com suas ideias iniciais, o que poderia, se esse processo for realizado de maneira inadequada, gerar viés. Já na revisão sistemática, o pesquisador não selecionará qualquer artigo, pois obedecerá a um sistema, no qual ele faz um levantamento da literatura de maneira sistemática.

Assim, é apresentado um modelo de Revisão Sistemática baseada em um Procedimento de Pesquisa Bibliográfica Estruturado – PPBE –, que cobre desde a definição da questão de pesquisa e da estratégia de busca até a síntese dos documentos selecionados nas fontes de informação. Este modelo de revisão é uma iniciativa dos doutorandos Graciane Borges; Hélder Firmino e Cristiane Mendes e da Prof^ª Dr^ª Gercina Ângela de Lima, todos oriundos da Universidade Federal de Minas Gerais (Brasil).

METODOLOGIA

Quanto à finalidade do estudo, esta pesquisa é caracterizada como *descritiva*, que, de acordo com Best (1972), é aquela que “delineia o que é”, considerando: 1) descrição; 2) registro; 3) análise; e 4) interpretação de fenômenos atuais. Também apresenta característica de uma *exploração técnica e*

sistemática, onde o investigador, baseando-se em conhecimentos teóricos anteriores, planeja cuidadosamente o método a ser utilizado, formulando problemas e hipóteses e registrando metodicamente os dados coletados para análise (Best, 1972). O modelo de revisão sistemática adotado é caracterizado pelo viés analítico-sintético e composto por três módulos: 1) Processo de Pesquisa Bibliográfica Estruturado – PPBE; 2) Processo de Gestão de Resultados – PGR; e 3) Processo de Redação – PR (ver Figura 1):



Figura 1 – Modelo de revisão sistemática

Referência: Borges et al. (2017, no prelo).

MÓDULO 1 - PROCESSO DE PESQUISA BIBLIOGRÁFICA ESTRUTURADO – PPBE

Etapa 1: questão de pesquisa

Delimitadora do assunto sobre o qual se deseja levantar a literatura.

- Passo 1 - Identificar o sujeito da pesquisa;
- Passo 2 - Identificar o objeto da pesquisa;
- Passo 3 - Especificar o sujeito e o objeto;
- Passo 4 - Redigir a(s) questão(ões) de pesquisa;
- Passo 5 - Destacar as palavras significativas.

Identifica-se o sujeito da pesquisa, que é a realidade a respeito da qual se deseja saber, ou seja, o universo de referência. Posteriormente, identifica-se o objeto, que é o tema sobre o sujeito. A especificação do sujeito e do objeto se dá por meio de adjetivos explicativos e restritivos e complementos nominais e circunstanciais (Salvador, 1980, p. 46-48). Elabora-se uma frase objetiva que reflita os limites da temática, destacando as palavras significativas que serão utilizadas na realização das buscas preliminares.

Etapa 2: estratégia de busca

Instrumento norteador para acesso às fontes de informação.

- Passo 1 - Definir controle terminológico;
- Passo 2 - Definir *strings* de busca;
- Passo 3 - Definir critério(s) de inclusão e exclusão;
- Passo 4 - Definir fontes de pesquisa;
- Passo 5 - Definir ferramenta de gestão bibliográfica;
- Passo 6 - Montar Protocolo de Levantamento;
- Passo 7 - Validar Protocolo de Levantamento.

Inicia-se o controle terminológico por meio de buscas preliminares realizadas com o uso das palavras significativas da questão de pesquisa. É levantada a bibliografia inicial de onde será extraído um conjunto de palavras-chave em linguagem natural. As buscas podem ser realizadas em acervo próprio e em fontes de informação validadas no meio científico. As palavras-chave são traduzidas em descritores com o uso de vocabulário controlado. A bibliografia inicial é incorporada à lista final de documentos selecionados. Definem-se as *strings* de busca, que são esquemas de termos combinados por meio de operadores *booleanos* e de caracteres especiais, utilizados no truncamento. São estabelecidos os critérios de inclusão e exclusão dos documentos: 1) intervalo temporal; 2) idioma; 3) tipo documental; 4) fator de impacto do periódico, considerando o critério de seleção do *Science Citation Index* (SCI), que é calculado anualmente pelo *Institute for Scientific Information/Thompson Scientific Reuters* para as revistas indexadas em sua base de dados e publicado pelo *Journal Citations Reports* (JCR) (Garfield, 1956 como referido em Thomaz, Assad, & Moreira, 2011); e 5) qualis-periódicos, que atesta a qualidade dos artigos por meio da análise da qualidade dos periódicos científicos (Qualis-Periódicos, 2017). Definem-se as fontes de informação, que, de acordo com Tomaél et al. (2004, p. 23), para serem avaliadas, deve-se identificar o indivíduo ou instituição responsável por sua compilação. Quanto à ferramenta de gestão bibliográfica a ser utilizada, sugere-se: Mendeley; Zotero e EndNote. Finalmente, procede-se com o preenchimento do Protocolo de Levantamento.

Etapa 3: levantamento e localização

Implementação do processo de pesquisa por meio do Protocolo de Levantamento.

- Passo 1 - Acessar fontes de pesquisa;
- Passo 2 - Executar Protocolo de Levantamento;
- Passo 3 - Realizar Teste de aderência;
- Passo 4 - Analisar listas de resultados;
- Passo 5 - Localizar documentos em texto completo.

Acessam-se as fontes de informação determinadas para a execução do protocolo. Analisam-se as listas de resultados das primeiras buscas por meio da realização de um Teste de Aderência, que irá verificar a pertinência dos documentos retornados na primeira página de resultados após aplicação de 50% das *strings* de busca. Obtendo-se avaliação de aderência fraca, deve-se retornar aos passos 2 e 3 da etapa de estratégia de busca e remodelar seus itens de composição. Procede-se com a localização dos documentos em texto completo.

Etapa 4: seleção

Listagem das referências que irão compor a revisão.

- Passo 1 - Realizar leitura exploratória;
- Passo 2 - Realizar leitura seletiva;
- Passo 3 - Compilar a Lista de Referências;
- Passo 4 - Refinar Lista de Referências.

Realiza-se a leitura exploratória do texto a fim de verificar em que medida a obra interessa à pesquisa. A leitura exploratória, de acordo com Gil (2006, p. 77-78), tem natureza de *reconhecimento do tema*, analisando-se: 1) folha de rosto; 2) resumo; 3) índice; 4) notas; 5) introdução; 6) prefácio; 7) conclusões; e 8) orelhas. Posteriormente, executa-se a leitura seletiva, com o intuito de identificar e selecionar aqueles documentos que respondam à(s) questão(ões) de pesquisa(s). Também de acordo com Gil (2006, p. 78), a leitura seletiva caracteriza-se por uma natureza crítica, preconizando: 1) ter em mente os objetivos da pesquisa; e 2) ler textos paralelos pertinentes ao problema a fim de ampliar a compreensão do tema. Obtêm-se a compilação da Lista de Referências, que deverá ser refinada por meio da validação de um especialista da área.

MÓDULO 2 - PROCESSO DE GESTÃO DE RESULTADOS – PGR

Etapa 1: arquivamento

Manipulação controlada dos documentos, evitando-se perda de arquivos e dificuldade de localização dos itens citados.

- Passo 1 - Padronizar título dos documentos;
- Passo 2 - Padronizar título dos formulários de fichamento;
- Passo 3 - Salvar documentos na ferramenta de gestão bibliográfica.

Padronizam-se os títulos dos arquivos armazenados e de seus respectivos formulários de fichamento. Realiza-se o armazenamento efetivo dos documentos selecionados na ferramenta de gestão bibliográfica definida.

Etapa 2: manutenção

Controle das regras definidas durante o período de execução da pesquisa.

- Passo 1 - Registrar Estratégia de Busca;
- Passo 2 - Definir pastas locais;
- Passo 3 - Criar alertas das *strings* de busca.

Começa-se pelo registro da estratégia de busca estabelecida. Posteriormente, compõe-se uma estrutura local de pastas e subpastas no computador utilizado para redação e devem-se estabelecer regras de *backup* em HD externo. A estrutura deverá ser compatível com a utilizada na ferramenta de gestão, compondo um sistema de dupla garantia de segurança dos acervos. Definem-se alertas das *strings* de buscas nas bases de dados selecionadas, a fim de manter o pesquisador atualizado acerca de documentos recentemente publicados sobre o tema.

MÓDULO 3 - PROCESSO DE REDAÇÃO – PR

Etapa 1: fichamento

Análise e síntese dos documentos selecionados, a fim de responder à(s) questão(s) de pesquisa definida(s).

- Passo 1 - Preparar Modelo de Formulário de fichamento;
- Passo 2 - Elaborar referência do documento;
- Passo 3 - Analisar o resumo do documento;
- Passo 4 - Definir parâmetros de análise;
- Passo 5 - Definir critérios de identificação;
- Passo 6 - Realizar síntese do documento;
- Passo 7 - Extrair citações diretas;
- Passo 8 - Elaborar analíticas.

Prepara-se o modelo de Formulário de Fichamento a ser utilizado. Passa-se à elaboração e normalização da referência do documento analisado. Transcreve-se o resumo original do documento e destacam-se termos representativos. Definem-se os parâmetros e critérios que irão nortear a análise e síntese do texto. Realiza-se a síntese do documento, viabilizada por meio da *leitura analítica* do texto. Para Gil (2006, p. 78-79), essa leitura caracteriza-se por uma natureza de análise de dados, objetivando: 1) ordenar e sumarizar as informações contidas nas fontes; e 2) responder ao problema de pesquisa. Para sua execução, o autor sugere: 1) ter objetividade, imparcialidade e respeito; 2) realizar a leitura integral do documento, usando dicionários e consultando trabalhos correlatos; 3) identificar as ideias principais do texto: ler uma frase e identificar as palavras-chave; ler um parágrafo e escolher a frase de síntese deste; selecionar parágrafos significativos e sintetizar suas ideias principais; 4) hierarquizar as ideias, organizando-as segundo sua ordem de importância no contexto da pesquisa: ideias principais e ideias secundárias; estabelecendo categorias de análise; e 5) sintetizar as ideias, recompondo o todo decomposto pela análise: eliminar o que é secundário e fixar no essencial à solução do problema. Após a leitura, passa-se à extração de citações diretas que serão utilizadas na redação da revisão, devendo-se reproduzir a citação literal, registrando no formulário: 1) autoria; 2) data de publicação; e 3) página. O último passo corresponde à elaboração de analíticas: breves comentários sobre citações, úteis para facilitar o encadeamento das ideias durante a redação.

Etapa 2: escrita

Finalização do procedimento geral.

- Passo 1 - Definir formato do texto;
- Passo 2 - Redigir o texto;

- Passo 3 - Revisar o texto;
- Passo 4 - Normalizar documentária.

De acordo com Naves (2017), a linguagem científica deve preconizar: 1) uso do verbo de maneira impessoal; 2) não adoção da primeira pessoa do singular ou do plural; 3) criação de elos entre os parágrafos, dando sequência e coerência ao texto; 4) uso de comparações, mostrando consensos, discordâncias e tendências. Define-se o formato do texto de acordo o objetivo da revisão, podendo ser: 1) revisão de literatura, priorizar parâmetros e abordagens utilizados; 2) fundamentação histórica, identificar a evolução da temática, marcos históricos e paradigmas; 3) fundamentação teórica, priorizar teorias utilizadas; e 4) fundamentação metodológica, priorizar métodos e técnicas utilizados. Passa-se à redação do texto por meio da técnica de *leitura interpretativa*. Para Gil (2006, p. 79-80), essa leitura caracteriza-se por uma natureza correlacional com outros conhecimentos, objetivando relacionar as ideias do autor com o problema de pesquisa em questão. Para execução, o autor sugere ligar os dados a conhecimentos significativos, como pesquisas empíricas e teorias comprovadas. Revisa-se o conteúdo de acordo com o padrão culto da língua vernácula. Finalmente, o pesquisador deverá providenciar a normalização do texto, obedecendo a instituições reguladoras nacionais e/ou internacionais. Devem-se formatar os elementos pré-textuais, textuais e pós-textuais, considerando: 1) elaboração de referências e 2) controle de citações.

Apresentado o detalhamento procedimental da revisão, passa-se, no capítulo seguinte, à apresentação dos resultados obtidos com a aplicação do método.

RESULTADOS

A partir de uma primeira aplicação do método, foi obtido o seguinte resultado:

- 15 documentos para subsidiar a revisão de literatura, sendo:
 - questão de pesquisa número 1, “AI aplicada às organizações”: 12 itens;
 - questão de pesquisa número 2, “AI baseada na técnica da análise facetada”: três itens.
- 25 documentos para subsidiar a fundamentação teórica;
- 28 documentos para subsidiar a fundamentação metodológica.

Após análise dos documentos, foi possível destacar algumas dificuldades do cotidiano das organizações que justificam o investimento em projetos de AI, tais como: a) problemas na gestão do conhecimento interno; b) a dificuldade no reuso de dados; c) falta de controle dos fluxos informacionais; d) vulnerabilidade do sistema de segurança da informação; e, ainda, e) inexistência de padronização terminológica.

A técnica de análise facetada mostrou-se adequada ao processo de modelagens de ambientes informacionais digitais sob a luz da Arquitetura da Informação, como demonstrado na totalidade da amostra referente à questão de pesquisa número 2.

Paralelamente, evidenciaram-se alguns benefícios da gestão de conteúdos: 1) maior compartilhamento de informações; 2) melhoria na comunicação interna; 3) padronização e controle terminológicos; 4) processo de tomada de decisão com mais assertividade; 5) controle dos fluxos informacionais; 6) melhores índices de precisão e menor tempo de resposta na recuperação de conteúdos; 7) possibilidade

real de reuso de dados e informações internos; 8) segurança da informação; 9) redução de perdas financeiras; 10) condições internas para a gestão do conhecimento da empresa; 11) sistematização do processo de armazenamento e recuperação de conteúdos.

CONCLUSÕES

O resultado obtido por meio da aplicação do método de Revisão Sistemática baseada em PPBE facilitou o mapeamento da literatura, tanto da temática mais abrangente quanto da temática mais específica e dentro do período de pesquisa coberto (1985 a 2017). Apesar de ter sido recuperado um número reduzido de textos referentes ao eixo temático central da pesquisa, eles trouxeram importantes aportes que irão auxiliar em pesquisas futuras.

O modelo mostrou-se consistente e será capaz de subsidiar o mapeamento das temáticas evidenciadas nesta pesquisa, sendo aplicável também a outras áreas do conhecimento na sistematização de percursos metodológicos que objetivam estudar domínios específicos.

AGRADECIMENTOS

Agradecimentos ao apoio das Agências de fomento: CAPES, FAPEMIG e CNPq/Brasil.

REFERÊNCIAS BIBLIOGRÁFICAS

Barbosa, A. P. (1969). *Teoria e prática dos sistemas de classificação bibliográfica*. Rio de Janeiro, RJ, Brasil: IBBD.

Best, J. W. (1972). *Como investigar en educación*. (2. ed. cap. 1). Madrid, Espanha: Morata.

Borges, G. S. B., Firmino, H. N. M., Netto, C. M., & Lima, G. A. (2017). *Revisão sistemática baseada em Procedimento de Pesquisa Bibliográfica Estruturado – PPBE*. Documento não publicado. Belo Horizonte, MG, Brasil. No prelo.

Davenport, T. H. (2000). Arquitetura da Informação. In Davenport, T. H. (Ed.), *Ecologia da informação: por que só a tecnologia não basta para o sucesso na era da informação* (pp. 200 – 223). São Paulo, SP, Brasil: Futura.

Garfield, E. (1956). Citation indexes: new paths to scientific knowledge. *Chem Bull*, 43(4), 11-2.

ISI Web of Knowledge. (2009). *Journal Citation Reports*. Recuperado de <http://admin-apps.isiknowledge.com/JCR/JCR>.

Gil, A. C. (2006). *Como elaborar projetos de pesquisa*. (4. ed.) São Paulo, SP, Brasil: Atlas.

Greenhalgh, T. (1997). Papers that summarize other papers (systematic review and meta-analyses). *British Medical Journal*, London, 315(7109), 672-675.

- Lima, G. Â. (2004). *Mapa Hipertextual (MHTX): um modelo para organização hipertextual de documentos*. (Tese de Doutorado não publicada). Universidade Federal de Minas Gerais, Brasil.
- Lima, G. Â. (2016). Arquitetura da informação. In Miranda, R. C. R. (Ed.). *Arquitetura da informação na Câmara dos Deputados*. Brasília, DF, Brasil: Câmara dos Deputados. pp. 47 - 62. (Série Retatos de Pesquisa).
- Lima, G. A. B. (2002). A análise facetada na modelagem conceitual de sistemas de hipertexto: uma revisão de literatura. *Perspectiva em Ciência da Informação*, 7(2), 189-196. Recuperado de portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/407/219
- Maculan, B. C. M. S. (2014). *Taxonomia facetada e navegacional: um mecanismo de recuperação*. Curitiba, PR, Brasil: Appris.
- Morville, P.; Rosenfeld, L. (2006). *Information architecture for the world wide web*. O'Reilly Media.
- Naves, M. M. L. (2017). *Manual para elaboração de monografias*. Atualização, notas e apêndices de Borges, G. S. B. 2. ed. Belo Horizonte, MG, Brasil: IMEDE.
- Qualis-Periódicos. (2017). Recuperado de <https://sucupira.capes.gov.br/sucupira/>
- Ranganathan, S. R. (1967). *Prolegomena to library classification*. (3. ed.). London: Asia Publishing House. Recuperado de <http://arizona.openrepository.com/arizona/handle/10150/105067>
- Rother, E. T. (2007). Revisão sistemática x revisão narrativa. *Acta Paulista de Enfermagem*, 20(2), v-vi.
- Salvador, Â. D. (1980). *Métodos e técnicas de pesquisa bibliográfica: elaboração de trabalhos científicos*. (8. Ed). Porto Alegre, Brasil: Sulina.
- Silva, M. B., & Miranda, Z. D. (2016). A formação de assuntos na teoria da classificação facetada de Ranganathan: uma análise conceitual. In Lucas, E. R. O, Corrêa, E. C., & Eggert-Steindel, D. G. (Orgs.). *As contribuições de Ranganathan para a biblioteconomia: reflexões e desafios* (pp. 72-83). São Paulo, SP, Brasil: FEBAB.
- Souza, R. F. (2016). Ranganathan e a Classificação dos Dois Pontos. In Lucas, E. R. O, Corrêa, E. C., & Eggert-Steindel, D. G. (Orgs.). *As contribuições de Ranganathan para a biblioteconomia: reflexões e desafios* (pp. 84-95). São Paulo, SP, Brasil: FEBAB.
- Speziali, P. (1973). Classifications of the sciences. In *Dictionary of the History of Ideas* (pp. 462-467). New York: Scribners
- Testoni, G. (2015). *Revisão sistemática e metanálise FOP UNICAMP*. Vídeo. (9:33 min). Recuperado de <https://www.youtube.com/watch?v=3DiYW6tM7kY>
- Tomaél, M. I. et al. (2004). Critérios de qualidade para avaliar fontes de informação na Internet. In Tomaél, M. I., & Valentim, M. L. P. (Orgs.). *Avaliação de fontes de informação na Internet* (pp. 19-40). Londrina, Brasil: Eduel.
- Vickery, B. C. (1980). *Classificação e indexação nas ciências*. Rio de Janeiro, RJ, Brasil: BNG/Brasilart.

Wurman, R. S. (2000). *Information anxiety 2*. Indianópolis, Indiana: QUE. 307.



O ATO LEITOR NO PROCESSO DE INDEXAÇÃO: UMA ABORDAGEM COGNITIVA

Ana Lúcia Terra

Instituto Politécnico do Porto; CIC.Digital – Porto, anatterra@iscap.ipp.pt

RESUMO Este trabalho incide numa análise do ato leitor que ocorre durante o processo de indexação a partir das operações cognitivas identificada no modelo de compreensão da leitura definido por Giasson (1993). Procura dar-se um contributo para aprofundar a reflexão e o conhecimento sobre a leitura documental. Em termos metodológicos, o trabalho fundamenta-se numa revisão crítica da literatura sobre o ato leitor no processo de indexação. De seguida, são analisados, na ótica da leitura documental, os cinco tipos de processos que ocorrem durante a leitura/compreensão de um texto, conforme definido por Giasson, a saber: 1) microprocessos, 2) integração, 3) macroprocessos, 4) elaboração e 5) metacognição.

PALAVRAS-CHAVE *Ato leitor, Indexação, Modelo de leitura, Processos cognitivos.*

ABSTRACT Using the reading comprehension framework created by Giasson (1993) to identify reading cognitive operations, this paper presents an analysis of the reading act that occurs on the indexing process. It seeks to contribute to deepen reflection and knowledge about documentary reading. In methodological terms, the work is based on a critical review of the literature on the reader act in the indexing process. Then, the five types of processes that occur during the reading and comprehension of a text, as defined by Giasson, are analysed, namely: 1) microprocesses, 2) integration, 3) macroprocesses, 4) elaboration and 5) metacognition.

KEYWORDS *Reading act, Indexing, Reading framework, Cognitive processes*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

1. INTRODUÇÃO

Tal como explicitado por vários autores, na análise de assunto enquadrada no processo de indexação, o ato leitor realizado pelo indexador é determinante para criar produtos de qualidade resultantes da indexação (Fujita, 1999; Fujita, Rubi, 2006; Mai, 2001; Moura, 2004). Contudo, apesar da sua relevância e das especificidades que apresenta, as características da leitura realizada pelo indexador carecem de um conhecimento mais aprofundado (Hudon, 2013; Anderson, Pérez-Carballo, 2011).

A leitura é uma atividade complexa e plural em grande parte devido ao facto de se tratar de uma comunicação diferida, já que o autor e o leitor não são, por norma, coincidentes no espaço e no tempo. O texto apresenta-se ao leitor fora do seu contexto original, pelo que alguns autores (Jouve, 1993) postulam que a “descontextualização” da mensagem escrita sustenta múltiplas interpretações, considerando que cada leitor traz consigo a sua experiência, a sua cultura e os seus valores pessoais bem como os do seu tempo. Esta perspetiva é especialmente crítica quando se considera que o indexador

é um leitor-mediador (Redigolo & Fujita, 2015) em que o resultado da sua leitura irá possibilitar aos utilizadores dos serviços de informação o acesso aos documentos através de representações do seu conteúdo intelectual. Assim, na leitura documental, considerada como uma prática discursiva, o universo do discurso do indexador não é coincidente nem com o do autor da obra, nem com os dos utilizadores do serviço de informação, tal como notou Michel Le Guern, no prefácio da obra de Muriel Amar (2000). Acrescenta-se, portanto, um grau superior de complexidade a esta leitura profissional que pode ser perspectivada no âmbito da comunicação interativa e da leitura “partilhada”, com ênfase na intersubjetividade. Este último conceito é especialmente relevante porque «l’intersubjectivité permet ainsi de prendre en compte à la fois le caractère “subjectif” de la lecture, c’est-à-dire l’espace de débordement reconnu à chaque lecteur du fait de sa singularité, et le souci d’une saisie “objective” du contenu du texte du fait du partage communicationnel attendu» (Fondin, 2006, 27).

Assim, o leitor-indexador posiciona-se como um leitor-analista, conforme a classificação de Thérien (1990), podendo ainda apresentar-se como um leitor-especialista. É um leitor-analista na medida em que lê num duplo registo: um permite-lhe conhecer o texto e o outro habilita-o a dar a conhecer o texto, no seu todo ou em algumas partes, num quadro mais abrangente que lhe serve de princípio de análise. Com efeito, o leitor-indexador lê o texto para captar o seu conteúdo e dá o texto a conhecer aos utilizadores do serviço de informação ao qual está vinculado, realizando a sua leitura no enquadramento específico que lhe é dado pelo âmbito desse serviço. Será ainda um leitor-especializado se apresentar conhecimentos científicos próprios de áreas delimitadas.

Além disso, o leitor-indexador não lê para ele mas para os outros, não escolhe o que lê e não lê os textos integralmente (Waller, 1999). Por seu lado, Farrow (1995) sublinha que a leitura do indexador é determinada por limitações de tempo, pelo sobrevoar dos textos no sentido de identificar pistas que permitam compreender as ideias essenciais, pela realização de uma tarefa profissional (não é uma leitura de aprendizagem) e pela criação imediata de produtos resultantes dessa leitura.

Fujita (1999) acrescenta também que os indexadores estão sujeitos a condições específicas de leitura, nomeadamente em termos de limite de tempo, de propósito definido, de produção de resultados, de vinculação a um sistema documental comprometido com as necessidades de pesquisa dos seus utilizadores, os quais irão condicionar os seus procedimentos de leitura. Neste sentido, precisam de realizar uma leitura bem estruturada e seguir um método sistematizado que apoie a identificação dos conceitos do documento, permitindo a compreensão da temática, e o reconhecimento dos termos representativos equivalentes usando a linguagem documental em uso no serviço de informação. A finalidade específica da leitura realizada pelo leitor-indexador é igualmente sublinhada por Moura (2004, p. 164) que postula uma “leitura técnica” consistindo «(...) na abordagem global dos itens informacionais, tem por objetivo recolher os dados que permitirão o estabelecimento da representação desses itens nos sistemas de informação» e que procura «(...) a reconstituição bruta da informação veiculada no texto original». É neste contexto funcional que a leitura documental é efetivada com objetivos bem delimitados mas sem uma explicitação consciente de como é realizada por parte do indexador, nomeadamente ao nível dos processos e das operações cognitivas em ação.

2. METODOLOGIA

Considerando as premissas acima explicitadas, apresentam-se algumas reflexões e respostas para a pergunta “que leitura faz o profissional da informação quando exerce funções de indexação?”.

Em termos metodológicos, o trabalho foi desenvolvido com base numa revisão crítica da literatura relevante de modo a situar o ato leitor no contexto da indexação. Este enquadramento serviu de pano de fundo para a aplicação, à leitura documental realizada durante o processo de indexação, do modelo de compreensão da leitura de Giasson (1993), especificamente no que toca aos processos cognitivos envolvidos e às suas componentes.

Note-se que o modelo de leitura de Giasson (Leitor-Texto-Contexto) tem guiado alguns estudos (cf. Fujita & Nardi, 1998, entre outros) sobre a leitura documental mas a sua tipificação das operações cognitivas que ocorrem durante o ato leitor não tem sido explorada neste âmbito. Assim, a presente abordagem constitui-se como inovadora e relevante para a teoria e para a prática, na medida em que visa aprofundar a reflexão e o conhecimento sobre a leitura documental, podendo sustentar no futuro estudos de ordem empírica no sentido de conceber uma metodologia e estratégias de leitura documental otimizadas.

3. PROCESSOS DE LEITURA DOCUMENTAL

De acordo com Giasson (1993), do ponto de vista do leitor, a leitura/compreensão de um texto envolve cinco tipos de processos: 1) microprocessos, 2) processos de integração, 3) macroprocessos, 4) processos de elaboração e 5) processos de metacognição. Os microprocessos sustentam a compreensão da informação contida numa frase. Os processos de integração servem para criar ligações entre as proposições ou as frases, pois centram-se na procura da coerência entre as frases. Os macroprocessos estão orientados para a compreensão global do texto e, nesse sentido, focalizam-se na compreensão da coerência entre as frases. Os processos de elaboração permitem ao leitor ir além do texto. Os processos metacognitivos definem a compreensão e permitem ao leitor alinhar-se com o texto ou situação.

De seguida, iremos abordar esta tipologia de processos cognitivos no contexto específico do ato leitor que ocorre quando da indexação.

MICROPROCESSOS

Os microprocessos envolvem competências ao nível da identificação das palavras, da leitura de grupos de palavras e da microseleção. Quando a identificação das palavras é automática, a atenção do leitor pode ficar disponível para outros processos, além de tornar a leitura mais rápida.

Para o trabalho de análise documental estão especialmente convocadas as duas últimas. De facto, a leitura por grupos de palavras implica a utilização de pistas sintáticas para identificar nas frases os elementos significativos de sentido, sendo que a compreensão assenta na passagem da memória de curta duração, onde são tratados os conjuntos de palavras, para a memória de longa duração, onde são armazenados. Já a microseleção consiste em decidir e memorizar a informação pertinente numa frase. Na leitura de um texto, uma parte da informação das frases lidas deve ser memorizada pela memória de curta duração, de modo a que a nova informação seja integrada na que já existe. Como o indexador tem

de fazer uma leitura seletiva, pois não é viável a leitura integral dos documentos, a leitura por grupos de palavras e a microseleção são muito relevantes. A primeira aumenta a cadência da leitura, poupando tempo e esforço cognitivo, enquanto a segunda proporciona a focalização da atenção e da memória nas componentes sintáticas e semânticas mais relevantes para extrair o(s) assunto(s) pertinentes no documento. Aliás, a literatura sobre análise de assunto sublinha que não é necessário, nem aconselhável realizar uma leitura linear, palavra por palavra, pois o leitor deve progredir no texto conforme lhe for possível antever o seguimento das ideias (Fujita & Nardi, 1998). Estará assim, a convocar a capacidade de previsão umas das competências dos processos de integração.

PROCESSOS DE INTEGRAÇÃO

Os processos de integração permitem compreender as pistas explícitas, que esclarecem a relação entre as frases (são os referentes ou anáforas e os conectores), e as relações implícitas entre as frases e as ideias, através de inferências baseadas no texto ou nas referências do leitor. As inferências permitem adivinhar o implícito e levam a uma compreensão do texto que ultrapassa o nível literal, pressupondo que o leitor recorre ao seu conhecimento pré-existente para conferir sentido a uma mensagem. Quanto mais intensa e profunda for a experiência profissional do indexador, maior será a sua capacidade de realizar inferências, podendo antecipar, ponderar ou relativizar a importância dos textos, de identificar aspetos inovadores e de seleccionar a informação útil (Waller, 1999). As inferências podem ser lógicas ou pragmáticas.

No contexto da análise documental, quanto melhor o indexador identificar as pistas explícitas mais rapidamente poderá centrar-se nas partes essenciais do texto para entender o seu sentido. Note-se ainda que a frequência da ocorrência de referentes ou anáforas ajudará a determinar o assunto do documento, sendo aliás um critério fundamental na indexação automática baseada na contagem de palavras. Contudo, este indício quantitativo não invalida a necessidade de uma aferição qualitativa do significado da ocorrência dessas palavras e dos seus equivalentes para determinar com exatidão a temática do item informativo.

No que respeita às relações implícitas, o indexador precisará de usar inferências lógicas de modo a compreender mais profundamente o sentido do texto e também irá recorrer a inferências pragmáticas, sendo que nestas últimas deve confirmar as suas suposições através de uma análise cuidada do documento objeto do processo de tratamento documental. Note-se ainda que, quanto maior for a sua experiência e o seu conhecimento das temáticas, mais inferências pragmáticas fará o indexador. Contudo, isto também aumentará a subjectividade já que este tipo de inferências está eminentemente ligado às características cognitivas e vivências de cada indivíduo.

MACROPROCESSOS

Os macroprocessos incidem sobre a apreensão global do texto, desvendando as ligações que o tornam um todo coerente. Estes macroprocessos incluem a identificação da ideia principal, a elaboração do resumo e a apreensão/utilização da estrutura do texto, todos essenciais no trabalho de tratamento documental.

Existe uma grande diversidade no que respeita à noção de ideia principal de um texto mas podem distinguir-se duas categorias fundamentais: a) a informação é importante porque o autor a apresenta

como tal, e nesse caso é uma informação textualmente importante; b) a informação é importante porque o leitor a considera como tal, tendo em conta a intenção da sua leitura, pelo que nessa situação a informação é contextualmente importante. No caso do indexador, importa que este foque a sua atenção na informação textualmente importante ainda que deva ter em conta a informação contextualmente relevante em função das necessidades dos utilizadores para os quais está a realizar o tratamento documental. Note-se ainda que pode ser relevante distinguir a temática do texto e a ideia principal do texto, a qual pode ser explícita, implícita (devendo ser inferida) ou ambígua (não estando claramente exposta).

A elaboração do resumo implica uma maior dificuldade de realização porque requer seleção e hierarquização para eliminar os elementos secundários e redundantes. O resumo consiste em reescrever um texto pré-existente, mantendo a informação importante, havendo equivalência informativa, através de uma economia de meios e a adaptação a uma nova situação de comunicação.

A estrutura do texto condiciona a organização das ideias e cada tipo de texto (narrativo, informativo, científico, etc) apresenta uma estrutura própria. O leitor precisa de apreender e utilizar essa estrutura do texto para compreender e memorizar a informação. No fundo, trata-se de formas de organização convencionadas que os leitores “armazenam” na sua mente e que ativam ao ler os textos de forma automática. Relativamente à leitura documental, conforme sublinhado por Fujita (1999, p. 109) «Considera-se, ainda, que o leitor com facilidade de reconhecer as superestruturas textuais capta melhor as ideias principais do texto, do que um leitor que lê linearmente. Apoiando-se na estrutura textual e em seus conhecimentos prévios o leitor infere significados e levanta hipóteses que o ajudarão a apreender a temática global». Esta perspetiva é igualmente sublinhada por Naves (2001) ao enfatizar que, para realizar a leitura com fins específicos, é necessário que o indexador conheça as estruturas de diferentes tipos de textos. Para Waller (1999) esta apreensão da estrutura do texto faz-se *antes* da leitura e focaliza-se nos elementos de apresentação do texto, como os caracteres tipográficos, a disposição do texto (organização dos capítulos, títulos, notas de rodapé, etc), as ilustrações (imagens, gráficos, legendas, etc), o uso de parágrafos e de alíneas, além da pontuação.

PROCESSOS DE ELABORAÇÃO

Os processos de elaboração permitem ao leitor ir além do texto, criando inferências não previstas pelo autor. Estes processos incluem cinco componentes: previsões, imagens mentais, respostas afetivas, ligações com os conhecimentos e raciocínio

As previsões dizem respeito às ideias e incidem sobre as frases, e não sobre as palavras, podendo ser relativas ao conteúdo do texto (o que é expectável encontrar no texto) ou à sua estrutura (o que já se sabe sobre o tipo de texto em análise). As imagens mentais podem ocorrer de forma espontânea durante a leitura mas a produção consciente de imagens mentais favorece a compreensão leitora. As imagens mentais consistem em criar na mente uma representação do que se lê. Esta atividade aumenta a capacidade de memória de trabalho durante a leitura, ao reunir os detalhes em grandes conjuntos, facilita a criação de analogias e de comparações, serve para estruturar e memorizar a informação, aumenta o grau de comprometimento com o texto e melhora o interesse pela leitura.

As respostas afetivas incidem sobre as emoções convocadas pela leitura, sendo que um leitor envolvido emocionalmente com uma obra irá compreendê-la melhor. Contudo, o envolvimento afetivo pode suscitar erros levando a enfatizar aspetos que não são realmente importantes. Assim, as reações

emotivas nem sempre são adequadas, dependendo do texto. A ligação com os conhecimentos significa que o leitor estabelece uma relação entre o texto e os seus conhecimentos pessoais, os quais não sendo indispensáveis à compreensão do texto, e não se afastando muito dele, podem facilitar a memorização da informação. Por fim, nos processos de elaboração estão incluídas as competências de raciocínio. De facto, compreender não é suficiente pois é necessário que os leitores formulem um julgamento crítico sobre o texto, o que implicará distinguir os factos das opiniões, conhecer as fontes de informação e estar sensibilizado para os “preconceitos” do autor.

No que respeita aos processos de elaboração no decurso da análise documental, se as previsões ajudam à compreensão do texto, elas devem basear-se em pistas, motivadas por factos ou pela estrutura do texto, e nessa medida são previsões e não procedimentos de adivinhação/imaginação. Contudo, o indexador deve usar com cuidado estas pistas, nomeadamente, no que toca ao título, o qual muitas vezes, por razões de marketing editorial, entre outras, traduz uma ideia pouco precisa do conteúdo que identifica. Por sua vez, a produção de imagens mentais implica que o leitor esteja atento ao texto e favorece a sua compreensão, apoiando a estruturação e a memorização da informação, além de facilitar a criação de analogias e de comparações, operações essenciais ao indexador não só para a apreensão do texto mas igualmente para as fases subseqüentes à análise. Com efeito, a criação de analogias e de comparações serão relevantes para o momento em que os assuntos serão traduzidos em termos da linguagem documental em uso na instituição.

A integração com os conhecimentos prévios do indexador é também uma operação fundamental para o procedimento de análise documental, sendo certo que a experiência em determinada área temática irá facilitar a realização de um trabalho de qualidade, apoiando a compreensão do conteúdo informativo do texto porque o leitor já possui referentes que o ajudam a enquadrar-se no assunto (Redigolo & Fujita, 2015). Assim, é mais fácil compreender um texto produzido no âmbito da mesma matriz cultural ou numa mesma área científica já que os referentes são os mesmos. Com efeito, quanto mais informação um sujeito tiver sobre uma temática mais apto estará para compreender, memorizar e produzir inferências a partir da sua leitura. Ainda assim, também é verdade que os conhecimentos prévios podem afastar o leitor dos aspetos essenciais do texto, desviando o objetivo da leitura. Por outro lado, é também essencial ter cautela com os conhecimentos erróneos pois a natureza do conhecimento prévio afeta a aquisição de novos conhecimentos. Neste sentido, se o indexador tiver conhecimentos pouco precisos ou errados sobre uma matéria terá tendência a distorcer o que ler, de modo a adequar esse conteúdo aos seus quadros mentais prévios, mesmo que estejam errados. O raciocínio é igualmente uma operação nuclear no processo de análise documental, sustentando a distinção entre factos e opiniões ou identificando os aspetos denotativos da linguagem (apreensão do sentido literal) e os aspetos conotativos da linguagem (apreensão do sentido figurado da linguagem).

PROCESSOS METACOGNITIVOS

Por fim, os processos metacognitivos gerem a compreensão e permitem ao leitor adaptar-se ao texto e à situação/contexto em que decorre a leitura. A cognição envolve a compreensão, a memorização e o tratamento da informação. Já a metacognição diz respeito ao conhecimento que o indivíduo tem sobre o seu próprio funcionamento cognitivo e as estratégias para controlar esse processo.

As competências metacognitivas desdobram-se em duas vertentes: na autoavaliação e na autorregulação. Pela autoavaliação o leitor torna-se ciente das suas capacidades e das limitações cognitivas, dos seus interesses, motivações, bem como das exigências da leitura e de estratégias adequadas para ultrapassar

problemas de leitura. Trata-se, portanto, de identificar perdas da compreensão. Com a autorregulação o leitor consegue resolver a perda de compreensão pois identifica quando compreende e quando não compreende, sabe o que compreende ou não e aquilo de que necessita para compreender e sabe que pode seguir outras estratégias para compreender. Assim, o leitor consegue aferir se compreende bem, posicionando-se numa perspectiva de autogestão da compreensão, identificando e resolvendo as lacunas de compreensão.

Na opinião de Fujita (1999, 108) «(...) a leitura para fins documentários exige que o leitor-indexador seja metacognitivo por ser necessário usar estratégias de interação com o texto de forma a propiciar uma compreensão suficiente para a sua representação». Mais recentemente, Redigolo e Fujita (2015) trabalharam o levantamento de estratégias metacognitivas através a aplicação do protocolo verbal a profissionais da informação dedicados à análise de assunto em monografias em bibliotecas universitárias. Neste contexto, concluíram que «as estratégias metacognitivas demonstram que o leitor tem consciência dos objectivos da leitura e utiliza-se de alguns artificios para realiza-la da melhor maneira, tentando superar suas dificuldades e explorando conhecimentos» (Redigolo & Fujita, 2015, p. 365).

CONCLUSÕES

Ainda que o modelo de leitura de Giasson (Leitor-Texto-Contexto) não tenha sido desenvolvido especificamente para explicar os processos de leitura documental, tem guiado alguns estudos nesta matéria (cf. Fujita & Nardi, 1998 entre outros). Contudo, até agora, não parece ter sido aplicada a sua tipificação dos processos cognitivos que ocorrem durante o ato leitor no contexto particular da leitura documental. Neste trabalho, expusemos como a aplicação dessa tipificação dos processos cognitivos e das suas componentes/operações pode ser bastante proveitosa em termos de compreensão dos processos cognitivos apresentados pelo indexador. Esta linha de análise precisa ainda de ser aprofundada em termos teóricos e com aplicações de índole prática mas a identificação dos processos feita por Giasson parece adequada para compreender os processo cognitivos que ocorrem durante a leitura documental, apesar das suas particularidades.

REFERÊNCIAS BIBLIOGRÁFICAS

- Amar, M. (2000). *Les fondements théoriques de l'indexation: une approche linguistique*. Paris: ADBS Éditions.
- Anderson, J. D., & Pérez-Carballo, J. (2001). Nature of indexing: how humans and machines analyze messages and texts for retrieval. Part I: research, and the nature of human indexing. *Information Processing and Management*, 37(2), 231–254.
- Farrow, J. (1995). All in the mind: concept analysis in indexing. *Indexer*, 19(4), 243–245.
- Fondin, H. (2006). La problématique de la lecture documentaire du texte informationnel. *Documentation et Bibliothèques*, Jan-Mar, 17–28.
- Fujita, M. S. L. (1999). A leitura do indexador: estudo de observação. *Perspectivas em Ciência da Informação*, 4(1), 101–116.

- Fujita, M. S. L., & Rubi, M. P. (2006). Um modelo de leitura documentária para a indexação de artigos científicos: princípios de elaboração e uso para a formação de indexadores. *DataGramaZero- Revista de Ciência da Informação*, 7(3).
- Fujita, M. S., & Nardi, M. I. A. (1998). A leitura em análise documentária. *Transinformação*, 10(3), 13–31.
- Giasson, J. (1993). *A compreensão na leitura*. Porto: Edições ASA.
- Hudon, M. (2013). *Analyse et représentation documentaires: introduction à l'indexation, à la classification et à la condensation des documents*. Québec: Presses de l'Université du Québec.
- Jouve, V. (1993). *La lecture*. Paris: Hachette.
- Mai, J.-E. (2001). Semiotics and indexing: an analysis of the subject indexing process. *Journal of Documentation*, 57(5), 591–622.
- Moura, M. A. (2004). Leitor-bibliotecário: interpretação, memória e as contradições da subjetividade. *Perspectivas em Ciência da Informação*, 9(2), 158–169.
- Naves, M. M. L. (2001). Estudo de fatores interferentes no processo de análise de assunto. *Perspectivas em Ciência da Informação*, 6(2), 189–203.
- Redigolo, F. M., & Fujita, M. S. L. (2015). A leitura profissional do catalogador e seu papel como mediadora da informação. *Informação & Informação*, 20(3), 356–376.
- Thérien, G. (1990). Pour une sémiotique de la lecture. *Protée: Théories et Pratiques Sémiotiques*, 18(2), 67–80.
- Waller, S. (1999). *L'analyse documentaire: une approche methodologique*. Paris: ADBS Éditions.



TRATAMENTO INFORMACIONAL DA IMAGEM BIOMÉDICA: ESTADO DA ARTE EM PERIÓDICOS DE CIÊNCIA DA INFORMAÇÃO

Joice Cleide Cardoso Ennes de Souza

Universidade Federal Fluminense – UFF, 0000-0002-3848-4923, joicecardoso@id.uff.br

RESUMO Introdução: Na área biomédica, as imagens de exames médicos auxiliam na tomada de decisão de diagnósticos clínicos. Com sua disseminação, constata-se o grande volume de imagens produzidas e o surgimento de questionamentos em como organizar e representar a informação imagética com vistas à recuperação. **Objetivos:** Identificar as bases teórico-metodológicas adotadas na representação da imagem no contexto da saúde em periódicos científicos na área da Ciência da Informação. **Métodos:** Consiste em pesquisa qualitativa, de caráter exploratório, com orientação analítico-descritiva, mediante a identificação de bases teórico-metodológicas para representação de imagem biomédica. Por oferecer rigor científico e profundidade no exame dos textos, adotou-se a análise de conteúdo para analisar os artigos recuperados no levantamento bibliográfico. **Resultados/conclusões:** Infere-se que a representação da imagem médica compreende descrição das regiões e como se apresentam, e interpretação do que é visualizado pelo profissional médico. Destaca-se o modelo conceitual piramidal de Jaimes e Chang (1999) como método de descrição de imagens médicas.

PALAVRAS-CHAVE *Tratamento informacional, Imagem biomédica, Periódicos científicos, Ciência da Informação, Ciências da Saúde.*

ABSTRACT Introduction: In the biomedical area, images produced for medical examinations assist in establishing medical diagnoses. Due to image dissemination, it's evident the great number of images produced nowadays, which provokes questions on how to organize and represent image information, seeking its eventual retrieval. **Purposes:** To identify the theoretical and methodological approach for representation of images in the context of health in journals in the area of Information Science. **Methodology:** Qualitative research, exploratory, descriptive-analytical oriented, by identifying theoretical and methodological approach for biomedical image representation. The articles retrieved in the bibliographic survey was analysed. Content analysis was adopted because it offered scientific rigor and depth in examining the texts. **Results/Conclusions:** It's possible to infer that medical image representation consists of describing the body parts and their current aspects, and of its analysis by a medical professional. We concluded that the conceptual pyramidal model by Jaimes and Chang (1999) stand out as the preferred method to describe medical images.

KEYWORDS *Information Gathering, Biomedical image, Journals, Information Science, Health Sciences.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Na atualidade, dissemina-se a utilização de imagens nas diferentes atividades cotidianas. Por apresentar natureza imediatista, informativa e ilustrativa, as imagens biomédicas são usadas tanto no contexto educacional, como na área clínica. Na área biomédica, as imagens produzidas (radiologias, tomografias, etc.) auxiliam na tomada de decisão de diagnósticos clínicos. No âmbito educacional do profissional da área da saúde, as imagens são recursos didáticos que complementam o processo de aprendizagem iniciado nas aulas laboratoriais. A quantidade de imagens provoca questionamentos em como organizar e representar a informação imagética com vistas a sua recuperação.

A origem da imagem médica remonta a 1895, com a descoberta da radiografia por Wilhelm Conrad Roentgen (Wong & Hoo Jr., 2002, p. 83). Desde então, criaram-se diferentes exames de imagem, como eletrocardiograma (ECG), mamografia, entre outros. As imagens possibilitam o estudo do corpo humano e a definição de diagnósticos clínicos.

Alguns exames produzem grande número de imagens, como a tomografia computadorizada. Os diferentes tipos de exames exigem que as imagens sejam capturadas, descritas, visualizadas, armazenadas, recuperadas, permitindo que sejam posteriormente consultadas, como objetiva os sistemas de comunicação e arquivamento de imagens ou PACS (picture archiving and communication system).

Consultando os bancos de imagens médicas, os profissionais médicos acessam e interpretam as imagens para fundamentar diagnósticos clínicos. Portanto, as imagens devem oferecer metadados e descrição para serem recuperadas. A especialidade médica e o órgão foco do exame podem gerar um número diferenciado de imagens que exige um tratamento informacional específico, fazendo-nos refletir sobre novos procedimentos para a gestão do acervo de imagens biomédicas. Segundo Wong e Hoo Jr. (2002), a imagem biomédica possui uma particularidade por representar informações funcionais de partes do corpo humano, além da estrutura e conteúdo anatômico.

Diante disso, nos perguntamos: quais são os princípios teórico-metodológicos observados na literatura em Ciência da Informação relacionados ao tratamento informacional de imagens biomédicas, levando em consideração a mediação da informação no processo comunicacional no âmbito da saúde? A partir da revisão de literatura, objetivamos identificar as bases teórico-metodológicas adotadas na representação da imagem no contexto da saúde, em periódicos científicos na área da Ciência da Informação.

A significação do conceito de representação coloca-se entre apresentação e imaginação, estendendo-se aos conceitos semióticos de signo, veículo do signo, imagem, significação e referência (Santaella & Nöth, 1998, p. 16). Quando se representa a informação de um documento, a representação manifesta a primeira significação do assunto contida na imagem, por exemplo. Dessa representação surgirão desdobramentos, ampliando o escopo do analista indexador e do profissional que recupera a informação.

Sobre a interdisciplinaridade entre a Ciência da Informação e a Ciências da Saúde, Silva (2013) explica que "é muito recente, além do que ocorre em questões muito particulares". O binômio informação e saúde é observado nos serviços de informação em saúde, com destaque para a satisfação das necessidades dos usuários, sejam profissionais médicos ou pacientes. O autor relata outros aspectos contemplados na interdisciplinaridade, como caracterização do usuário dos serviços, competência em

informação na área de saúde, e políticas de informação científica e tecnológica no âmbito da saúde. Contudo, Silva não pontua o tratamento informacional da documentação gerada na área da saúde.

Define-se imagem biomédica como representação visual que assume o papel de ícone por apresentar semelhança com a realidade, de forma a proporcionar qualidade representativa para o usuário da área de saúde no seu desempenho profissional. Biomedicina é a ciência voltada para os estudos e pesquisas de doenças humanas, suas causas e meios de tratá-las. As imagens biomédicas auxiliam na definição dos diagnósticos patológicos e clínicos, sendo aplicadas como recurso didático na formação do profissional da área de saúde.

Enquanto documentos, as imagens biomédicas compõem o prontuário médico do paciente, definido pelo art. 1º da Resolução nº 1638/2002 do Conselho Federal de Medicina (CFM) como

[...] o documento único constituído de um conjunto de informações, sinais e **imagens registradas**, geradas a partir de fatos, acontecimentos e situações sobre a saúde do paciente e a assistência a ele prestada, de caráter legal, sigiloso e científico, que possibilita a comunicação entre membros da equipe multiprofissional e a continuidade da assistência prestada ao indivíduo (Brasil, 2002, grifo nosso).

Apesar dos estudos qualitativos sobre representação de imagens biomédicas no âmbito da Ciência da Informação serem escassos, destacamos: Bentes Pinto (2008); Bentes Pinto e Ferreira (2010); Simpson et al. (2014); Souza, Souza e Almeida (2015); Souza e Paes-De-Almeida (2014). Na área das Ciências da Saúde, registramos Ehsani et al. (2008), Smith, Barnes e Chiosea (2011).

Entende-se como tratamento informacional os procedimentos que visam a organização e representação do documento, objetivando sua recuperação. Para organizar, é necessário ter princípios de análise do documento de modo a representar a informação. Valle Gastaminza (2002) esclarece que, para analisar a imagem fotográfica como objeto de uma dada coleção, é preciso definir seus atributos e estudá-los. Para o autor, atributo de uma imagem é "qualquer tipo de característica, componente ou propriedade do objeto que possa se representado em um sistema de processamento de informação" (Valle Gastaminza, 2002, tradução nossa). Os atributos não se restringem às características visuais, compreendendo também as cognitivas, afetivas, interpretativas, espaciais, semânticas, simbólicas ou emocionais.

METODOLOGIA

Para fundamentar o estudo dos artigos recuperados, adotou-se a análise de conteúdo por nos oferecer rigor científico e profundidade no exame dos textos. Ao proporcionar as técnicas necessárias para analisar criticamente os artigos, tal metodologia minimiza a subjetividade observada no processo comunicativo.

Segundo Bardin (2006, p. 42), a análise de conteúdo é

um conjunto de técnicas de análise das comunicações, visando obter, por procedimentos, sistemáticos e objetivos de descrição do conteúdo das mensagens, indicadores (quantitativos ou não) que permitam a inferência de conhecimentos relativos às condições de produção/recepção (variáveis inferidas) destas mensagens.

Bardin (2006, p. 95) nos apresenta três etapas a serem seguidas: pré-análise, exploração do material e tratamento dos resultados, inferência e interpretação.

Pré-análise é a etapa da organização, em que são tomadas decisões operacionais como a sistematização de ideias e a constituição do corpus da pesquisa. Nesta etapa selecionam-se os documentos, formulam-se as hipóteses e os objetivos a serem alcançados, além de definir os indicadores a serem adotados na análise. Esta fase compreende as atividades: leitura flutuante (primeiro contato com os textos), escolha dos documentos com base nas regras de exaustividade (abrangência do tema dentro do recorte proposto), representatividade (amostras que representem o universo pesquisado), homogeneidade (dados sobre um mesmo tema), pertinência (os documentos devem se referir ao tema pesquisado) (Bardin, 2006).

A segunda etapa consiste na exploração do material a partir da codificação, classificação e categorização. É o momento em que se definem as categorias a serem usadas e a identificação das unidades de registro e das unidades de contexto nos textos. É a etapa decisiva para interpretação e inferências, pois fundamenta-se na descrição analítica do corpus, a partir do marco teórico e das hipóteses propostas (Bardin, 2006).

A codificação, segundo Bardin (2006, p. 103), é o tratamento do material a ser analisado, sendo necessário selecionar a unidade de registro a ser adotada, ou seja, a unidade a ser codificada. No escopo deste estudo, adotaremos o tema como unidade de registro, sendo identificado a partir do processo de análise de assunto, que tem como “objetivo identificar e selecionar os conceitos que representam a essência de um documento” (Fujita, 2003, p. 85). As unidades de contexto do tema, que são usadas para compreender o sentido de uso da unidade de registro, são obtidas com base na declaração de assunto oriunda da pergunta 'de que trata o artigo?'. A próxima etapa é a categorização das unidades de registro, entendida como o processo de identificar "qualquer propriedade ou medida qualitativa ou quantitativa de uma entidade" (Ranganathan, 1967, p.53, tradução nossa). Constitui no primeiro corte classificatório de um dado domínio. Após a identificação das categorias, classificam-se as unidades de registro de acordo com características ou atributos semelhantes. Por último, houve o tratamento dos resultados, inferência e interpretação com base no marco teórico proposto. Constatou-se o pequeno número de artigos que versam especificamente sobre o tratamento informacional de imagens biomédicas no âmbito da Ciência da Informação.

Para realização do estudo, empreendeu-se revisão sistemática da literatura com pesquisas efetuadas nos periódicos oferecidos pelo Portal Periódicos Capes (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, 2017), restringindo a área de conhecimento por Ciências Sociais e Aplicadas e subárea Ciência da Informação, conforme Quadro 1. Para a seleção dos artigos, adotamos os critérios: publicação no período 2012-2017; artigos publicados em inglês; artigos oferecidos em versão completa.

Quadro 1. Periódicos pesquisados no Portal Periódicos Capes

American Libraries	Annual Meeting Proceedings - ASIS&T
Annual review of Information Science and Technology	Applied Medical Informatics
Biomedical Digital Libraries	Bulletin of the American Society for Information Science
Bulletin of the American Society for Information Science and Technology	Bulletin of the Medical Library Association
eJournal of Health Informatics	Evidence Based Library and Information Practice

Health Care on the Internet	Health Information and Libraries Journal
Health Information Management	Health Libraries Review
IFLA Journal	International Journal of Healthcare Information Systems and Informatics
Journal of Classification, Journal of Digital Information Management	Journal of Electronic Resources in Medical Libraries
Journal of System and Information Technology	Journal of the American Society for Information Science
Journal of the American Society for Information Science and Technology	Journal of the Medical Library Association
Journal of the Society of Archivists, Knowledge Organization	Library Trends
The Indexer	

Realizou-se o levantamento quantitativo nos sites dos periódicos, usando como estratégias de busca: "medical image", "biomedical image", "Image retrieval" and biomedical, "Image retrieval" and medical, "health information" and image, "health information", "health sciences information", "medical visual information".

Os artigos recuperados foram organizados pelos nomes dos periódicos para depois efetuar a pré-análise do conjunto com a leitura do resumo do artigo e posterior seleção daqueles que atendiam ao objeto do presente estudo, compondo a amostra. Selecionamos artigos que focavam: tratamento informacional da imagem biomédica, gestão da documentação médica e instrumentos de representação adotados. Com a leitura do resumo do artigo, obtivemos a declaração de assunto para identificação dos conceitos.

RESULTADOS E DISCUSSÃO

No Quadro 2 verificamos o número expressivo de publicações sobre imagens médicas no Journal of the American Society for Information Science and Technology, seguido pela Knowledge Organization e Bulletin of the American Society for Information Science and Technology.

Quadro 2. Levantamento quantitativo de artigos por periódico

Periódicos	Nº de artigos recuperados
Annual Meeting Proceedings - ASIS&T	1
Applied Medical Informatics	1
Bulletin of the American Society for Information Science and Technology	3
Journal of Digital Information Management	1
Journal of System and Information Technology	1
Journal of the American Society for Information Science and Technology	12

Journal of the Medical Library Association	1
Knowledge Organization	5

Já no Quadro 3, verificamos o número de artigos sobre informação e imagens médicas publicado por ano. Constata-se o crescimento a partir de 2015, revelando o aumento de interesse no assunto.

Quadro 3. Levantamento quantitativo de artigos por ano

Ano	Nº de artigos recuperados
2017	2
2016	7
2015	10
2014	1
2013	3
2012	2
Total	25

Dos artigos recuperados, Boyd (2017) e Beaudoin (2016) não apresentam conteúdo específico da área de saúde. Distinguímos diferentes abordagens em relação à informação médica: sistemas de informação em saúde (McDermott, 2016; Najaftorkaman, Ghapanchi, Talaei-Khoei, & Ray, 2014; Richardson, 2016; Andrade & Lara, 2016); acesso à documentação médica pelo paciente (McNamara et al., 2015; Tamine, Chouquet, & Palmer, 2015; Huvila, Cajander, Daniels, & Ahlfeldt, 2015; Oh, Joo, & Jeong, 2015); uso e qualidade da informação médica (Zhang, Sun, & Xie, 2015; Wolf & Veinot, 2015); papel dos metadados na operacionalização do conhecimento médico e no acesso pelo usuário (Panzer, 2016); investigação de estruturas de produção, organização e comunicação de conhecimento com foco na saúde (Marteleto & Carvalho, 2015); uso de vocabulário padronizado na indexação na área da saúde (Pinto, Rabelo, & Girão, 2014); organização e sistema de gerenciamento de documentos, indexação, classificação de conteúdo médico (Chiaravalloti, Pasceri, & Taverniti, 2012); indexação de assunto de documentos biomédicos (Chebil, Soualmia, Omri, & Darmoni, 2015; Lu & Mao, 2015).

Em relação à imagem biomédica, registramos os assuntos: indexação de imagem biomédica (Wang et al., 2012; Kim, 2013; Apostolova et al., 2013; Splendiani & Ribera, 2016); recuperação de imagens (Wang & Erdelez, 2013; Ayyachamy, 2015).

Wang et al. (2012) examinam quais níveis de atributos de imagens são exigidos pelos usuários de imagens médicas com diferentes níveis de domínio de conhecimento. Segundo os autores, a representação da imagem médica é complexa pela natureza da imagem, por ser tácita e sem palavras, contendo diversos níveis sintáticos e semânticos, para não mencionar o próprio domínio da medicina. Uma amostra de 240 imagens de radiografia foram descritas por usuários da área médica com expertises diferentes, divididos em novatos, intermediários e especialistas. Para a análise das descrições, Wang et al. (2012) desenvolveram um modelo para estruturar as informações visuais, com 10 níveis divididos em sintáticos e semânticos: Sintáticos: Tipo/Técnica (usada na produção da imagem. Ex.: raio-x), Distribuição Global (conteúdo global descrito com termos de nível baixo, como cor), Estrutura Local (componentes individuais da imagem, como os elementos visuais básicos. Ex.: textura), Composição Global (análise da imagem como um todo, usando elementos básicos. Ex.: simetria); Semânticos: Objeto Genérico (descrição geral de objetos, visualização da imagem na sua totalidade. Ex.: perna direita), Cena Genérica (descrição genérica da cena. Na radiologia, é a posição do paciente. Ex.: lateral),

Objeto Específico (nomes dos objetos específicos, exigindo conhecimento específico. Ex.: coluna lombar), Cena Específica (descrição específica da cena por atributos da imagem, exigindo conhecimento específico. Ex.: posição correta do raio central), Objeto Abstrato (conhecimento especializado ou interpretativo sobre os objetos representados, incluindo informações sobre o paciente. Ex.: paciente pediátrico) e Cena Abstrata (o que a imagem representa no todo, incluindo diagnóstico patológico, como fratura).

Como resultado, Wang et al. (2012) constataram que o conhecimento de um dado domínio tem relação significativa com o uso de atributos de imagem semânticos nas descrições dos usuários da imagem. Os especialistas adotam mais atributos de imagem de alto nível que exigem conhecimento prévio ou do diagnóstico (Objetos Abstratos e Cenas) do que os novatos. Já estes, identificam os objetos básicos, o que não exige conhecimento radiológico prévio (Objetos genéricos). Todos os participantes da pesquisa adotaram atributos de imagem de níveis semânticos na descrição, de modo a poder recuperar futuramente a imagem, em especial Cena Genérica, Específica e Abstrata.

Wang et al. (2012) concluem que o nível de descrição da imagem deve coincidir com o nível de interesse que o usuário objetiva recuperar na imagem, sendo necessário providenciar pontos de acesso multifacetados, de modo que aquele recupere a informação a partir de diferentes aspectos como qualidade da imagem, etapas da doença, entre outros. Tal fato confirma que o processo de indexação deve ser realizado pelos profissionais da área, assessorados por profissionais da informação, como exposto por Panzer (2016), que destaca a importância de se compatibilizar a linguagem médica com a linguagem do usuário para garantir o acesso.

Entendemos que, no estudo apresentado por Wang et al. (2012), os usuários, apesar de possuírem três níveis de especialização, eram da área médica da radiologia, justificando a ausência de gradações semânticas na busca empreendida pelos usuários. Contudo, acreditamos que tal resultado seria diferente se os usuários, mesmo da área médica, fossem de especialidades diferentes.

Em pesquisa posterior, Wang & Erdelez (2013) estudaram como o domínio do conhecimento influencia na elaboração das atividades de busca no tocante à seleção do termo, como também na escolha das táticas a serem adotadas. Os tipos de etapas de pesquisa de imagens foram classificados em geral, específico e abstrato, baseado nas categorias de análise de Shatford (1986), apresentando resultados interessantes. Ao contrário da hipótese inicial, os especialistas fazem buscas gerais para recuperar um conjunto de imagens, possibilitando a seleção da mais adequada a sua necessidade informacional. Já os participantes novatos e intermediários adotaram o nível específico nas buscas, o que revela uma preocupação em não recuperar imagens que fujam dos termos usados na estratégia de busca.

A partir da pesquisa de Wang & Erdelez (2013), observamos que a identificação do nível de representação das imagens radiológicas varia entre usuários de uma mesma comunidade: a especificidade atende às demandas dos usuários novatos e intermediários por apresentarem informações precisas; no caso dos especialistas, o nível é genérico. Tal conclusão indica que deve-se atentar tanto à identificação dos aspectos a serem observados nas imagens, quanto ao nível a ser adotado na representação.

Por sua vez, Kim (2013) destaca a importância da participação dos alunos da área médica na indexação das imagens. Segundo o autor, o treinamento do usuário final pode melhorar tanto a qualidade da indexação como a recuperação das imagens biomédicas. Para isso, o autor comparou a complexidade do conceito (completude de cobertura das unidades indexáveis identificadas em imagens individuais) e

efetividade na recuperação de termos de indexação gerados por estudantes de medicina ao indexarem imagens de histologia. Instruções sobre indexação foram fornecidas em vídeo ao grupo de modo a avaliar a qualidade da indexação feita pelo usuário final. Kim (2013) concluiu que, quanto à atribuição de termos de indexação, o grupo que recebeu instruções atribuiu menos termos do que o grupo sem instruções, mas a diferença entre os grupos foi pequena. Ao analisar o mapa conceitual gerado pela pesquisa, observou-se que o grupo que não recebeu instruções gerou mais tipos semânticos do que o outro grupo, sendo que este seguiu os tipos semânticos apresentados no vídeo. Os tipos semânticos atribuídos pelos estudantes podem servir como guia para a indexação de imagens histopatológicas, destacando aspectos como componentes celulares, anticorpos, coloração, tecido e ampliação da imagem, além de informações constantes na legenda.

Apostolova et al. (2013) focam na representação de imagens de exames compostos por uma série de imagens ou multi-imagens, como a tomografia computadorizada. Enquanto segmentos de um mesmo exame, a recuperação individual de cada imagem constitui um problema, pois cada imagem possui características distintas da outra. Os autores objetivam automatizar a segmentação de imagens que compõem os exames, de modo a preparar a imagem para a indexação semântica automática. Para isso, relatam o método, combinando o resultado do conteúdo da imagem com etapas de processamento baseadas em reconhecimento automático de texto para separar as imagens que compõem o exame multi-imagem, e atribuir a cada uma delas uma legenda. A pesquisa de Apostolova et al. (2013) apresenta solução para operacionalizar o tratamento de informação de exames com múltiplas imagens, apesar de não indicar os aspectos que serão observados por ocasião da indexação automática.

Splendiani e Ribera (2016) analisam imagens e legendas publicadas em artigos acadêmicos de Biomedicina, Computação e Matemática, com o intuito de avaliar até que ponto os textos poderiam ser adotados como descrições alternativas das imagens. O estudo considera o acesso à informação das imagens, a partir da extração automática das legendas, por leitores com deficiência visual. Os autores entendem que as legendas são confiáveis por serem redigidas pelos autores, além de serem significativas por ilustrarem os textos. As palavras das legendas foram extraídas automaticamente e categorizadas manualmente segundo o modelo teórico piramidal de Jaimes e Chang (1999), também usado por Wang et al. (2012). Como resultado, Splendiani e Ribera (2016) apresentaram que o número de caracteres nas legendas que acompanham imagens em artigos médicos é bem maior do que nas áreas de Computação e Matemática, o que demonstra a necessidade da imagem médica apresentar uma legenda descritiva para seu pleno entendimento pelo leitor. Splendiani e Ribera (2016) identificaram nas legendas das imagens biomédicas os aspectos: características visuais primitivas (linha, círculo), elementos de localização espacial na imagem (direita, acima); tipo de imagem; e verbos. Os autores afirmam que as legendas da área médica poderiam ser mais detalhadas, descrevendo o que está presente na imagem e também sobre o que ela trata.

Em artigo sobre recuperação de imagens médicas utilizando textura no auxílio para definir o diagnóstico clínico, Ayyachamy (2015) propõe que o atributo textura seja automaticamente extraído das imagens por softwares. O autor esclarece que o uso do atributo na indexação de imagens por características de baixo nível possui dois propósitos: a imagem pode ser considerada um mosaico composto por diferentes regiões de textura, que serão usadas como exemplos para recuperar áreas similares; e que a textura pode ser empregada automaticamente para se referir ao conteúdo de uma imagem.

CONCLUSÕES

Ao fazer amostragem de artigos sobre "informação em saúde", "informação em ciências da saúde", "imagem biomédica" realizamos análise semântica para identificar os artigos que efetivamente abordavam o tratamento informacional da imagem biomédica.

Inferimos que representar a imagem médica não se restringe a descrever as regiões e como se apresentam. Há informação implícita nas imagens que, ao serem "lidas" pelo profissional médico, indicam ou não a incidência de doenças. Daí a complexidade da representação das imagens médicas: o ato de representar exige a interpretação do que é visualizado pelo profissional, para definição do tratamento e procedimentos a serem realizados.

Apesar do processamento de imagem médica assistido por computador ser disseminado na literatura, nossa pesquisa objetivou identificar na literatura as categorias de alto nível apontadas pelos autores para análise e representação de imagens médicas. Compreendemos que a recuperação baseada em conceitos (CBIR) na área da Medicina se deve ao volume de imagens, fazendo com que a extração de características individuais e de padrões agilize o processamento das imagens e, conseqüentemente, do diagnóstico médico.

No ato de representar, verificamos que algumas ações são inerentes ao indivíduo, como a confecção de legendas informativas para imagens médicas em periódicos, e a análise de aspectos para representação e recuperação de imagens.

Observamos também que categorias presentes em terminologias médicas podem configurar em aspectos usados para representar documentos imagéticos, como o SNOMED-CT (Nomenclatura sistematizada de termos clínicos de Medicina), que apresenta 19 categorias: diagnóstico clínico/doença, procedimentos, entidades observáveis, estrutura corporal, organismo, substância, produtos farmacêuticos/biológicos, espécime, objetos físicos, força física, evento, localização geográfica ou ambientes, contexto social, estágios e escalas, conceitos especiais e qualificadores (Pinto, Rabelo, & Girão, 2014).

Por fim, destacamos o modelo conceitual piramidal de Jaimes e Chang (1999), usado por Wang et al. (2012) e Splendiani e Ribera (2016) como método para análise, apresentando-se como o modelo de descrição de imagens médicas, no período de 2012-2017.

AGRADECIMENTOS

Pesquisa realizada com auxílio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq

REFERÊNCIAS BIBLIOGRÁFICAS

Andrade, J. & Lara, M. L. G. (2016). Interoperability and mapping between knowledge organization systems: Metathesaurus—Unified Medical Language System of the National Library of Medicine. *Knowledge Organization*, 43(2), 107-112.

Apostolova, E., You, D., Xue, Z., Antani S., Demner-Fushman, D., Thoma, G.R. (2013). Image retrieval from scientific publications: Text and image content processing to separate multi-panel Figures. *Journal of the American Society for Information Science*, 64(5), 893–908. doi: 10.1002/asi.22810

Ayyachamy, S.(2015). Registration based retrieval using texture measures. *Applied Medical Informatics*, 37(3), 1-10. Recuperado em 1 fevereiro, 2017 de <https://ami.info.umfcluj.ro/index.php/AMI/article/view/537>

Barcellos, C. (2015). Uso de imagens nos artigos científicos: Visualizar, reter, divulgar, aprender. *Revista Eletrônica de Comunicação, Informação & Inovação em Saúde*, 9(1). Recuperado em 27 janeiro, 2017 de <http://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/924/1569>

Bardin, L. (2006). *Análise de conteúdo*. Lisboa: Edições 70.

Bentes Pinto, V. (2008). Indexação morfossemântica de imagens no contexto da saúde visando à recuperação de informações. *História, Ciências, Saúde – Manguinhos*, 15(2), 313-330. Recuperado em 27 janeiro, 2017 de <http://www.scielo.br/pdf/hcsm/v15n2/05.pdf>

Bentes Pinto, V. & Ferreira, J.L.O. (2010). O que dizem as imagens do campo da saúde: Um exercício de construção ontológica. In Bentes Pinto, V.; Soares, M.E. (Org.). *Informação para a área de saúde: prontuário do paciente, ontologia de imagem, terminologia, legislação e gerenciamento eletrônico de documentos* (vol. 1) (pp. 39-65). Fortaleza: Edições UFC.

Beaudoin, J. E. (2016). Content-based image retrieval methods and professional image users. *Journal of the Association for Information Science and Technology*, 67(2), 350-365. Recuperado em 27 janeiro, 2017 de <http://onlinelibrary.wiley.com/doi/10.1002/asi.23387/pdf>

Boyd, M. (2017), Standards for archives. *Bulletin of the American Society for Information Science and Technology*, 43, 24–28. doi:10.1002/bul2.2017.1720430210

Brasil.(2002). Conselho Federal de Medicina. Resolução 1.638 de 2002. Recuperado em 15 março, 2017 de <http://www.sbp.org.br/arquivos/Resolucao%20CFM%201638%202002.pdf>

Chebil, W. , Soualmia, L. F., Omri, M. N., & Darmoni, S. J.(2015). Indexing biomedical documents with a possibilistic network. *Journal of the Association for Information Science and Technology*, 67(4), 928–941. doi: 10.1002/asi.23435

Chiaravalloti, M. T., Pasceri, E., & Taverniti, M. (2012). URT "Indexing and classification systems" projects and biomedical knowledge standards. *Knowledge Organization*, 39(1), 3-12.

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.(2017). Portal de Periódicos CAPES MEC. Recuperado de <http://www.periodicos.capes.gov.br/>

Ehsani, S., Kiehl, T.R., Bernstein, A., Gentili, F., Asa, S.L., Croul, S.E. (2008). Creation of a retrospective searchable neuropathologic database from print archives at Toronto's University Health Network. *Laboratory Investigation*, 88, 89-93. doi: 10.1038/labinvest.3700694

Fujita, M. S. L.(2003). A identificação de conceitos no processo de análise de assunto para indexação. *Revista Digital de Biblioteconomia e Ciência da Informação*, 1(1), 60-90. Recuperado em 10 março, 2017 de <http://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/2089/2219>

Huvila, I., Cajander, A., Daniels, M. & Ahlfeldt, R.M.(2015). Patients' perceptions of their medical records from different subject positions. *Journal of the Association for Information Science and Technology*, 66(12), 2456–2470. doi: 10.1002/asi.23343

Jaimes, A., & Chang, S. (1999). A conceptual framework for indexing visual information at multiple levels. In Proceedings of the SPIE- IS&T Internet Imaging (Vol. 3964, pp. 2–15). San Jose, CA: SPIE. doi:10.1117/12.373443

Jomsri, P.(2016, December). A combination indexing for image social bookmarking system to improve. *Journal of Digital Information Management*, 14(6), 423-431. Recuperado em 5 fevereiro, 2017 de http://dline.info/fpaper/jdim/v14i6/jdimv14i6_8.pdf

Kim, S. (2013). An exploratory study of user-centered indexing of published biomedical images. *Journal of the Medical Library Association (JMLA)*, 101(1), 73-76. Recuperado em 28 janeiro, 2017 de <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3543137/>

Lu, K. & Mao, J. (2015). An automatic approach to weighted subject indexing-an empirical study in the biomedical domain. *Journal of the Association for Information Science and Technology*, 66(9), 1776-1784. doi.org/10.1002/asi.23290

Marteletto, R. M. & Carvalho, L. S. (2015). Health as a knowledge domain and social field: Dialogues with Birger Hjørland and Pierre Bourdieu. *Knowledge Organization*, 42(8), 581-590.

McDermott, K. (2016, Oct/Nov). Achieving data liquidity across health care requires a technical architecture. *Bulletin of the Association for Information Science & Technology*, 43(1),19-22. doi: 10.1002/bul2.2016.1720430104

McNamara, M., Arnold, C., Sarma, K., Aberle, D.R., Garon, E., Bui, A.A.T.(2015). Patient portal preferences: Perspectives on imaging information. *Journal of the Association for Information Science and Technology*, 66(8), 1606–1615. doi: 10.1002/asi.23269.

Najaftorkaman, M., Ghapanchi, A.H., Talaei-Khoei, A., & Ray, P. (2015). A taxonomy of antecedents to user adoption of health information systems: A synthesis of thirty years of research. *Journal of the Association for Information Science and Technology*, 66(3), 576–598. doi: 10.1002/asi.23181.

Oh, K.E., Joo, S., & Jeong, E. (2015). Online consumer health information organization: Users' perspectives on faceted navigation. *Knowledge Organization*, 42(3), 176-186.

Panzer, M. (2016). Increasing patient findability of medical research: Annotating clinical trials using standard vocabularies. *Bulletin of the Association for Information Science and Technology*, 43(2), 40-4. doi: 10.1002/bul2.2017.1720430213

Pinto, V. B., Rabelo, C. R. O., & Girão, I. P. T. (2014). SNOMED-CT as standard language for organization and representation of the information in patient records. *Knowledge Organization*, 41(4), 311-318.

Ranganathan, S.R. (1967). *Prolegomena to library classification*. Bombay: Ásia Publishing House.

- Richardson, N. C. I. (2016). Software-as-a-medical device: Demystifying connected health regulations. *Journal of Systems and Information Technology*, 18(2), 186 - 215. <http://dx.doi.org/10.1108/JSIT-07-2015-0061>
- Santaella, L.; Noth, W. (1998). *Imagem: cognição, semiótica, mídia*. São Paulo: Iluminuras.
- Silva, J. L. C. (2013). Das concepções disciplinares na Ciência da Informação e/ou de suas configurações epistemológicas: o desiderato percebido da interdisciplinaridade. *Investigación bibliotecológica*, 27(59), 67-92. Recuperado em 14 de março, 2017, de http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0187-358X2013000100004
- Shatford, S.(1986). Analyzing the subject of a picture: a theoretical approach. *Cataloging and Classification Quarterly*, 6(3), 39-62.
- Simpson, M.S. et al. (2013). Multimodal biomedical image indexing and retrieval using descriptive text and global feature mapping. *Information Retrieval*, 17, 229-264. doi:10.1007/s10791-013-9235-2
- Smith , M.A., Barnes, E.L., & Chiose, S.I. (2011). Pathology archive: evaluation of integrity, regulatory compliance, and construction of searchable database from print reports. *American Journal of Clinical Pathology*, 135(5), 753-759. doi: 10.1309/AJCP3CVA2NAVUUUVU.
- Souza, E.G., Souza, J.C.C.E., & Almeida, E.C. P. (2015). Estrutura de metadados para banco de imagens em Patologia. In Guimarães, J.A.C.& Dodebei, V. (Org.). *Organização do conhecimento e diversidade cultural* (vol. 3) (pp.413-424). Marília: ISKO-Brasil: FUNDEPE.
- Souza, J.C.C.E. & Paes-De-Almeida, E. C. (2014). Abordagem teórico-metodológica para organização de imagens em patologia. In *Anais do 15º Encontro Nacional de Pesquisa em Pós-Graduação em Ciência da Informação* (pp 778-797). Belo Horizonte, MG, ECI, UFMG.
- Splendiani, B. & Ribera, M.(2016). Accessibility of graphics in STEM research articles: Analysis and proposal for improvement. *Journal of the Association for Information Science and Technology*, 67(6), 1507–1520. doi: 10.1002/asi.23464
- Tamine, L., Chouquet, C. & Palmer, T.(2015). Analysis of biomedical and health queries: Lessons learned from TREC and CLEF evaluation benchmarks. *Journal of the Association for Information Science and Technology*, 66(12), 2626-2642. doi: 10.1002/asi.23351
- Valle Gastaminza, F. (2002). Dimensión documental de la fotografía. In *Congreso Internacional Sobre Imágenes e Investigación Social*. México. Conferência. Recuperado em 17 fevereiro, 2011 de <http://www.ucm.es/info/multidoc/prof/fvalle/Confemex.htm>.
- Wang, X. & Erdelez, S. (2013). Medical image users' search tactics across different search tasks. In *ASIST '13 Proceedings of the 76th ASIS&T Annual Meeting: Beyond the Cloud: Rethinking Information Boundaries*. Montreal, Quebec, Canada. Recupera em 5 fevereiro, 2017 de <http://onlinelibrary.wiley.com/doi/10.1002/meet.14505001168/pdf>
- Wang, X., Erdelez, S., Allen, C., Anderson, B., Cao, H., Shyu, C. (2012). Role of domain knowledge in developing user-centered medical-image indexing. *Journal of the American Society for Information Science and Technology*, 63(2), 225–241. doi: 10.1002/asi.21686

Wolf, C. T. & Veinot, T. C. (2015). Struggling for space and finding my place: An interactionist perspective on everyday use of biomedical information. *Journal of the Association for Information Science and Technology*, 66, 282–296. doi:10.1002/asi.23178

Wong, S. & Hoo Jr, K. S. (2002). Medical imagery. In Castelli, V. & Bergman, L. D. (editors) *Image Databases: Search and Retrieval of Digital Imagery* (pp 83-103). John Wiley and Sons, New York.

Zhang, Y.; Sun, Y., & Xie, B. (2015). Quality of health information for consumers on the web: a systematic review of indicators, criteria, tools, and evaluation results. *Journal of the Association for Information Science and Technology*, 66(10), 2071-2084. doi: 10.1002/asi.23311



O RESUMO CIENTÍFICO COMO RECURSO DE ACESSO EQUITATIVO À INFORMAÇÃO

Helena Leitão¹, Maria da Graça Simões²

¹Faculdade de Letras da Universidade de Coimbra, 0000-0002-8417-6657, helena.leitao@uc.pt

²Faculdade de Letras da Universidade de Coimbra, 0000-0001-5575-0261, gsimoes@fl.uc.pt

RESUMO A garantia de um acesso equitativo ao conhecimento é fundamental para a evolução científica, consolidação da democracia e do bem-estar. A sua exequibilidade passa pela existência de meios que permitam aos indivíduos ter acesso à informação e ao conhecimento sem constrangimentos de qualquer natureza. Objetiva-se analisar em que medida o resumo científico, como recurso mediador entre o utilizador e a informação original, contribui para o acesso equitativo à informação. Partiu-se de um paradigma qualitativo cujo método foi a revisão de literatura sistemática, a partir da qual se contextualizou, definiu e caracterizou o resumo científico, o conceito de equidade, articulando-se este com o de acesso à informação, refletindo-se ainda sobre a possibilidade de o resumo agilizar o acesso equitativo à informação. Os resultados inferem que o resumo, através das suas propriedades intrínsecas, permite o acesso equitativo à informação e ao conhecimento, na medida em que apresenta a informação condensada e organizada numa macroestrutura. De um modo geral, a terminologia e o idioma são inteligíveis, potencia os recursos económicos e a poupança de tempo, concorrendo para a atualização, a divulgação, o consumo e a produção científica, fomentando a participação ativa da sociedade no desenvolvimento científico.

PALAVRAS-CHAVE *Resumo científico, acesso equitativo à informação, bem-estar social.*

ABSTRACT Ensuring equitable access to knowledge is fundamental to scientific progress, consolidation of democracy and well-being. Its feasibility is the existence of means that allow individuals to have access to information and knowledge without constraints of any nature. The objective is to analyze the extent to which the scientific abstract, as a mediating resource between the user and the original document, contributes to equitable access to information. It was based on a qualitative paradigm whose method was the revision of systematic literature, from which the scientific abstract and the concept of equity were contextualized, defined and characterized, and the concept of equity was articulated with the concept of access to information. It was also reflected on the possibility of the scientific abstract facilitating equitable access to information. The results infer that the abstract, through its intrinsic properties, allows equitable access to information and knowledge, insofar as it presents information condensed and organized in a macrostructure, in general, the terminology and the language is intelligible to whom needs her, potentiates the saving of economic resources and time, competing for the updating of the user and promotes the active participation of society in the development, dissemination, consumption and scientific production.

KEYWORDS *Scientific abstract, equitable access to information, social well-being.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A sociedade contemporânea centrada nas questões dos direitos humanos, em particular, no respeito pela diferença e na justiça social, confronta-se com a necessidade de esbater as barreiras que concorrem para a igualdade e equidade. Apesar de se terem desenvolvido um número significativo de estudos sobre igualdade no acesso à informação, muitos dos quais abordando o movimento do *Open Access*, o número diminui quando falamos de acesso equitativo à informação e ao conhecimento. As possibilidades de acesso geradas por este movimento contribuem para um envolvimento ativo dos cidadãos na construção e discussão da ciência, proporcionando oportunidades de educação e comunicação que reduzem a distância entre público leigo e a informação de natureza científica, gerando a chamada *Ciência cidadã* (Kimura & Kinchy, 2016; Kullenberg & Kasperowski, 2016; Preece, 2017; Ruiz-mallén, Riboli-sasco, Ribault, Heras, & Laguna, 2016). Para Voronin, Myrzahmetov & Bernstein (2011) o acesso à informação é a condição básica para que toda comunidade participe na construção da ciência. Caldas (2010) vai ao encontro desta perspetiva, afirmando que quase todos os acontecimentos quotidianos são influenciados pelo desenvolvimento da ciência e da tecnologia, pelo que é fundamental que o discurso científico seja compreendido pelo máximo de pessoas e que cada indivíduo possua ferramentas que lhe permitam a tomada de decisão consciente e informada, podendo abrir-se espaço para debates públicos sobre a ciência principalmente quando esta gera impactos sociais imediatos.

Harnad (201?) já antecipa a questão do acesso equitativo à informação quando preconiza que o *Open Access* deve ser gratuito, imediato, permanente, em texto integral e *online*, permitindo entre outros, um aumento da visibilidade, viabilidade, compreensão, aplicação e uso da informação que, em última análise, se traduz num crescente impacto no progresso e financiamento de investigações. Esta atitude assume um papel muito relevante para a ciência cidadã, na medida em que, para este movimento, o livre acesso aos conteúdos é uma condição básica, pois só através da sua concretização é possível conhecer, discutir e participar na ciência. Neste contexto, destaca-se o resumo como um recurso de eleição para o acesso à informação seja em acesso aberto ou restrito (Leitão, Borges & Simões, 2015).

Um resumo consiste numa versão abreviada, objetiva e concisa do conteúdo do documento que o origina (ANSI Z39.14-1997, ISO 214:1976 (F)). Representa um documento quanto ao conteúdo e à estrutura, podendo, eventualmente, substituí-lo (Simões, 2015) e não apresenta interpretações sobre o documento base (Simões & Ferreira, 2013). Lancaster (1991) define-o como uma representação sucinta e exata do conteúdo de um documento acrescentando que, ao contrário do extrato, o resumo apesar de eventualmente incluir palavras presentes no documento, é um texto original criado pelo resumidor e não uma transcrição direta. Cremmins (1982) refere as características distintivas que permitem compreender de forma mais precisa o conceito de resumo: o facto de ser estruturado, conciso e coerente, indo assim ao encontro ao que é referido nas normas. Maizell, Smith e Singer (1971, p.1), definem o resumo de uma forma mais simples referindo que se trata de uma *condensation that presents succinctly the objectives, scope and findings of a document*.

Para Grant (2015), a estrutura do resumo deve iniciar-se com a questão de investigação e terminar com a resposta à mesma, registando-se entre estes dois pontos a descrição do processo. Anisimova & Grebenshchikova (2016) concluem que a composição deve incluir informações acerca do autor e das referências bibliográficas mais relevantes quando o objetivo é a publicação numa revista de resumos. Isto porque, não havendo acesso ao texto integral é importante que se possa extrair o máximo de informações formais e referências possível. Em determinadas áreas ou em determinados casos é

preferido um resumo estruturado, recorrendo-se nesse caso a subcabeçalhos - *background, objectives, methods, results and conclusions* (Weinert, 2010).

O resumo assume-se um recurso fundamental quando a informação primária não se encontra numa língua acessível, pois é redigido também em inglês – *abstract* (Simões, Borges, Freitas & Ferreira, 2013). É fundamental no acesso e divulgação da literatura científica, como, quando se usam na submissão de trabalhos a congressos, sendo estes eventos fundamentais para a divulgação da investigação científica (Jacinto, Helvoort, Boots, Skoczynski & Bjerg, (2014), propósito corroborado por Weinert (2010), quando refere que, por isso, não devem ser subestimados pelos investigadores. Ao elaborar-se um resumo deve pensar-se a quem se dirige; atualmente, algumas revistas procuram no resumo um intermediário entre o documento original e um público mais alargado, situação que vai ao encontro do contexto da ciência cidadã.

Os conceitos de igualdade e equidade, como refere a *Sex- and gender-based analysis (SGBA) e-learning resource* (2016) encontram-se frequentemente relacionados. A diferença entre os dois resume-se, basicamente, à ideia de que a igualdade consiste numa distribuição idêntica dos recursos a todos os indivíduos, e a equidade preocupa-se em alocar os recursos consoante as especificidades de cada um, para que todos alcancem o mesmo nível (Mann, 2014; Klees & Qargha, 2014; Minor, 2016), sendo que estes termos apenas são considerados *sinónimos* se os indivíduos partirem do mesmo ponto e necessitarem do mesmo suporte. McCowan (2016), apesar de defender a igualdade de oportunidades, admite que há especificidades que fazem com que seja importante oferecer condições especiais e específicas a alguns indivíduos.

Ao prever-se um mundo focado na garantia do cumprimento do acesso à justiça social, à qual a equidade se encontra estreitamente ligada (Vind, 2015; Bing, 2015; Jaeger et al., 2015), e quando a tendência económica se orienta para um segmento social e solidário, Godoy Viera, Foresti & Rosa (2016) entendem que a informação e o conhecimento são determinantes na concretização deste processo. Este *movimento* traduz-se num crescente desafio para a Ciência da Informação, na medida em que objetiva o desenvolvimento humano sustentável e inclusivo, assumindo-se os serviços de informação como meios privilegiados na sua concretização (Marçal, Amante, Pinto & Neto (2015).

A American Library Association (ALA) entende que a equidade no acesso à informação está relacionada com a possibilidade de cada indivíduo aceder à informação que necessita independentemente da sua idade, nível educacional, língua, limitações físicas entre outras, chamando a atenção que *It also means they are free to exercise their right to know without fear of censorship or reprisal*. O acesso aberto à informação é uma situação que traz alguma inquietude aos cidadãos, sobretudo no que se refere à iniquidade, nomeadamente quando esta ocorre em áreas sensíveis, como é o caso da saúde. Neste sentido, Álvarez et. al (2014) ao refletirem sobre este assunto, referem que o acesso aberto à informação é fundamental para os profissionais desta área, na medida em que dependem da informação para prestar melhores cuidados aos seus pacientes.

Atualmente, existem três problemas no acesso e consumo de informação - os conflitos culturais, os idiomas, a quantidade e a qualidade de informação disponível, sendo, por isso, fundamental criarem-se mecanismos capazes de os ultrapassar para se aceder ao conhecimento (Kucuk, 2016). Segundo este autor, os governos devem responsabilizar-se por prover mecanismos para se atingir a equidade, alertando ainda para a iliteracia digital - *unfortunately, many consumers are not yet digitally literate and drown in unreliable and unnecessary information*. Neste contexto, Luo (2016, p.189) alerta para o papel do bibliotecário de referência no acesso equitativo à informação, na medida em que é este

profissional da informação que estabelece mais interações com o público, constituindo-se neste sentido mediador entre o cidadão e a informação: *reference librarians endeavor to bring to users unbiased and equitable services with the promise of confidentiality and the protection of intellectual freedom*. Como mediador entre estes dois agentes: público e informação, o bibliotecário deverá dominar um conjunto de matérias específicas, situação que na maioria dos casos não acontece, neste sentido o resumo, poderá ser um recurso fundamental, ao descrever de um modo sucinto os temas abordados nos documentos. Deste modo, o resumo assume-se como uma ferramenta potenciadora da mediação.

O facto de alguns consumidores serem incapazes de fazerem um uso correto dos meios digitais e das informações que recuperam (Kucuk, 2016), concorre para que o resumo se assuma como um elemento fundamental neste processo. Também Bing (2015) conclui que a comunidade internacional acredita que a biblioteca é o melhor local para o acesso à informação sendo o *most powerful social security system*. Assim, as comunidades criam bibliotecas públicas de forma a garantirem aos seus cidadãos o direito ao acesso gratuito e livre à informação, que agilizado pelos instrumentos de organização de informação permite o acesso efetivo ao conhecimento.

É nesta dinâmica de complexidade sociocultural e científica em que se promove o livre acesso à informação e a excelência do seu uso racional, que se pondera o papel que os resumos assumem neste contexto, atendendo, em particular, para a função que poderão assumir no desenvolvimento e consolidação da equidade social. Assim, objetiva-se analisar, como o resumo contribui para o acesso equitativo à informação e ao conhecimento, procurando-se: a) definir e caracterizar o resumo científico; b) definir do ponto de vista sociológico os conceitos de igualdade e equidade; c) articular estes conceitos com o acesso à informação; d) refletir sobre a possibilidade de o resumo agilizar o acesso equitativo à informação.

METODOLOGIA

Optou-se por um paradigma qualitativo com recurso à revisão de literatura sistemática e à análise de conteúdo como técnica que, segundo Bardin (2013, p.121) compreende uma *pré-análise*, a *exploração do material* e por fim *o tratamento dos resultados a inferência e a interpretação*. A interpretação dos textos, relativos à equidade no acesso à informação fez-se com base em categorias de análise construídas a partir das características do resumo. A recolha bibliográfica foi feita na base de dados B-on¹. A escolha desta base prende-se com a sua abrangência e relevância no contexto da investigação em Portugal, e por integrar publicações com elevado fator de impacto em distintas áreas, o que concorrerá para resultados consistentes. Na recolha foram considerados artigos publicados entre 2010-2016. O estudo foi elaborado em duas fases:

- i) Recolha dos textos: a) Determinação dos termos de pesquisa (*Resumo científico/Scientific Abstract; Equidade/Equity; Igualdade/Equality; Acesso à informação/Information Access*); b) Construção das equações: orientou-se a pesquisa do particular para o geral, recorrendo-se a filtros cronológicos. Fixou-se para os temas: *resumo* o período de 2010 a 2016, sendo um tema essencial para o enquadramento e determinação das categorias de análise, e para *equidade no acesso à informação* um

¹ Disponível em <http://www.b-on.pt/>

período entre 2014 a 2016 (mais restrito dado o número de resultados); c) Seleção dos documentos para o *corpus do trabalho* com base nos objetivos do estudo, partindo do resumo e palavras-chave. De um universo constituído por todos os documentos que compõem a base, identificou-se uma amostra inicial de 1585 da qual, foram selecionados 10 artigos sobre equidade no acesso à informação que constituem o *corpus* do estudo.

Quadro1 : Artigos sobre equidade selecionados para análise

Autor	Data	Título	Publicação
Vind	2015	<i>Freedom of speech : MORE than free and equal access to information</i>	<i>Scandinavian Public Library Quarterly</i>
Bing	2015	<i>Nanjing Library 's Efforts on Intellectual Freedom</i>	<i>Chinese Librarianship: an International Electronic Journal</i>
Luo	2016	<i>Ethical Issues in Reference : An in-depth view from the librarians' perspective</i>	<i>Reference and User Services Quarterly</i>
Pohle	2015	<i>UNESCO and INFOethics: Seeking global ethical values in the Information Society</i>	<i>Telematics and Informatics</i>
Jaeger et al.	2015	<i>The Virtuous Circle Revisited: Injecting Diversity, Inclusion, Rights, Justice, and Equity into LIS from Education to Advocacy</i>	<i>Library Quarterly</i>
Godoy Viera, Foresti e Rosa	2016	<i>Informação e conhecimento na economia solidária</i>	<i>Biblios: Journal of Librarianship and Information Science</i>
Hinojosa Álvarez, Serra Larin, Gutiérrez Soto, & Luis González	2014	<i>Initiatives for equitable access to scientific information at the National Institute of Hygiene, Epidemiology and Microbiology of Cuba</i>	<i>Revista cubana de higiene y epidemiología</i>
Mathiesen	2015	<i>Human Rights as a Topic and Guide for LIS Research and Practice</i>	<i>Journal of the Association for Information Science and Technology</i>
Kiney	2016	<i>The Internet, Public Libraries, and the Digital Divide</i>	<i>Public Library Quarterly</i>
Lazzarin & Souza	2015	<i>Aspectos que interferem no acesso à informação e interação dos usuários cegos com o OPAC em bibliotecas universitárias.</i>	<i>Perspectivas em Ciência da Informação</i>

Fonte: Original das autoras

- ii) Leitura e interpretação dos documentos selecionados: A partir da leitura dos textos relativos aos resumos, extraiu-se a informação para a determinação das categorias de análise – estrutura, dimensão, linguagem, idioma, foco no utilizador, acesso e mediação - que foram aplicadas à leitura dos textos relativos à equidade no acesso à informação. Por fim cruzou-se a informação de modo a construir os resultados e as conclusões.

RESULTADOS

A partir da análise de conteúdo chegaram-se aos resultados que se apresentam no *Quadro 2* simplificada e, no anexo 1 de forma desenvolvida. Esta opção teve sobretudo a ver com a dimensão do quadro desenvolvido.

Quadro 2: Presença (X) das características do resumo nas obras relativas à equidade no acesso à informação

Título	Autor/Data	Estrutura	Dimensão	Linguagem	Idioma	Foco no utilizador	Acesso	Mediação	Total
<i>Freedom of speech : MORE than free and equal access to information</i>	Vind, T. (2015)				X	X	X	X	4
<i>Nanjing Library 's Efforts on Intellectual Freedom</i>	Bing, W. (2015)						X	X	2
<i>Ethical Issues in Reference : An in-depth view from the librarians' perspective</i>	Luo, L. (2016)					X	X	X	3
<i>UNESCO and INFOethics: Seeking global ethical values in the Information Society</i>	Pohle, J. (2015)				X	X	X	X	4
<i>The Virtuous Circle Revisited: Injecting Diversity, Inclusion, Rights, Justice, and Equity into LIS from Education to Advocacy</i>	Jaeger et al., (2015)			X	X	X	X	X	5
<i>Informação e conhecimento na economia solidária</i>	Godoy Viera, A. F.; Forestí, F. & Rosa, V. (2016)					X	X	X	3
<i>Initiatives for equitable access to scientific information at the National Institute of Hygiene, Epidemiology and Microbiology of Cuba</i>	Hinojosa Álvarez,M; Serra Larin, S.; Gutiérrez Soto, T. & Luis González, I. (2014)					X	X	X	3
<i>Human Rights as a Topic and Guide for LIS Research and Practice</i>	Mathiesen, K. (2015)	X				X	X	X	4
<i>The Internet, Public Libraries, and the Digital Divide</i>	Kiney, B. (2016)						X	X	2

<i>Aspectos que interferem no acesso à informação e interação dos usuários cegos com o OPAC em bibliotecas universitárias</i>	Lazzarin, F.A & Souza, M. R. F. (2015)	X	X	X		X	X	X	6
Total (ocorrência absoluta)		2	1	2	3	8	10	10	36
Total (ocorrência percentual)		5,56%	2,78%	5,56%	8,33%	22,22%	27,78%	27,78%	100%

Fonte: Original das autoras

Os resultados demonstram que as propriedades dos resumos podem ser fundamentais no processo de acesso equitativo à informação, embora com níveis distintos de expressividade. Nos textos considerados as propriedades *mediação* e o *acesso* apresentam dez ocorrências, respetivamente (27,78%); o *foco no utilizador* 8 ocorrências (22,22%), o *idioma*, três (8,33%); a *linguagem e a estrutura* duas (5,56%), e a *dimensão* uma (2,78%).

Do ponto de vista da sua *estrutura*, o resumo, respeitando a estrutura do documento original, permite aceder aos seus conteúdos com fiabilidade e pertinência (Simões, Borges, Freitas & Ferreira, 2013; Grant, 2015), concorrendo para um acesso organizado (Lazzarin & Sousa, 2015; Mathiesen, 2015), generalizado e pertinente à informação. A secção de palavras-chave acaba por permitir a construção de hiperligações proporcionando um maior número de respostas passíveis de resolver um maior número de questões e, dessa forma, apoiar equitativamente um maior número de pessoas. Quanto às *dimensões* do resumo, a informação condensada, abreviada e concisa (Lazzarin & Sousa, 2015) proporciona uma leitura célere e com menor margem para erros, propriedade relevante para indivíduos com limitações visuais, contribuindo ainda para uma maior atualização da informação a todos os cidadãos. No que concerne à *Linguagem*, o facto de ser redigido numa linguagem natural e o mais clara possível (Lazzarin & Sousa, 2015) torna-o perceptível a muito mais indivíduos e grupos (Jaeger et al., 2015) sejam eles especialistas ou *público leigo* (Dubé & Lapane, 2014); Relativamente ao *Idioma* é importante referir que, em ciência, sendo exigida na maioria dos casos uma versão em idioma inglês, facilita a compreensão do conteúdo para um maior e mais diversificado número de pessoas (Jaeger et al., 2015; Leitão, 2015; Pohle, 2015; Simões et al., 2013; Vind, 2015). Quanto ao *Foco no utilizador*, a expressão *reader in mind* (Ketcham, Hardy, Rubin & Siega, 2010, p. 4) é sintomática do interesse desta propriedade. O resumo é construído pensando essencialmente no utilizador. Por ser objetivo e usar um discurso simples concorre para a abrangência dos utilizadores, tornando a sociedade mais inclusiva, mais equitativa (Godoy Viera, Foresti & Rosa, 2016). Ao nível do *Acesso* à informação, promover-lo de uma forma livre e aberta, é no atual paradigma da Sociedade da Informação, uma necessidade premente (Bing, 2015; Luo, 2016; Vind, 2015). Na maioria dos casos, o resumo encontra-se em acesso aberto, independentemente do contexto do documento original (aberto ou restrito) o que faz dele um recurso de excelência para a promoção de equidade no acesso. Relativamente à *Mediação*, Bing (2015) e Vind (2015) defendem a importância das bibliotecas no acesso democratizado ao conhecimento, assumindo o serviço de referência, pelo seu caráter de interação com o público e personalização, uma função de mediador entre a informação e o consumidor (Mathiesen, 2015; Pohle, 2015), assumindo assim um papel de destaque, no acesso equitativo ao conhecimento (Jaeger et al., 2015; Luo, 2016). Neste contexto, o resumo comporta duas funções: na seleção dos documentos, e no apoio ao profissional da informação, quando este não possui competências cognitivas num determinado domínio do conhecimento, revelando-se um recurso fundamental para a compreensão dos conteúdos.

CONCLUSÃO

Atualmente com o crescente interesse no acesso aberto à informação, existe a preocupação de orientar a produção científica para o acesso livre e gratuito, constituindo, deste modo, um meio efetivo para a democratização do conhecimento. Esta circunstância, converge para a promoção do desenvolvimento intelectual e social das comunidades, situação que concorre para a sua participação ativa na evolução da ciência.

Igualdade e equidade no acesso ao conhecimento são dois conceitos fundamentais, que se distinguem fundamentalmente na perspectiva da distribuição dos recursos. Apesar de ambos preverem a justiça social, apenas o segundo a concretiza, na medida em que cada sujeito tem condições distintas de outro, sendo diferente em variados aspetos. Assim, a sociedade da informação, além de prever a igualdade no que concerne às condições de acesso ao conhecimento por parte dos consumidores, deve, neste propósito, considerar, entre outras, as especificidades físicas, linguísticas e socioeconómicas de cada cidadão, pois estas podem ser uma entropia à igualdade de acesso à informação e conhecimento. Para contornar estas eventuais situações é necessário criar estratégias que permitam o acesso equitativo à informação a todos os cidadãos, respeitando as suas diferenças. O resumo pelas suas características constitui-se um recurso democrático ao acesso livre e equitativo ao conhecimento.

A partir da revisão de literatura conclui-se que o resumo é uma versão abreviada, precisa e fiel ao original representando o seu conteúdo e espelhando a sua estrutura. Acompanhando-o quase sempre pode, em algumas situações, substituí-lo. Assim, constitui-se como uma estrutura de organização do conhecimento muito completa inferindo-se pelos resultados que o resumo possa contribuir para a promoção da equidade no acesso à informação, e tal situação tem a ver com as suas propriedades intrínsecas. A IFLA (2011) quando recomenda algumas condições para a implementação do acesso aberto refere a necessidade de estimular a literacia funcional. O resumo como produto estruturado e redigido numa linguagem transversal permite uma maior compreensão dos conteúdos, contribuindo ativamente para a diminuição da iliteracia funcional, da exclusão e da pobreza social, e permite a participação democrática e a equidade social. A poupança de recursos potenciada pelo resumo que fomenta um acesso à informação a dois níveis – formal e de conteúdo - permite que um indivíduo se organize dentro de uma crescente quantidade de informação de forma a seleccionar aquilo que é mais relevante e pertinente. Assim, além de uma poupança de tempo fundamental para os investigadores, verifica-se uma poupança de recursos financeiros na aquisição de documentos em acesso restrito cuja leitura integral não seja necessária para a compreensão do conteúdo, já que o resumo cumpre o propósito. Neste sentido, a condição socioeconómica passa a não ter tanto impacto no acesso à ciência, como é espetável (Godoy Viera, Foresti & Rosa, 2016), situação que ganha relevo em bases de acesso restrito (Leitão, Borges & Simões, 2015). Isto concorre ainda para a *participação dos cidadãos* na ciência que é direta ou indiretamente financiada pelos mesmos já que todos devem ter acesso ao seu desenvolvimento e participar ativamente na sua discussão (Vind, 2015), propósito da *ciência cidadã*. Neste contexto, é fundamental que os cidadãos possam aceder à informação científica de forma breve, através de uma linguagem perceptível, precisa e objetiva de modo a poder compreender de onde se partiu (objetivos) e o que se alcançou (resultados e conclusões), sendo estes aspetos a essência do resumo. Os aspetos referidos concorrem para uma alfabetização funcional que é, do ponto de vista da alfabetização, o grande desafio das sociedades de primeiro mundo.

A Sociedade da Informação, ao prever, em tese, a igualdade e equidade nas condições de acesso ao conhecimento, já que na prática tal propósito está longe de ser alcançado (Castells, 2003), deve criar

estratégias que permitam ultrapassar as necessidades de cada indivíduo. É neste cenário que o resumo se assume, dadas as suas características, um excelente recurso ao acesso democrático à informação, como nos demonstram os resultados. Infere-se que o resumo permita a todos os cidadãos, dependendo, naturalmente, dos sistemas de recuperação da informação, acederem aos mesmos conteúdos. Como estrutura de organização do conhecimento que é, muito completa, intui-se que possa contribuir para o acesso equitativo à informação ao apresentar: a informação estruturada, dimensões reduzidas, sem, contudo, perder a riqueza conceptual do documento original, ser redigido em inglês, o que lhe confere uma transversalidade terminológica compreensível no mundo científico, potenciar o acesso ao conteúdo independentemente deste estar aberto ou restrito e fomentar a mediação entre a informação e o consumidor.

REFERÊNCIAS BIBLIOGRÁFICAS

- Anisimova, A. E., & Grebenschikova, E. G. (2016). The science studies abstracts journal of the RAS institute of scientific information for social sciences as a source of scientific information. *Scientific and Technical Information Processing*, 43(1), 78–82. <http://doi.org/10.3103/S0147688216010111>
- American Library Association. *Equity of Access*. Disponível na WWW: <http://www.ala.org/advocacy/access/equityofaccess>
- Associação Brasileira De Normas Técnicas (2003). NBR 6028: Informação e documentação – Resumo. Disponível na WWW: http://unicentroagronomia.com/destino_arquivo/norma_6028_resumo.pdf
- Bing, W. (2015). Nanjing Library 's Efforts on Intellectual Freedom. *Chinese Librarianship: An International Electronic Journal*, (39), 7–19. Disponível na www: <https://doaj.org/article/3cdebe9f3705497b99cb59fa81a1da7a>
- Bocato, V. R. C. (2011). Os sistemas de organização do conhecimento nas perspectivas atuais das normas internacionais de construção. *InCID*, 2(1), 165–192. <http://dx.doi.org/10.11606/issn.2178-2075.v2i1p165-192>
- Caldas, G. (2010). Divulgação científica e relações de poder. *Informação & Informação*, 15, p. 31-42. Disponível na WWW: <http://www.uel.br/revistas/uel/index.php/informacao/article/viewFile/5583/6763>
- Castells, M. (2003). *A sociedade em rede*. Lisboa: Fundação Caloust Gulbenkian.
- Cremmins, E. T. (1982). *The art of abstracting*. Philadelphia: iSi Press
- De Lara, M. L. G. (2013). Documentary languages and knowledge organization systems in the context of the semantic web. *Transinformacao*, 25(2), 145–150. <http://doi.org/10.1590/S0103-37862013000200005>
- Dubé, C. E., & Lapane, K. L. (2014). Lay abstracts and summaries: Writing advice for scientists. *Journal of Cancer Education*, 29(3). <http://doi.org/10.1007/s13187-013-0570-1>
- Faria, M. I., & Pericão, M. G. (2008). *Dicionário do Livro: Da escrita ao livro electrónico*. Coimbra: Almedina.

- Golub, K. D. T. M. L. Z. Z. M. (2014). Terminology Registries for Knowledge Organization Systems: Functionality, Use, and Attributes. *International Review of Research in Open and Distance Learning*, 65(9), 1901–1916. <http://doi.org/10.1002/asi>
- Grant, I. (2015). The scientific abstract : a beginner's guide. *Journal of Small Animal Practice*, 56. doi:10.1111/jsap.12357
- Harnad, S. Open Access. *Eprints*. Southampton: University of Southampton. Disponível na WWW: <http://www.eprints.org/openaccess/>
- International Federation Of Library Associations And Institutions (2011). *Declaração da IFLA sobre livre acesso – esclarecendo sobre a posição e a estratégia da IFLA*. Disponível na WWW: <http://www.ifla.org/files/assets/hq/news/documents/ifla-statement-on-open-access-pt.pdf>
- Instituto Português Da Qualidade (1988). *NP 418 – Documentação: Resumos analíticos para publicações e documentação*. Disponível na WWW: <http://ancacid.yolasite.com/resources/NP418.pdf>
- Instituto Português Da Qualidade (1988). *NP 4285-3 – Documentação e informação: vocabulário*. Disponível na WWW: <http://hdl.handle.net/10405/30979>
- International Standard Organization (1976). *ISO 214:1976: Documentation -Abstracts for publications and documentation*. Disponível na WWW:http://www.iso.org/iso/catalogue_detail.htm?csnumber=4084
- Jacinto, T., Van Helvoort, H., Boots, A., Skoczyński, S., & Bjerg, A. (2014). Doing science: Writing conference abstracts. *Breathe*, 10(3), 265–269. <http://doi.org/10.1183/20734735.103214>
- Jackson, W. A. (2015). Distributive Justice with and Without Culture. *Journal of Cultural Economy*, 8(6), 673–688. <http://doi.org/10.1080/17530350.2015.1054414>
- Jaeger, P. T., Cooke, N. A., Feltis, C., Hamiel, M., Jardine, F., & Shilton, K. (2015). The Virtuous Circle Revisited: Injecting Diversity, Inclusion, Rights, Justice, and Equity into LIS from Education to Advocacy. *Library Quarterly*, 85(2, SI), 150–171. <http://doi.org/10.1086/680154>
- Kakali, C. (2014). A Utilization Model of Users' Metadata in Libraries. *Journal of Academic Librarianship*, 40(6), 565–573. <http://doi.org/10.1016/j.acalib.2014.08.004>
- Ketcham, C. M., Hardy, R. W., Rubin, B., & Siegal, G. P. (2010). What editors want in an abstract. *Laboratory Investigation; a Journal of Technical Methods and Pathology*, 90(1), 4–5. <http://doi.org/10.1038/labinvest.2009.122>
- Kimura, A.; Kinchy, A.. (2016). Citizen Science : Probing the Virtues and Contexts of Participatory Research. *Engaging Science, Technology and Society*, 2, 331–361. <http://doi.org/10.17351/ests2016.099>
- Klees, S. J., & Qargha, O. (2014). Equity in education: The case of UNICEF and the need for participative debate. *Prospects*, 44, 321–333. <http://doi.org/10.1007/s11125-014-9295-0>
- Kucuk, S. U. (2016). Consumerism in the Digital Age. *Journal of Consumer Affairs*, 50(3), 515–538. <http://doi.org/10.1111/joca.12101>

Kullenberg, C., & Kasperowski, D. (2016). What Is Citizen Science ? – What Is Citizen Science? – A Scientometric Meta-Analysis. *PLoS One* 1–16. <http://doi.org/10.1371/journal.pone.0147152>

Lancaster, F. W. (1991). *Indexação e Resumos: Teoria e Prática*. Brasília: Briquet de Lemos/Livros.

Lazzarin, F. A., & Sousa, M. R. F. de. (2015). Aspectos que interferem no acesso à informação e interação dos usuários cegos com o {OPAC} em bibliotecas universitárias. *Perspectivas Em Ciência Da Informação*, 20(1), 75–91. <http://doi.org/10.1590/1981-5344/1882>

Leitão, H., Borges, M. M., & Simões, M. da G. (2015). *Resumos Científicos em Saúde – Estudo exploratório de resumos de artigos das revistas Nature e PLOS One*. Universidade de Coimbra. Retirado da WWW:<https://estudogeral.sib.uc.pt/jspui/handle/10316/29972>

Luo, L. (2016). Ethical Issues in Reference : An in-depth view from the librarians' perspective. *Reference and User Services Quarterly*, 55(3), 189–198. Disponível na WWW: <https://journals.ala.org/index.php/rusq/article/view/5928/7514>

Mann, B. (2014). *Equity and Equality Are Not Equal*. Disponível na WWW:<https://edtrust.org/the-equity-line/equity-and-equality-are-not-equal/>

Marçal, B.; Amante, M. J.; Pinto, C.; Neto, L. (2015). Conteúdos inclusivos: o papel das bibliotecas na sensibilização dos produtores de informação. In *12º Congresso Nacional BAD*. Évora.

Mccowan, T. (2016). Three dimensions of equity of access to higher education. *Compare: A Journal of Comparative and International Education*, 46(4), 645–665. <http://doi.org/10.1080/03057925.2015.1043237>

Minor, C. (2016). On Digital Equity and Oyster Knives: How Digital Literacy can Disrupt the Status Quo. *Literacy Today*, Abril.

Nagda, S. (2013). How to Write a Scientific Abstract. *Indian Prosthodontic Society*, 13(3). <http://doi.org/10.1007/s13191-013-0299-x>

National Information Standards Organization (1997). ANSI Z39.14-1997 : *Guidelines for abstracts*. Disponível na WWW: http://www.niso.org/apps/group_public/download.php/6610/Guidelines%20for%20Abstracts.pdf

Pohle, J. (2015). UNESCO and INFOethics: Seeking global ethical values in the Information Society. *Telematics and Informatics*, 32(2), 381–390. <http://doi.org/10.1016/j.tele.2014.05.006>

Preece, J., & Preece, J. (2017). Citizen Science : New Research Challenges for Human – Computer Interaction Citizen Science : New Research Challenges for Human – Computer Interaction. *International Journal of Human--Computer Interaction*, 32(8), 585–612. <http://doi.org/10.1080/10447318.2016.1194153>

Ruiz-mallén, I., Riboli-sasco, L., Ribault, C., Heras, M., & Laguna, D. (2016). Citizen Science : Toward Transformative Learning. *Science Communication*, 38(4), 523–534. <http://doi.org/10.1177/1075547016642241>

SGBA. (2016). Distinguish between Equity and Equality. Disponível em WWW:<http://sgba-resource.ca/en/concepts/equity/distinguish-between-equity-and-equality/>

Simões, Maria da Graça; Ferreira, C. (2013). Resumo e abstract : evolução morfológica e semântica. *Biblios: Revista Da Faculdade de Letras Da Universidade de Coimbra*, (XI), 531–550.

Simões, M. da G. (2015). Resumo documental e literatura científica: origem, desenvolvimento e consolidação. *Páginas a&b*, 3(3), 15-36. Disponível na WWW: <http://ojs.letras.up.pt/index.php/paginasaeb/article/view/663>

Simões, M. da G., Borges, M. M., Freitas, M. C. V. de, & Ferreira, C. (2013). O resumo estruturado como recurso na disseminação da produção científica na área da saúde. *I Congresso ISKO Espanha E Portugal / XI Congresso ISKO Espanha 2013*, 879–900. Retirado da WWW:[https://estudogeral.sib.uc.pt/jspui/handle/10316/28546%5Cnhttp://files/4181/Simões et al. - 2013 - O resumo estruturado como recurso na disseminação .pdf%5Cnhttp://files/4182/28546.html](https://estudogeral.sib.uc.pt/jspui/handle/10316/28546%5Cnhttp://files/4181/Simões%20et%20al.%20-%202013%20-%20O%20resumo%20estruturado%20como%20recurso%20na%20disseminação%20.pdf%5Cnhttp://files/4182/28546.html)

Souza, R. R., Tudhope, D., & Almeida, M. B. (2010). Towards a Taxonomy of KOS: Dimensions for Classifying Knowledge Organization Systems. In *ISKO conference*. Retirado da WWW:<http://search.proquest.com/docview/1347770189?accountid=142596>

Godoy Viera, A. F., Foresti, F., & Rosa, V. (2016). Informação e conhecimento na economia solidária. *Biblios: Journal of Librarianship and Information Science*, 62(62), 34. <http://doi.org/10.5195/biblios.2016.263>

Vind, T. (2015). Freedom of speech : MORE than free and equal access to information. *Scandinavian Public Library Quarterly*, (3).

Voronin, Y., Myrzahmetov, A., & Bernstein, A. (2011). Access to Scientific Publications: The Scientist's Perspective. *Plos ONE*, 6(11), e27868. doi:10.1371/journal.pone.0027868

Weinert, C. (2010). Are all abstracts created equal?? *Applied Nursing Research*, 23(2), 106–109. <http://doi.org/10.1016/j.apnr.2008.06.003>

Yoon, K., Dols, R., Hulscher, L., & Newberry, T. (2016). An exploratory study of library website accessibility for visually impaired users. *Library and Information Science Research*, 38(3), 250–258. <http://doi.org/10.1016/j.lisr.2016.08.006>

ANEXO 1 – QUADRO 2 DESENVOLVIDO

Título	Autor/Da- ta	Estrutur- a	Dimensã- o	Linguae- m	Idioma	Foco no utilizado r	Acesso	Mediaçã- o	Total
<i>Freedom of speech : MORE than free and equal access to information</i>	Vind, T. (2015)				(...)But many people do not have that opportunity, because they have not acquired the competences or the language necessary .	(...)which has an impact on the individual citizen's own forms of expression (...)	Free and equal access to information is one of the library's most important task (...)	(...)This is where the libraries can step in , either in collaboration with primary and secondary schools or associations, or independently.	4
<i>Nanjing Library 's Efforts on Intellectual Freedom</i>	Bing, W. (2015)						In a civilized society, free and equal access to information resources is every citizen's basic constitutional rights.	(...)the library is not the sole means to guarantee citizens' right to freedom of knowledge , it undoubtedly plays a very important role.	2
<i>Ethical Issues in Reference : An in-depth view from the librarians' perspective</i>	Luo, L. (2016)					(...) bring to users unbiased and equitable services (...)	Two issues regarding equitable access to information and service are prevalent	(...)bring to users unbiased and equitable services with the promise of confidentiality and the protection of intellectual freedom (...)	3
<i>UNESCO INFOethics: Seeking global ethical values in the Information Society</i>	Pohle, J. (2015)				(...) the efforts to foster linguistic diversity (...)	(...) to the level of the individual user and the capacities of individuals to efficiently	(...) the question of free access to information as a new human right .(...)	(...) and the relationship between users and owners of information .(...)	4

						<i>make use of communication and information(...)</i>			
<i>The Virtuous Circle Revisited: Injecting Diversity, Inclusion, Rights, Justice, and Equity into LIS from Education to Advocacy</i>	Jaeger et al., (2015)			<i>(...)diverse user groups (...)</i>	<i>A variety of guest speakers (...)</i>	<i>In order to find out what people want, (...)</i>	<i>(...) equity are dependent on information access(...)</i>	<i>(...) services that will truly make a difference in the lives of those in need(...)</i>	5
<i>Informação e conhecimento na economia solidária</i>	Godoy Viera, A. F.; Foresti, F. & Rosa, V. (2016)					<i>(...)mas especialmente por estudar as pessoas que necessitam utilizar e “interagir” com a informação para sua “subsistência” e resolução de problemas.</i>	<i>para facilitar o acesso e a disseminação da informação em todas as dimensões do movimento econômico solidário</i>	<i>lidar não apenas com o enorme volume de informação mas especialmente por estudar as pessoas que necessitam</i>	3
<i>Initiatives for equitable access to scientific information at the National Institute of Hygiene, Epidemiology and Microbiology of Cuba</i>	Hinojosa Álvarez, M; Serra Larin, S.; Gutiérrez Soto, T. & Luis González, I. (2014)					<i>(...) priorizadas y basadas en el nuevo paradigma de los servicios orientado a los usuarios, (...)</i>	<i>(...) a las alternativas de acceso a la información científica (...)</i>	<i>(...) partir de una biblioteca convencional, su departamento de información científica para dar cobertura a las necesidades informativas de sus usuarios.</i>	3
<i>Human Rights as a Topic and Guide for LIS Research and Practice</i>	Mathiesen, K. (2015)	<i>(...)related to information organization, access, and control.</i>				<i>Library and information science (LIS) has always been a discipline with an ethical</i>	<i>(...) people around the globe can access a vast array of information (...)</i>	<i>(...) by “facilitating the effective communication of desired information between human</i>	4

						<i>mission at its core: improving people's lives(...)</i>		<i>generator and human user" through information services and systems</i>	
<i>The Internet, Public Libraries, and the Digital Divide</i>	Kiney, B. (2016)						<i>Although the Internet improves information access for people with disabilities, it also poses barriers to access(...)</i>	<i>(...) patrons may ask librarians to help them evaluate Internet sources (...)</i>	2
<i>Aspectos que interferem no acesso à informação e interação dos usuários cegos com o OPAC em bibliotecas universitárias.</i>	Lazzarin, F.A & Souza, M. R. F. (2015)	<i>Elementos semânticos para marcar a estrutura (...)</i>	<i>(...) os conteúdos que não forem textuais devem ter alternativas em texto de forma clara e concisa;</i>	<i>(...)os conteúdos que não forem textuais devem ter alternativas em texto de forma clara e concisa;</i>		<i>(...)as dificuldades enfrentadas por estes usuários poderão ser minimizadas e suas necessidades melhor compreendidas.</i>	<i>(...)a garantia do acesso à informação para os diversos segmentos da sociedade e (...)</i>	<i>(...)orientação ao usuário de forma a conduzi-lo ao início do conteúdo principal(...)</i>	6
Total (ocorrência absoluta)		2	1	2	3	8	10	10	36
Total (ocorrência percentual)		5,56%	2,78%	5,56%	8,33%	22,22%	27,78%	27,78%	100%



PRINCÍPIOS PARA MODELAGEM DE DOMÍNIOS EM SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO (SOC)

Simone Torres¹, Maurício Barcellos Almeida², Maria da Graça Melo Simões³

¹Universidade Federal de Minas Gerais – UFMG, simone.torres@almg.gov.br

²Universidade Federal de Minas Gerais – UFMG, mba@eci.ufmg.br

³Universidade de Coimbra – UC, gsimoes@fl.uc.pt

RESUMO Os Sistemas de Organização do Conhecimento (SOC) são instrumentos usados na representação do conteúdo de documentos para fins de armazenamento, organização, recuperação e compartilhamento. A elaboração de um SOC implica em modelagem do conhecimento, ou seja, em criação de modelos semânticos, descrições simplificadas da realidade de um dado domínio. A pesquisa aqui apresentada propõe-se ao desafio de contrapor duas abordagens teórico-metodológicas, dedicadas à construção de tesouros e ontologias, a fim de que sejam identificadas suas complementaridades e consolidados os resultados, proporcionando novos conhecimentos. A pesquisa foi classificada como qualitativa, e a metodologia de pesquisa empregou procedimento técnico do estudo de caso. Como resultados apresenta uma síntese das observações realizadas na forma de um fluxograma das etapas de construção de tesouros e ontologias e um quadro sinóptico dos princípios para modelagem de domínios nesses dois SOC. Conclui-se que ambas as metodologias abordam de forma similar a modelagem de domínios, criando modelos semânticos baseados em conceitos, definições e relações semânticas e que os referenciais teórico-metodológicos de tesouros e ontologias, no que se refere à modelagem de domínios, são complementares.

PALAVRAS-CHAVE *Modelagem de Domínios do Conhecimento; Sistemas de Organização do Conhecimento (SOC); Tesouros; Ontologias; Documentação Jurídica.*

ABSTRACT Knowledge Organization Systems (KOS) are instruments applied in the representation of documents content for purpose of information storage, organization, retrieval and sharing. The development of a KOS implies in modeling knowledge, which means to create a simplified description of reality usually called semantic model. The present research takes the challenge of comparing two theoretical-methodological approaches, which have been used to build thesaurus and ontologies, in order to identify complementarities and to consolidate new knowledge. The research is classified as qualitative, and the methodology has the case study as technical procedure. As result, we presented a synthesis of our observations in the form of a fluxogram that shows the phases of construction of both thesaurus and ontologies. The results are organized in a synoptic table in which we carry out a comparison between modeling principles of these two aforementioned KOS. We concluded that both methodologies advance the activity of domain modeling in a similar way, in creating semantic models based on concepts, definitions and semantic relations. Thus, our findings indicate that the theoretical and methodological approaches are complementary.

KEY-WORDS *Knowledge domain modeling; Knowledge Organization Systems (KOS); Thesaurus; Ontologies; Legal documentation*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

O conhecimento, insumo para o desenvolvimento da humanidade, vem sendo preservado, organizado e disponibilizado desde a Antiguidade. Dada a importância da organização, preservação e acesso dele para a permanência e progresso humano, temos uma área das ciências dedicada exclusivamente ao seu estudo - a Ciência da Informação. Dentre os diversos ramos de especialidade da Ciência da Informação, está a área de Organização e Representação do Conhecimento¹, que se dedica ao estudo das teorias, metodologias, instrumentos e produtos para acesso ao conhecimento, o qual pode ou não incidir em transferência para o indivíduo, já que esta é uma ação que dele depende, mas caso ocorra pode culminar na produção de novos conhecimentos. Esses conhecimentos, se registrados, são transformados novamente em objeto de organização e representação em um ciclo infinito no âmbito da Ciência da Informação.

Os instrumentos utilizados para a representação são denominados na atualidade de Sistemas de Organização do Conhecimento (SOC), conhecidos também por *KOS*, acrônimo da denominação em inglês *Knowledge Organization System*. Os SOC são um conjunto de instrumentos que tratam da representação do conteúdo dos documentos para fins de armazenamento, organização, recuperação e compartilhamento. Podemos citar como exemplos de SOC os sistemas de classificação, as listas de cabeçalho de assunto, os tesouros, as taxonomias, os mapas conceituais e as ontologias (Hodge, 2000; Simões, 2008; Lara, 2015). Alvarenga (2001) reflete que as inovações mudam os meios, sofisticam-se os instrumentos e surgem nomes novos, mas a essência permanece. Assim, conforme ressaltado por Lara (2015, p.105), mais do que problematizar ou reorganizar os diferentes SOC, é necessário se apropriar da carga simbólica de cada um, para “tirar deles insumos para a construção de novas perspectivas de organização” (Lara, 2015, p. 105).

Dentre os SOC, destacam-se os tesouros e as ontologias pela complexidade semântica de sua estrutura, pela importância para organização da informação em contextos digitais e pela extensa utilização em serviços de informação. Esses dois instrumentos compartilham o objetivo geral de representar o conhecimento e guardam peculiaridades relacionadas, principalmente, aos operadores a que se destinam e à fundamentação filosófica em que se baseiam. As peculiaridades de tesouros e ontologias se devem principalmente ao momento histórico em que foram criados, refletindo a tecnologia circunstancial e o grau de amadurecimento das metodologias e teorias subjacentes. Suas semelhanças e diferenças já foram apontadas na literatura (Gilchrist, 2003; Sales & Café, 2010; Almeida & Mendonça & Aganette, 2014; Kless *et al*, 2014).

Devido à complexidade dos domínios de aplicação e dos processos de construção, os SOC têm despertado o interesse de pesquisadores oriundos de diversas áreas do conhecimento: Ciência da Informação, Ciência da Computação, Filosofia, Comunicação, entre outras. Dentre as diferentes abordagens encontradas, destacam-se no cenário atual duas vertentes: uma de natureza mais inovadora e fortemente influenciada pela Ciência da Computação e outra mais tradicional, fundamentada na Biblioteconomia. Considera-se que as ontologias e os tesouros são os instrumentos que melhor representam essas duas vertentes no âmbito dos SOC.

¹ Também conhecida no Brasil como Organização da Informação.

No âmbito da Ciência da Informação, a representação do conhecimento está ligada à organização de informações e documentos, sobretudo na padronização da terminologia usada para classificar e buscar a informação de forma precisa. Já, na Ciência da Computação, a representação do conhecimento se desenvolveu como um ramo da Inteligência Artificial, voltado à captura e a explicitação do conhecimento em sistemas. Tem como objetivo possibilitar inferências automáticas e utiliza axiomas no processamento da informação, por meio de ontologias (Robredo, 1986; Sowa, 1999; Almeida, 2006). Cabe ressaltar que, neste trabalho, valemo-nos da noção de representação do conhecimento proposta pela Ciência da Informação, sendo as ontologias consideradas como um dos instrumentos disponíveis para este fim.

A elaboração de um SOC implica modelar o conhecimento. A modelagem é o ato de modelar, ou seja, delinear segundo um modelo, criando ou reproduzindo algo de forma a sobressaltar seu relevo ou seus contornos. É fornecer descrições simplificadas por meio de modelos criados a partir do processo cognitivo de realizar abstrações de recortes da realidade, tendo como objeto teorias ou fenômenos observados. A atividade de criação de modelos é realizada por indivíduos ou grupos nas diversas áreas do conhecimento e integra as raízes do método científico, já que todas as teorias são aproximações da verdadeira natureza das coisas. Os constructos, ou seja, as representações mentais são produzidas de forma descritiva, concretizados sob a forma de modelos abstratos (modelos matemáticos, diagramas gráficos, etc.) e de modelos concretos (maquetes e miniaturas de projeto). A modelagem deve estar ancorada em princípios que norteiam o ato de modelar e que, ao mesmo tempo, possibilitem ao modelizador a liberdade de não ficar refém de um dado modelo específico os (Guarino, 1998; Campos, 2004; Sánchez & Cavero & Marcos, 2005; Almeida, 2006). Os SOC são modelos abstratos, estruturas semânticas formadas por conceitos, e as relações entre esses conceitos, estabelecidas com base na significação em determinado domínio do conhecimento.

As correntes teóricas da Ciência da Informação e Ciência da Computação estão se desenvolvendo paralelamente e o conhecimento se sobrepondo, em relação à modelagem de domínios para construção de SOC. Por um lado, a tarefa dos bibliotecários está sendo automatizada e, por outro, as técnicas de Inteligência Artificial (IA) estão sendo aplicadas a grandes massas de informação que têm de ser ordenadas, pesquisadas e classificadas, antes que as deduções estendidas sejam possíveis. Para lidar com esses problemas, organismos de normalização, associações profissionais e associações industriais desenvolveram padrões para facilitar a partilha. Entretanto, os padrões são parte do problema, já que os campos não podem ficar isolados, todavia todos os campos da ciência têm suas próprias normas especializadas, terminologia e convenções. Assim, o primeiro requisito é o desenvolvimento de padrões para relacionar padrões (Sowa, 1999).

Dentre os domínios do conhecimento, o domínio jurídico é formado pela comunidade discursiva dos operadores do Direito (agentes públicos, advogados, estudantes do Direito), no entanto é de interesse de todos os cidadãos, já que as normas jurídicas regulam a convivência em sociedade. As instituições estatais são as principais responsáveis pela geração, preservação, organização e disponibilização do conhecimento jurídico.

O Direito exerce a função social de regular a conduta em sociedade, sendo formado por complexos processos – produção de normas jurídicas, solução de litígios e o estudo destes processos – gerando um gigantesco volume de informações. O conhecimento jurídico, além de ser registrado para fins probatórios, atua como subsídio em todos os processos que integram o sistema jurídico e seu papel é fundamental na criação de normas que regulam a convivência em sociedade, no suporte ao arbítrio dos

juizes, nos negócios jurídicos, nas ações da Administração Pública, na transmissão de conhecimentos da Ciência do Direito.

Os documentos jurídicos mais comumente destacados pela literatura da Documentação Jurídica, em função de consistirem nas principais fontes do Direito, são os documentos legislativos, os documentos judiciais e a doutrina jurídica. A documentação legislativa é formada pelos atos legais (constituição, leis, decretos, resoluções etc.) e pela documentação produzida no decorrer do processo legislativo (proposições legislativas e documentos acessórios do processo legislativo). A informação registrada em documentos jurídicos não é apenas uma descrição da realidade, mas um instrumento de ação sobre ela, que gera efeitos sociais (Torres, 2013; Torres & Almeida, 2013).

Ao se estudar a modelagem de domínios do conhecimento, pode optar-se por uma das abordagens e a partir dela construir o SOC. O presente trabalho consiste na apresentação parcial dos resultados obtidos em investigação científica realizada no âmbito de pesquisa de doutoramento realizada por Torres (2017) no período de agosto de 2013 a fevereiro de 2017. A pesquisa aqui apresentada, propõe-se ao desafio de contrapor as duas vertentes, a fim de que sejam identificadas suas complementaridades e consolidados os resultados, proporcionando novos conhecimentos que possam integrar o que de melhor as duas abordagens teórico-metodológicas têm a oferecer à Ciência da Informação. Em síntese, o problema colocado pode ser assim delineado: *ainda que criados com finalidades e referenciais teóricos diferenciados, a construção de tesouros e ontologias compartilham etapas de modelagem do conhecimento? Em que medida os conhecimentos produzidos por essas diferentes vertentes da Ciência da Informação são complementares e podem ser reunidos em um só corpo teórico, útil à modelagem de domínios do conhecimento?*

O presente trabalho está organizado da seguinte forma, além da introdução, na seção 2 é apresentada a metodologia de pesquisa adotada, classificada quanto à abordagem do problema em qualitativa, empregando o procedimento técnico de estudo de caso aplicados ao domínio jurídico. Na seção 3, são apresentados parte dos resultados obtidos na pesquisa: um fluxograma das etapas de construção de tesouros e ontologias e um quadro sinóptico com os princípios para modelagem do conhecimento em SOC. Na seção 4, são tecidas considerações finais, em que se conclui que ambas as metodologias abordam de forma similar a modelagem de domínios, criando modelos semânticos baseados em conceitos, definições e relações semânticas. E ainda, que os referenciais teórico-metodológicos de tesouros e ontologias, no que se refere à modelagem de domínios, são complementares.

METODOLOGIA

Para dar cumprimento aos objetivos a que nos propusemos nessa investigação, foi utilizado o método *indutivo* que, segundo Lakatos e Marconi (2005), parte de dados particulares suficientemente constatados para inferir uma verdade geral e universal. Conforme Freitas (2013, p.1082), as escolhas realizadas traduzem as decisões tomadas e devem ser justificadas à luz da razão. Desta forma, por meio do referencial teórico e do estudo de caso foi possível propor princípios para a modelagem de domínios do conhecimento em SOC. A estratégia utilizada na pesquisa foi dividida conforme proposto por Minayo (2011), em três grandes etapas: (1) a fase exploratória, (2) fase empírica e (3) a análise e tratamento dos dados, conforme abaixo.

A fase exploratória consistiu no planejamento da pesquisa. Na fase empírica realizou-se a leitura na íntegra e a extração de trechos dos textos em função da relevância para os objetivos da pesquisa. Na fase de análise e tratamento dos dados, as informações coletadas foram consolidadas em um fluxograma com as etapas e um quadro sinóptico com os princípios para construção de tesouros e ontologias.

O *corpus* da pesquisa foi formado por metodologias de construção de tesouros e ontologias. No caso dos tesouros foi usada a ISO 25.964-1(2011), no segundo caso das melhores práticas à modelagem de ontologias proposta por Arp, Smith e Spear (2015), em combinação com a *OntoForInfoScience*, metodologia proposta por Mendonça (2015). Foram usados ainda textos da área de Ciência da Informação selecionados em função de serem produzidos por autores expoentes, sendo considerada a bibliografia consultada suficiente em relação aos objetivos da pesquisa. Ressalta-se que não houve uma limitação temporal previamente estabelecida, tendo em vista que os tesouros e ontologias se desenvolveram em períodos e contextos históricos diferenciados. Dado os diversos usos para o termo ontologia, a pesquisa ficou limitada às ontologias enquanto instrumentos aplicados à organização do conhecimento, para fins de recuperação da informação. A pesquisa bibliográfica foi realizada, especificamente, sobre as seguintes temáticas: i) Organização e representação do conhecimento; ii) Sistemas de organização do conhecimento (SOC); iii) Modelagens de domínios do conhecimento; iv) Metodologias para a construção de tesouros e ontologias.

RESULTADOS

O processo de construção, tanto de tesouros, quanto de ontologias é formado por diversas etapas. Em ambos os casos, para criar o modelo e fornecer descrições simplificadas do recorte da realidade, objeto da investigação, foi necessário o estudo prévio (1) das metodologias que orientam a construção do SOC, perpassando pelas teorias subjacentes e (2) dos conhecimentos do domínio, buscando compreender a natureza de seus fenômenos. Esse estudo foi um requisito para criar condições, a fim de que a estrutura semântica criada fosse compatível com a estrutura do SOC e, ao mesmo tempo, mantivesse minimamente um raciocínio dialético pautado no conhecimento acumulado pela comunidade científica desse campo do conhecimento.

Os dois modelos abstratos foram elaborados por meio da descrição que possibilitou a criação de um modelo lógico-semântico, a partir da identificação dos conceitos relevantes desse domínio, a coleta de definições que delimitam o significado desses conceitos e a identificação e estabelecimento das relações semânticas existentes entre eles. Todo esse processo permitiu alocar cada um dos conceitos coletados dentro de uma estrutura semântica fundamentada na significação dos conceitos dentro do contexto em análise.

Tendo em vista a necessidade de conhecer as teorias que subsidiam a construção de tesouros e ontologias, considera-se importante apontar algumas observações realizadas durante a revisão de literatura. Em relação às Teorias de Representação, foi possível constatar a partir da obra de Campos e Gomes (2014) que a elaboração de modelos conceituais é uma atividade que está apoiada em correntes filosóficas, e ainda que a corrente adotada não esteja explicitada, influencia o processo de modelagem e os resultados obtidos. No âmbito das Teorias da Representação foram destacadas duas relevantes correntes filosóficas: o Realismo, apoiado nas obras de Barry Smith e seus seguidores e o conceitualismo, apoiado na obra de Ingetraut Dahlberg e seus seguidores.

O Realismo defende a existência de Universais (essências) e Particulares (instâncias), em que o conceito pode ser entendido como um Universal. O Conceitualismo, por sua vez, afirma que os Universais não existem na realidade, mas apenas nas mentes humanas, sendo criados de forma *ad hoc*. No Conceitualismo, o conceito é compreendido como categorias, que podem ser compartilhados por uma pluralidade de pessoas, mas que não correspondem a Universais ou invariantes juntos à realidade. Constata-se que, ainda que as duas correntes tenham divergências em relação ao conceito de conceito, em ambos os casos é trabalhada a criação de classes hierárquicas que reúnem indivíduos ou suas partes em função de características gerais compartilhadas. Compartilham também a criação de relacionamentos de associação não hierárquica entre as classes, em função da proximidade semântica dos conceitos. Nesse sentido, as duas abordagens não são incompatíveis, o que permite que os processos de modelagem do conhecimento possam ser comparados neste trabalho.

O processo de construção, tanto de tesouros, quanto de ontologias é formado por diversas etapas. A ISO 25.964-1(2011) prevê as seis etapas à gestão da construção e manutenção dos tesouros: (1) planejamento, (2) compilação, (3) construção, (4) elaboração da introdução, (5) difusão e (6) atualização. Enquanto Arp, Smith e Spear (2015, p. 50) propõem cinco passos para a modelagem de uma ontologia de domínio: (1) demarcar o escopo da ontologia, (2) reunir informações, (3) ordenar os termos em uma hierarquia do geral para o mais específico (4) verificar o resultado; (5) formalizar o artefato representacional em uma linguagem utilizável por computador. E Mendonça (2015) propõe 8 etapas à construção de ontologias: (1) especificação da ontologia; (2) aquisição e extração do conhecimento; (3) conceitualização; (4) fundamentação ontológica; (5) formalização; (6) avaliação; (7) documentação e (8) disponibilização. A partir dessas abordagens, foi elaborado um fluxograma apresentado na figura 1, consolidando as etapas de construção de tesouros e ontologias.

No fluxograma se buscou respeitar as peculiaridades de cada instrumento e, ao mesmo tempo, evidenciar as etapas compartilhadas. Tendo em vista a divergência na terminologia adotada à denominação de etapas similares nas duas metodologias, esclarece-se que foi adotado o termo considerado como o mais representativo do conteúdo da etapa. É possível observar que na construção de tesouros é necessária uma etapa que torne o instrumento compreensível para humanos (elaboração da introdução e definição das formas de apresentação). Nas ontologias, da mesma forma, a construção deve contemplar uma etapa que a torne compreensível às máquinas (criação de definições e atributos de classes em linguagem lógica).

Em relação à escolha das metodologias, ressalta-se que no caso do tesouro foi relativamente fácil identificar a ISO 25.964-1(2011) e adotá-la. Não houve dúvida com relação à escolha da norma, pois se trata de um padrão internacional, mundialmente aceito. No que se refere às ontologias, na seleção da metodologia o principal problema surgido foi o da ausência de um padrão metodológico universal amplamente aceito. A coexistência de diversas metodologias acarretou dificuldades na seleção do melhor método a ser adotado, já que cada um apresenta características diferenciadas de acordo com sua finalidade, área de aplicação e teorias sobre as quais estão fundamentadas. Constata-se que essa diversidade pode dificultar ou mesmo inviabilizar a reutilização e integração de ontologias, bem como a interoperabilidade² de sistemas de informação, já que as ontologias foram construídas de maneiras diversas. Foi necessário realizar um levantamento dessas metodologias, compreender o âmbito de

² Interoperabilidade é a capacidade de dois ou mais sistemas ou componentes de trocar informações e usar as informações que foram trocadas (ISO 25.964-2/2014).

aplicação e só então fazer uma escolha pela combinação das melhores práticas à modelagem de ontologias proposta por Arp, Smith e Spear (2015), em combinação com a *OntoForInfoScience*, metodologia proposta por Mendonça (2015).

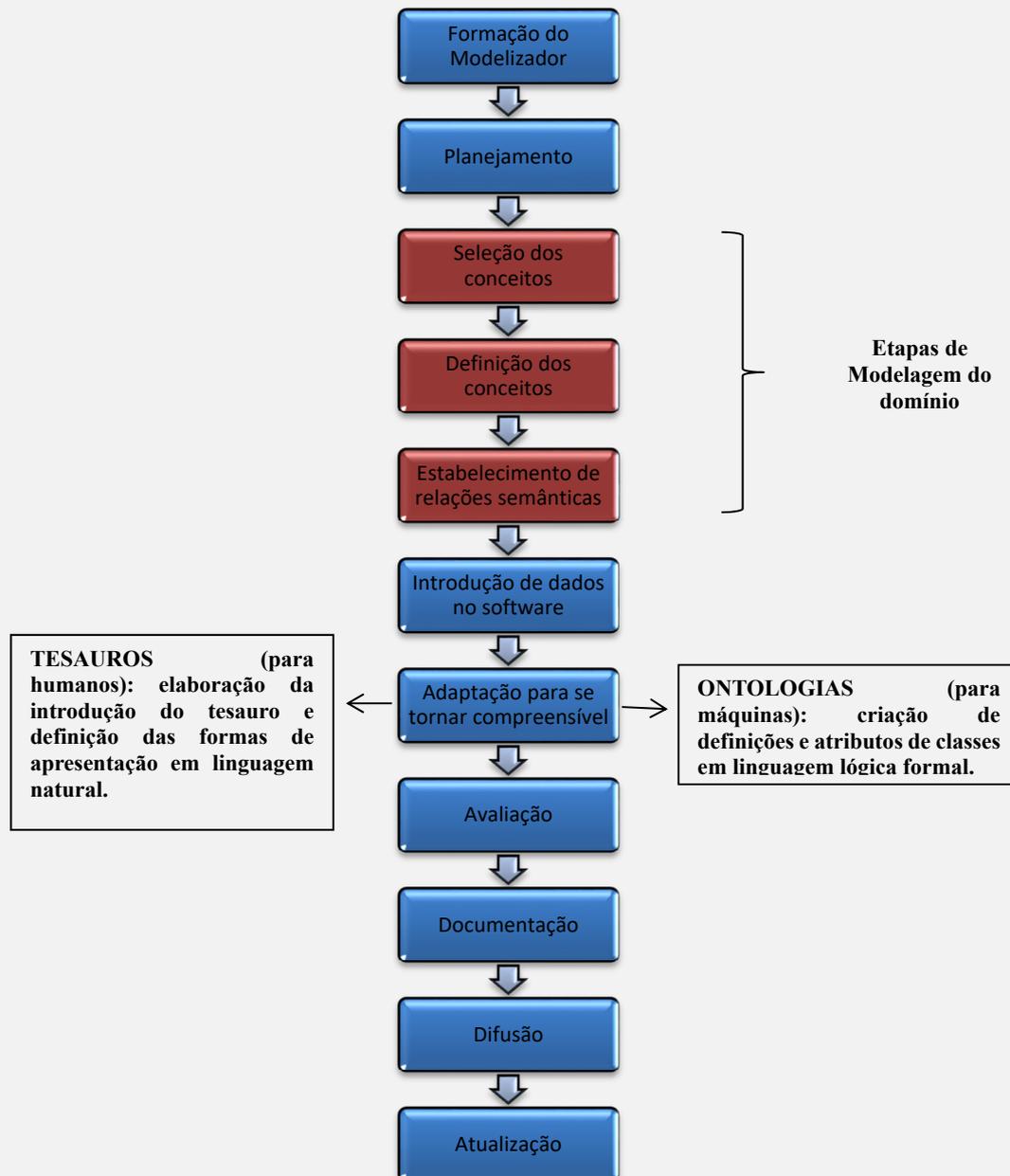


Figura 1. Etapas da construção de tesauros e ontologias

Referência: Torres (2017, p.175).

Quanto à proposição de princípios para modelização de domínios, pela literatura consultada nos foi proporcionado encontrar propostas, tanto na teoria de tesauros quanto na de ontologias. Em relação à teoria dos tesauros, Campos (2004, p.25) propôs quatro princípios à “modelização” de domínios do conhecimento: (1) método de raciocínio (método dedutivo, top-down ou o método indutivo, bottom-up), (2) objeto de representação (a menor unidade de manipulação/representação de um dado contexto), (3) relação entre objetos (Relação categorial, Relação Hierárquica, Relação Partitiva, Relação entre

Categorias, Relação de Equivalência) e (4) formas de representação gráfica (formas para comunicação do modelo). Já, na teoria das ontologias, Arp, Smith e Spear (2015, p. 50) propõe oito princípios gerais de boas práticas que devem ser observadas para a “modelagem de ontologias de domínio”: (1) realismo (a ontologia deve descrever a realidade), (2) perspectivismo (há múltiplas descrições da realidade), (3) falibilismo (as ontologias são revisáveis à luz de novas descobertas), (4) adequação (as entidades em um determinado domínio devem ser seriamente consideradas), (5) reuso (utilização de ontologias existentes), (6) equilíbrio entre utilidade e realismo (achar um ponto de equilíbrio nas decisões pragmáticas), (7) processo contínuo (necessidade de atualização), (8) começar pelo fruto mais baixo (começar a estruturação pelos conceitos mais fáceis de entender e definir). A tabela 1 apresenta um quadro sinóptico dos princípios gerais para a modelagem de domínios do conhecimento, coletados a partir da bibliografia consultada.

Tabela 1 - Quadro sinóptico: princípios gerais para a modelagem de domínios do conhecimento

Princípios gerais para a modelagem de domínios do conhecimento	Referências
<ol style="list-style-type: none"> 1. Conhecimento prévio: é necessário que o modelizador possua formação em teorias de representação do conhecimento, metodologias para criação de SOC e conhecimentos mínimos sobre o domínio que será tratado. 2. Perspectivismo: deve partir-se da premissa de que modelar um domínio é construir um modelo a partir de uma perspectiva de representação de uma realidade, dentre outras infinitas possibilidades de representação dessa mesma realidade; 3. Falibilismo: os modelos, assim como as teorias científicas, são revisáveis à luz de novas descobertas; 4. Contextualização: a modelagem de domínios deve ter em vista selecionar conceitos, criar definições e estabelecer relacionamentos representativos para o domínio tratado, tendo em vista os objetivos da construção do SOC. 5. Garantia literária: o modelo deve corresponder aos conhecimentos aceitos pela maior parte da comunidade científica do domínio registrado na literatura da área; 6. Garantia de uso: os termos utilizados para representar os conceitos selecionados devem corresponder aos termos utilizados pela comunidade científica que será atendida pelo SOC; 7. Garantia estrutural: a estrutura criada deve corresponder tanto quanto possível a um raciocínio dialético, ou seja, os elementos que dela fazem parte devem ser admitidos por sua relevância à estrutura semântica criada. 8. Reuso: outras estruturas de organização do conhecimento sobre o domínio ou domínios correlatos devem ser consideradas como referência e reutilizadas sempre que possível. 9. Método de raciocínio: pode optar-se por modelar o domínio a partir do método indutivo ou do dedutivo. No método indutivo, a terminologia é obtida a partir da incidência de termos na literatura e na consulta de usuários. Já, no método dedutivo, a obtenção se dá pelo processo consensual realizado através de comitês formados por peritos no assunto. 10. Iniciar pelos conceitos mais simples: trabalhar, primeiramente, os conceitos relevantes que são mais fáceis de entender e definir, em seguida trabalhando com os conceitos mais complexos e controversos. 11. Formas de apresentação: deve haver uma preocupação com as formas de apresentação do modelo construído para que se torne compreensível à comunidade atendida, considerando sua perspectiva enquanto espaço de representação. 12. Atualização: a construção do sistema de organização do conhecimento é um processo contínuo em função da necessidade de atualização, consequência da evolução dos conhecimentos. 	<p>Cutter (1904); Hulme (1955); Lancaster (1972, 1987); Foskett, (1973); Cavalcanti (1978), Dodebei (1979, 2002); Gomes (1990); Svenonius (2000); Campos (2001, 2004), Hjørland, (2007); Arp; Smith; Spear (2015).</p>

Referência: Torres (2017, p.197).

Em função de limite de espaço, maior aprofundamento a esses resultados e aos demais aspectos abordados na pesquisa poderão ser consultados em Torres (2017).

CONSIDERAÇÕES FINAIS

A comparação das metodologias e as etapas propostas em cada uma delas permitiu constatar que as metodologias à construção de tesouros e ontologias, ainda que guardem peculiaridades relacionadas à fundamentação filosófica, aos objetivos e, principalmente, aos destinatários de cada instrumento, compartilham etapas de modelagem de domínios que podem se complementar mutuamente. Ambas abordam de forma similar a modelagem de domínios do conhecimento, criando modelos semânticos baseados em conceitos, seus significados e seus relacionamentos. Compartilham ainda o objetivo de representar recortes da realidade, a partir de uma visão de mundo que deve ser dialética e que tenha como objetivo final a organização do conhecimento, possibilitando consolidar etapas e princípios a partir das duas abordagens.

A construção de modelos semânticos aceitos pela comunidade científica para a representação de um domínio do conhecimento exige métodos científicos. A cientificidade dos métodos garante não apenas sua adequada representação, mas resultados satisfatórios, seja na organização e recuperação da informação em bases de dados referenciais e documentais, seja na recuperação de conhecimento em bases de conhecimento e *Web*. Em qualquer modalidade de aplicação, esses modelos semânticos devem ser capazes de estabelecer um raciocínio válido e aceito pela comunidade a que se destinam. Nos tesouros, os seres humanos processam o conhecimento já, nas ontologias, a ênfase é no processamento da informação por computador, o que demanda axiomatizações profundas e precisas, pois não se pode contar com os recursos da cognição humana.

Como limitações da pesquisa, e ao mesmo tempo, possibilidade de estudos futuros, pontua-se a necessidade de análise de outras metodologias propostas para construção de ontologias, já que estas não dispõem de um padrão internacionalmente aceito e variam conforme a fundamentação teórica adotada. Considera-se, entretanto, que a pesquisa atingiu aos objetivos propostos e que tesouros e ontologias têm sofrido influências recíprocas e se tornado cada vez mais relevantes à representação do conhecimento.

REFERÊNCIAS BIBLIOGRÁFICAS

Almeida, M. B. (2006). *Um modelo baseado em ontologias para representação da memória organizacional*. Tese de Doutorado em Ciência da Informação, Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte.

Almeida, M. B. & Mendonça, Fabrício M. & Aganette, Elisângela C. (2014). Interfaces entre ontologias e conceitos seminais da ciência da informação: em busca de avanços na organização do conhecimento. *Anais do XIV Encontro Nacional de Pesquisa em Ciência da Informação*.

Alvarenga, L. (2001). A teoria do conceito revisitada em conexão com ontologias e metadados no contexto das bibliotecas tradicionais e digitais. *DataGramaZero*, 2 (6). Recuperado em 02 março de 2013 de http://www.dgz.org.br/dez01/Art_05.htm.

Arp, R., Smith, B., & Spear, A. D. (2015). *Building ontologies with basic formal ontology*. Mit Press.

- Campos, M. L. A. (2001). *Linguagem documentária: teorias que fundamentam sua elaboração*. Niterói: EDUFF. Recuperado em 15 maio 2010 de <http://www.conexaorio.com/bit/tesauro/index.htm>.
- Campos, M. L. A. (2004). Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. *Ciência da Informação*, Brasília, 33 (1), 22-32.
- Cavalcanti, C. R. (1978). *Indexação & tesauro; metodologia & técnicas: edição preliminar*. ABDF.
- Cutter, C. A. (1904). *Rules for a dictionary catalog*. 4. ed. Washington: Government Printing Office, 1904, 173 p. Recuperado em 24 novembro 2015 de <https://archive.org/details/rulesforadictio06cuttgoog>.
- Dodebei, V. L. D. (1979). *Construção de Thesauri: experimento empírico para a coleta de termos em formação profissional*. Dissertação de Mestrado do Instituto Brasileiro de Informação em Ciência e Tecnologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- Dodebei, V. L. D. (2002). *Tesauro: linguagem de representação da memória documentária*. Niterói: Intertexto; Rio de Janeiro: Interciência.
- Foskett, A. C. (1973). *A abordagem temática da informação*. São Paulo: Polígono.
- Freitas, M. C. V. (2013). Investigação qualitativa: contributos para a sua melhor compreensão e condução. *Indagatio Didactica*, 5(2), 1080- 1101.
- Gomes, H. E. (1990). *Manual de elaboração de tesouros monolíngues*. Brasília: Programa Nacional de Bibliotecas de Instituições de Ensino Superior.
- Gilchrist, A. (2003). Thesauri, Taxonomies and Ontologies. *Journal of Documentation*, 59 (1), 7-18.
- Guarino, N. *Formal ontology in information systems*. 1998. Disponível em: <<http://www.mif.vu.lt>>. Acesso: 02 abr. 2015.
- Hjørland, Birger. Semantic and Knowledge organization. *ARIST*, v. 41, p. 367-405, 2007
- Hodge, Gail. *Systems of knowledge organization for digital libraries: beyond traditional authority files*. Washington, D. C.: The Digital Library Federation Council on Library Information Resources, 2000.
- Hume, D. (1955). *An enquiry concerning human understanding, and selections from A treatise of human nature*. La Salle, IL: The Open Court Publishing Company.
- International Organization for Standardization (2011). *ISO 25964-1: thesauri and interoperability with other vocabularies - part 1 - thesauri for information retrieval*. Genebra: ISO.
- Kless, D., Milton, S., Kazmierczak, E., & Lindenthal, J. (2015). Thesaurus and ontology structure: formal and pragmatic differences and similarities. *JASIST*, 66 (7),1348-1366.
- Lakatos, E. M. & Marconi, M. A.(2005). *Fundamentos de metodologia científica*. 3. ed. São Paulo: Atlas.
- Lancaster, F. W. (1987). *Construção e uso de tesouros: curso condensado*. Brasília: IBICT.

Lancaster, F. W. (1972). *Vocabulary control for information retrieval*. Washington, D.C.: Information Resources Press.

Lara, M. G. (2015). Propostas de tipologias de KOS: uma análise das referências de formas dominantes de organização do conhecimento. *Encontros Bibli*, 20 (1), 89-107.

Mendonça, F.M. (2015). *OntoForInfoScience: metodologia para construção de ontologias pelos cientistas da informação: uma aplicação prática no desenvolvimento da ontologia sobre componentes do sangue humano (HEMONTA)*. Tese de Doutorado, Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte.

Minayo, M. C. S. (2011). O desafio da pesquisa social. In: Deslandes, Suely Ferreira; Gomes, Romeu; Minayo, Maria Cecília de Souza (org). *Pesquisa social: teoria, método e criatividade*. 30 ed. Petrópolis: Vozes.

Sales, R.; Café, L. (2010). Semelhanças e diferenças entre tesouros e ontologias. *Pesquisa Brasileira em Ciência da Informação e Biblioteconomia*, 3 (2).

Robredo, J. (1986). *Documentação de hoje e amanhã*. 2. ed. rev. e ampl. Brasília: Ed. Autor, 201-263.

Sánchez, D. M. & Cavero, J. M. & Marcos, E. (2005). *On models and ontologies*. 2005. Recuperado em 10 abril 2016 de <http://kybele.esctet.urjc.es/PHISE05/papers/sesionIV/SanchezCaveroMarcos.pdf>.

Simões, M.G.M. (2008). *Da abstração à complexidade formal: relações conceptuais num tesouro*. Coimbra: Almedina, 2008.

Sowa, J. F. (1999). *Building, sharing and merging ontologies*. Recuperado em 19 junho de 2016 de <http://users.bestweb.net/~sowa/ontology/ontoshar.htm>.

Svenonius, E. (2000). *The intellectual foundation of information organization*. Cambridge, MA: MIT Press.

Torres, Simone (2013). *A caracterização do documento jurídico para a organização da informação*. Dissertação de mestrado, Escola de Ciência da Informação, Belo Horizonte, Universidade Federal de Minas Gerais.

Torres, S. & Almeida, M. B.(2013). *Introdução ao estudo da documentação jurídica: a caracterização do documento jurídico*. Saarbrücken: Novas Edições Acadêmicas.

Torres, Simone (2017). *Modelagem de domínios em Sistemas de Organização do Conhecimento (SOC): uma investigação em tesouros e ontologias para a informação legislativa*. Tese de Doutorado em Ciência da Informação, Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte.



COMPUTAÇÃO EM NUVEM E SISTEMAS DE GESTÃO DOCUMENTAL: AVALIAÇÃO DE RISCOS E RECOMENDAÇÕES

Ariovaldo Veiga de Almeida¹, Maria Cristina Vieira de Freitas²

¹Universidade de Coimbra, 0000-0001-9563-2657, ari@student.uc.pt

²Universidade de Coimbra, 0000-0002-8849-8792, cristina.freitas@fl.uc.pt

RESUMO Atualmente, as informações são geradas numa velocidade espantosa. Segundo estudo da empresa *International Data Corporation* (IDC) estima-se que até o ano 2020 o crescimento de informações digitais deve chegar a 40000 exabyte e que aproximadamente 40% do total de informações geradas estarão armazenadas na Nuvem. Assim sendo, os sistemas de informação passam a gerir uma grande quantidade de dados armazenados localmente ou na Nuvem, intensificando-se os riscos de perda ou extravio de informações. Este estudo, de natureza exploratória e descritivo-prescritiva, tem como objetivo identificar características essenciais e modelos de serviços na Nuvem, bem como fatores associados à gestão de riscos na sua implementação e na sua utilização. Como resultado, identificamos os principais riscos e formas de os mitigar, balizados por diretrizes internacionais, concluindo que as boas práticas recomendadas pelas normas analisadas não configuram manuais de procedimentos, devendo, cada organização, avaliar de forma detalhada e específica as suas próprias necessidades técnicas e económicas, na adoção de serviços em Nuvem, amparando-se em estudos de custos, riscos e benefícios.

PALAVRAS-CHAVE *Armazenamento de dados, Gestão de riscos, Computação em Nuvem, Sistemas de gestão documental.*

ABSTRACT Currently, the information is generated at an astounding speed. According to a study done by International Data Corporation (IDC), it is estimated that, by the year 2020, the growth of digital information should reach 40,000 exabytes and that approximately 40% of the total information generated will be stored in the Cloud. As such, information systems manage a large amount of data stored locally or in the Cloud, intensifying the risks of loss of information. This exploratory and descriptive-prescriptive study aims to identify essential characteristics and service models in the Cloud, as well as factors associated with risk management in its implementation and use. As a result, based on international guidelines, we identify the main risks and ways to mitigate them, concluding that the good practices recommended by the standards analyzed do not configure a procedure to be followed, each organization should evaluate in detail its own technical and economic needs in adopting cloud services, based on analyses of cost, risk, and benefits studies.

KEYWORDS *Data storage, Risk management, Cloud computing, Document management systems.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Atualmente, o crescimento das informações digitais é muito significativo, com a estimativa de se tornar 50 vezes maior em 2020 (Luo, Wu, Gopukumar & Zhao, 2016; Machovec, 2014; Madrid, 2013; Sosa-Sosa & Hernandez-Ramirez, 2012).

Um número crescente de organizações públicas ou privadas possui Sistemas de Gestão Documental (SGD), usados para produzir, gerir e armazenar documentos e informações, mediados por tecnologias computacionais. Estes sistemas são considerados estratégicos, pois simplificam e flexibilizam a forma como a infraestrutura computacional e de armazenamento de dados é disponibilizada, trazendo benefícios nos custos de implementação, administração, manutenção e escalabilidade. Como garantia do seu bom funcionamento, as organizações devem implementar formas eficazes de controlo, para que os documentos, bem como as informações, permaneçam confiáveis, compreensivos, sistemáticos, íntegros e usáveis pelo tempo que forem requeridos (Antonio et al., 2014; Bushey, Demoulin, & McLelland, 2015; Ferreira, 2012; Machovec, 2014; Noh, 2015; Pires, 2016; Poulo, 2013; Shaw, 2013; Sosa-Sosa & Hernandez-Ramirez, 2012; Wilson, 2012).

Neste contexto, releva a segurança. Segundo Bernstein (1996 cit. por Massingham, 2010, p. 465) “[...] o risco é uma escolha e não destino” e “[...] mesmo quando não podemos eliminar o risco, devemos, pelo menos, antecipá-lo e, assim, colocarmos em ação processos que possam reduzir o seu impacto”. Ainda, segundo Massingham (2010), tragédias como o ataque terrorista às torres do *World Trade Center* ou a fraude fiscal da empresa Enron®, ocorridas em 2001, ou, ainda, desastres naturais, como o tsunami ocorrido no Oceano Índico, em 2004, aumentam em nós a consciência do risco e das suas consequências.

Os riscos associados à inacessibilidade, vazamento, roubo, extravio ou perda de informações devem ser avaliados e fazer parte da política de segurança das organizações, pois afetam diretamente o funcionamento e a prestação de contas, colocando em risco, no limite, a sua existência (Ali, Khan, & Vasilakos, 2015; Borglund, 2015; Chaves, 2011; Chou, 2015; Dias, Rodrigues, & Pires, 2012; Massingham, 2010; Ostrzenski, 2013; Popovic & Hocenski, 2010; Vurukonda & Rao, 2016; Zissis & Lekkas, 2012).

Partindo destes pressupostos, este estudo centra-se numa revisão teórica de textos recentemente publicados, o que lhe confere uma perspetiva atualizada do problema, procurando responder aos seguintes objetivos: i) identificar características e modelos de serviços na Nuvem, ii) bem como fatores relacionados à gestão de riscos na sua implementação e utilização, especialmente nos SGD; iii) apresentar um leque mínimo de boas práticas e de orientações a adotar, para mitigar esses mesmos riscos, amparando-se em modelos de referência internacionalmente válidos.

METODOLOGIA

Trata-se de um estudo de revisão, baseado numa abordagem exploratória e descritivo-prescritiva, de cunho tendencial, apoiada numa amostra intencional de textos científicos e de documentos normativos, recolhidos por meio de pesquisa bibliográfica e documental. Os documentos recuperados foram publicados, na sua maioria, no período de 2011 a 2017. As consultas foram realizadas em múltiplas

bases de dados disponibilizadas no Portal agregador de conteúdos B-On¹. Para complementar esses resultados, recorreu-se a outros estudos disponíveis e com elevados índices de citação, no Google Scholar². Resumidamente, as buscas foram realizadas do seguinte modo (figura 1):

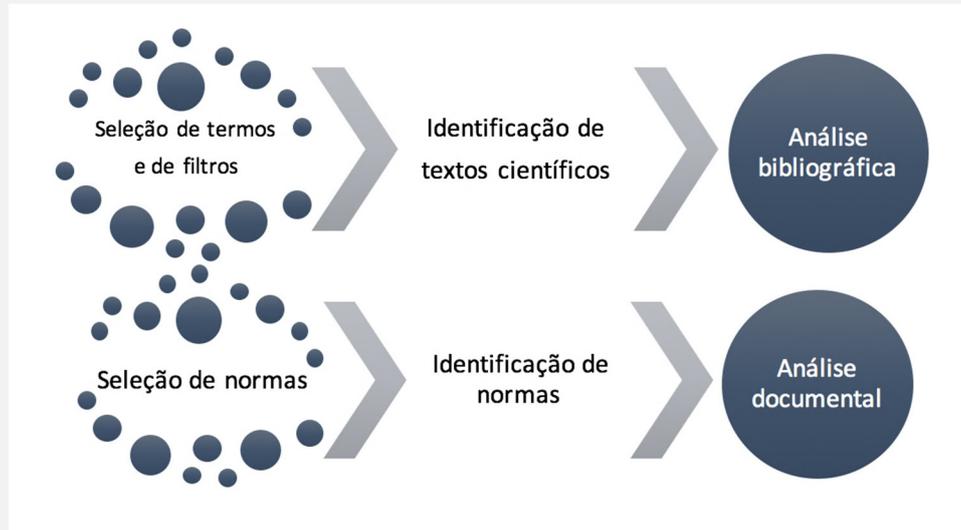


Figura 1. Procedimentos de busca adotados.

Fonte: Elaboração nossa.

Os termos de busca utilizados, isoladamente ou de forma combinada, foram: “*records management*”, “*Cloud storage*”, “*risk management*”, “*archives*” e “*documentation*”. Os filtros utilizados foram: Biblioteconomia e Ciência da Informação (área); Inglês e Português (idioma); Revistas acadêmicas, texto integral e revisto por especialistas (fontes). Feitas estas diligências, foram recuperados 220 artigos, dos quais foram selecionados 38 para compor a revisão, com base na leitura das seguintes secções: resumo, introdução e conclusão.

Entre as várias organizações internacionais preocupadas com criação de normas e de procedimentos para mitigar os riscos do uso de serviços de Nuvem, neste estudo, foram avaliados os documentos disponibilizados por apenas três delas, escolhidas pela sua abrangência e atuação, à escala mundial, bem como pelas citações contabilizadas nos vários documentos utilizados na pesquisa bibliográfica. Deste modo, as normas técnicas foram identificadas em portais de instituições credenciadas, a saber: i) *International Organization for Standardization (ISO)*, ii) *National Institute of Standards and Technology (NIST)* e iii) *Cloud Security Alliance (CSA)*.

Terminada a fase de recolha de dados, passou-se à fase de análise, que, neste estudo, limitou-se à reflexão das ideias principais contidas nos textos pesquisados, que ajudaram a responder aos objetivos de pesquisa delineados, gerando-se, assim, os tópicos incluídos nos resultados, discussão e conclusão.

¹ Disponível em: <http://www.b-on.pt>.

² Disponível em: <https://scholar.google.com>.

RESULTADOS

A COMPUTAÇÃO E O ARMAZENAMENTO DE DADOS NA NUVEM: CARACTERÍSTICAS, MODELOS E IMPLICAÇÕES

Em computação, o termo “Nuvem” terá sido usado pela primeira vez em 1996, pela empresa Compaq® (Regalado, 2011) num plano de negócios da empresa. Academicamente, estima-se que tenha sido empregue primeiramente pelo Prof. Ramnath Chellappa da Escola de Negócios da *Emory University*, no encontro do *Institute For Operations Research and the Management*, em 1997 (INFORMS, 1997). A computação em Nuvem, entretanto, terá raízes mais antigas. Em 1961, o Prof. John McCarthy, pioneiro cientista de computação, já havia abordado o conceito de “*Utility Computing*”, segundo o qual a computação seria adquirida num sistema *pay per use*, de forma semelhante aos outros serviços consumidos diariamente e pagos regularmente (tais como: energia elétrica, água e gás).

Algumas tentativas de descrever o fenómeno têm falhado por não considerar variáveis como diversas tecnologias, configurações e modelos de serviços e de implementação que podem ser usados. Nesse sentido, o NIST, em 2011, publicou um modelo conceitual de Nuvem (Meel & Grance, 2011, p2-3) (figura 2), explicitando cinco características essenciais, três modelos de serviços e quatro modelos de implementação, culminando, no ano seguinte, com as recomendações e chamadas de atenção para aspetos como: oportunidades e riscos, pontos fortes e pontos fracos (Badger, Patt-corner, & Voas, 2012), sendo estes os componentes típicos da Análise *SWOT*³.

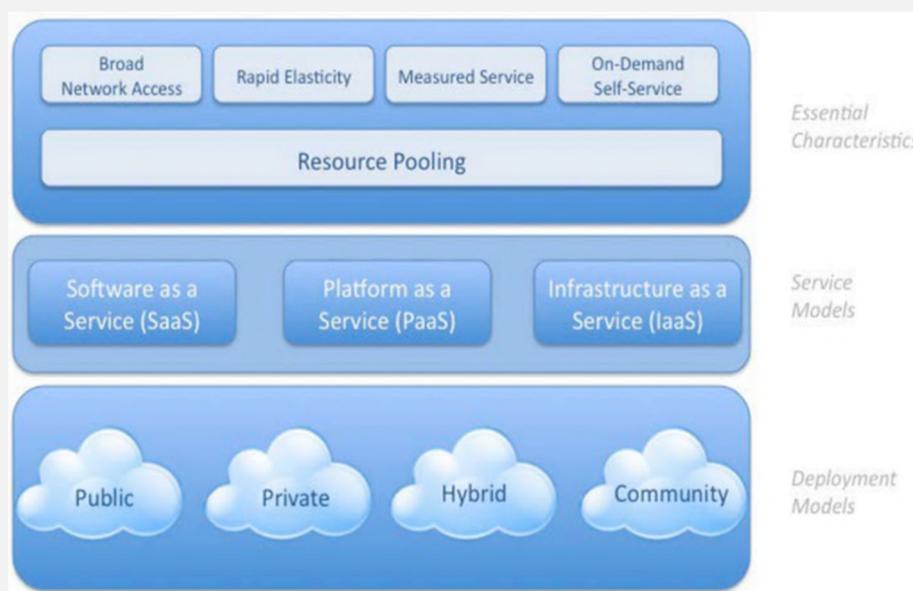


Figura 2. Modelo conceitual de computação em Nuvem, segundo o NIST (2011).

Fonte: (Archer et al., 2011, p.13)

Como vemos na figura 2, as cinco características essenciais da computação em Nuvem apresentadas pelo NIST, de 2011, são: a disponibilidade imediata, o autosserviço, o amplo acesso à rede, a

³ Análise SWOT (*Strengths, Weaknesses, Opportunities e Threats*) ferramenta estrutural da administração utilizada para avaliar os ambientes, sendo usada como base para a gestão e planejamento estratégico de uma empresa.

possibilidade de adquirir um conjunto de recursos agregados, a rápida elasticidade e a mensurabilidade dos serviços. A estas, a instituição CSA, através do relatório publicado em 2011 (Archer et al., 2011), acrescenta uma sexta característica, que corresponde à segregação de serviços (*Multi-Tenency*). Ainda, segundo a mesma diretriz do NIST, de 2011, são três as categorias ou modelos básicos de serviços oferecidos pela computação em Nuvem: o software como serviço, a plataforma como serviço e a infraestrutura como serviço. Finalmente, a mesma diretriz identifica quatro modelos básicos de implementação de uma infraestrutura de computação em Nuvem: privado, público, comunitário e híbrido.

À medida que a computação em Nuvem se torna mais madura e passa a ser utilizada pelas organizações, gerir as oportunidades e os desafios de segurança e, assim, minimizar os riscos, apresenta-se como crucial para o desenvolvimento dos seus processos de negócio.

Ao contrário do que usualmente acontece num ambiente computacional tradicional, a migração para a Nuvem implica perda da barreira de proteção da infraestrutura. Com a sua adoção, a segurança passa a concentrar-se nos aplicativos e na forma como são manipulados os dados, que precisarão de uma segurança própria que os acompanhe e os proteja. Isso implicará o seu completo isolamento, já que precisam ser mantidos em segurança, para que fiquem resguardados quando vários clientes usarem os recursos compartilhados pela infraestrutura.

Assim que migram seus ambientes de computação para a Nuvem, com suas respectivas identidades, infraestrutura e informações, as instituições se veem na iminência de abrir mão de certos níveis de controlo. No caso das Nuvens públicas, a migração exige um modelo de segurança de informação que concilie a capacidade de expansão e de multialocação de recursos computacionais, sendo necessário confiar nos sistemas e nos provedores, para além de executar verificações periódicas nos sistemas. Fazem parte desse processo de confiança e de verificação: o controlo dos acessos, a segurança dos dados, a gestão e a monitorização contínua de eventos e de informações. Também é importante que a virtualização e a criptografia permitam níveis alternativos de separação entre corporações, utilizadores e comunidades de interesse. Em resumo, trata-se de todos os elementos de segurança que são compreendidos por um departamento específico, implantados com a tecnologia existente e com possibilidade de extensão para a Nuvem.

Assim, a adoção da Nuvem exige um olhar acurado para os riscos que o uso desses serviços pode acarretar. Num tal sentido, as organizações devem empregar práticas de gestão de segurança e de controlo apropriadas à computação em Nuvem, essenciais para operar e manter uma solução segura de um serviço desse porte e com essas características. As atividades de segurança e de privacidade implicam, como vimos, monitorizar os ativos do sistema de informações da organização e avaliar a implementação de políticas, padrões, procedimentos e diretrizes, usados para estabelecer e preservar a confidencialidade, a integridade e a disponibilidade dos recursos nos sistemas de informação.

A PERSPETIVA NORMATIVA DA GESTÃO DE RISCOS NAS ORGANIZAÇÕES E NOS SISTEMAS DE GESTÃO DOCUMENTAL

O risco é definido como um evento ou uma sucessão de eventos não esperados, que afetam as organizações e os seus sistemas, manifestando-se pela perceção dos seus efeitos (ISO 18128-2014, 2014). Num tal sentido, e de acordo com (Massingham, 2010, p.469) os riscos podem ser classificados pelos seguintes parâmetros (figura 3):

Table Risk consequence	
Description	Definition
Catastrophic	Failure would prevent the organization from meeting the primary operational requirements
Critical	Failure would significantly degrade the organization's ability to perform its primary mission
Major	Failure would result in temporary loss of one or more significant capabilities within the organization
Minor	Failure would result in temporary degradation or loss of one or more capabilities within the organization

Figura 3. Tabela de classificação de riscos.

Fonte: Massingham, (2010, p. 469).

Note-se que, sendo mais graves, os riscos considerados catastróficos e críticos, pelo impacto sistémico, inibem as organizações do exercício da sua missão, podendo, no limite, decretar o seu encerramento. Assim sendo, é importante identificar e avaliar, o quanto antes, esses tipos de riscos, escolhendo-se as formas de os controlar e mitigar.

No campo normativo, destacam-se pelo menos dois dispositivos elaborados pela ISO e diretamente relacionados à gestão de riscos. O primeiro, de carácter geral – ISO 31000:2009: *Risk management: principles and guidelines* –, oferece princípios e orientações, bem como elenca os processos envolvidos na gestão dos riscos organizacionais. O seu uso estende-se aos diversos tipos de instituições, independentemente do tamanho, da atividade ou do setor de atuação. Essa norma oferece os requisitos necessários à identificação de oportunidades e de ameaças, para além de apoiar as organizações em tarefas de alocação e de uso efetivo de recursos no tratamento dos riscos. Os seus principais pontos, conforme (Lark, 2015), podem ser visualizados no esquema a seguir (figura 4):

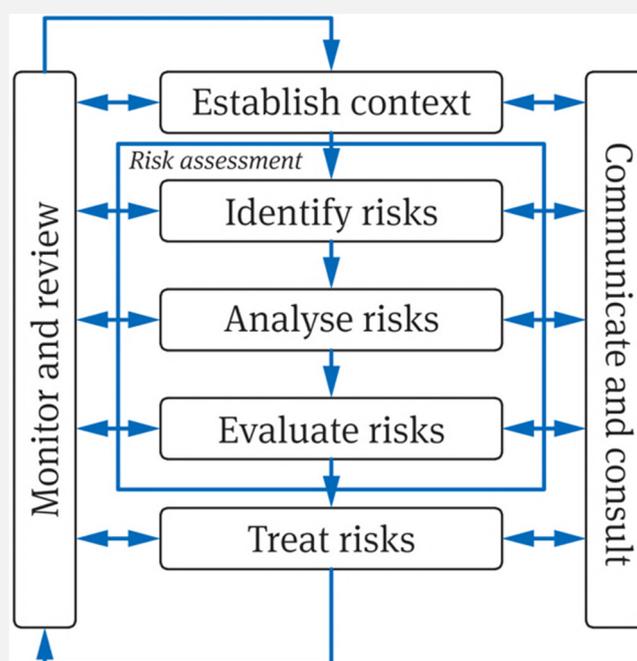


Figura 4. Processo de gestão de riscos segundo a norma ISO 31000:2009.

Fonte: (Lark, 2015, p.14)

Note-se que o processo preconizado pela norma segue um fluxo que inicia com o estabelecimento de um contexto de definição do risco a ser avaliado e prossegue com a sua identificação, análise, avaliação e tratamento. Em cada passo do processo, estão previstas, por um lado, atividades de avaliação periódica, por meio de monitorização e de revisão, e, por outro, atividades de comunicação e de consultoria. Dito diagrama demonstra, ademais, a complexidade e a necessidade de conectividade e de continuidade entre as diversas fases do processo, bem como o seu carácter proativo e reativo.

E, porque há riscos que não se evitam, mas que se controlam, segundo essa mesma norma, a sua gestão deve ir além da adoção de medidas preventivas. Trata-se, pois, de uma tomada de consciência dos riscos que, sendo ou não evitáveis, afetam a organização, de modo monitorizar a sua evolução nos diversos aspetos envolvidos (físicos, ambientais, financeiros e sociais). Num tal sentido, a avaliação deve responder às seguintes questões (figura 5):

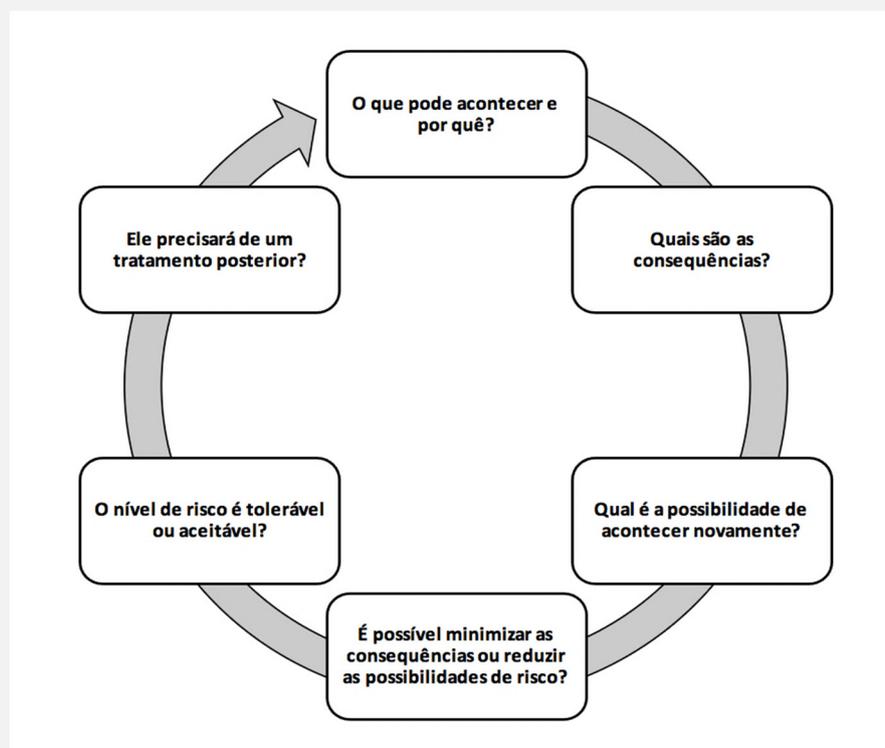


Figura 5. Monitorização dos riscos.

Fonte: Elaboração própria, com base em: (ISO31000-2009, 2009).

Com um carácter específico e de aquisição mais recente, é a norma ISO/TR18128:2014: *Information and documentation: risks assesment for records processes and systems*. Publicada em forma de relatório técnico, trata-se de uma transposição de conceitos e de procedimentos utilizados na norma ISO 31000:2009 para o âmbito da gestão documental. Portanto, esta norma tem como objetivo apoiar diretamente os responsáveis pelos documentos de arquivo, em tarefas de identificação, análise e avaliação dos riscos relacionados aos sistemas⁴ e aos processos de gestão da documentação produzida

⁴ Na aceção dada pela norma ISO 18128:2014, sistema é qualquer aplicação que cria ou armazena os documentos de arquivo, assegurando que os mesmos, ao serem produzidos, usados e geridos pela organização continuem a atender às suas necessidades de negócio pelo tempo que forem necessários.

e acumulada pelas organizações⁵, independentemente do suporte ou do formato. Pelo seu teor, oferece subsídios e diretrizes que contribuem para a implementação de um método de análise eficaz na identificação dos riscos relacionados a esses processos e sistemas, incluindo os fatores externos e os internos suscetíveis de ser encontrados em vários ambientes organizacionais, bem como os seus potenciais efeitos e as formas de os mitigar (ISO 18128-2014, 2014), como podemos visualizar na figura 6.

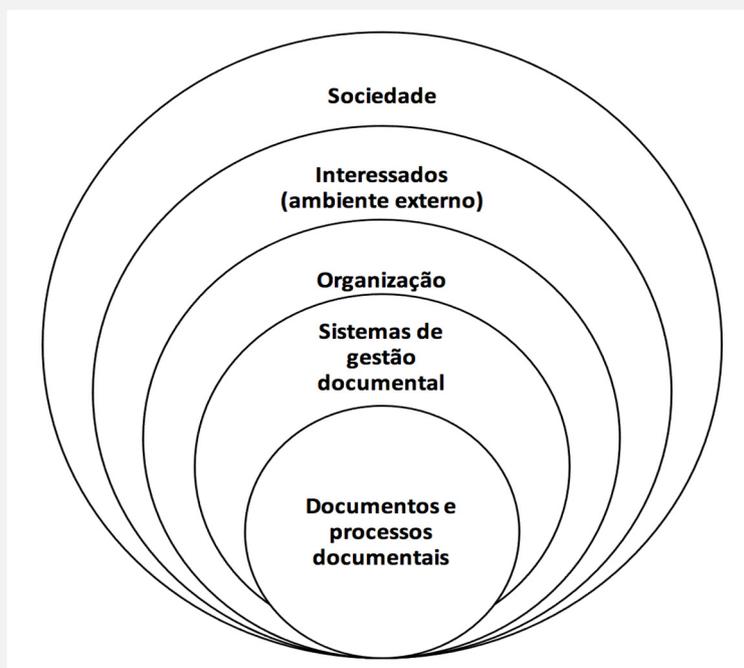


Figura 6. Camadas de intervenção no processo de gestão de riscos nos SGD.

Fonte: Adaptado de ISO 18128:2014 (p. 10).

Saliente-se que a responsabilidade pela identificação e pela gestão dos riscos, nesse caso, recai sobre o profissional da respetiva área (i.é., os arquivos organizacionais), sendo esta uma atividade distinta daquelas em que são identificados e avaliados outros tipos de riscos, que também afetam o bom funcionamento da organização.

Quanto aos impactos, a consequência mais grave associada a um evento de risco em sistemas dessa natureza será a perda ou a danificação irremediável dos documentos, tornando-os inutilizáveis, não confiáveis, inautênticos ou incompletos e, como tal, falhando o próprio sistema em atender aos propósitos da organização. Dada a gravidade, ditada pela extensão da perda, sugere-se que as análises de riscos a realizar nesses sistemas sejam incorporadas às pautas prioritárias das organizações. Os benefícios mais evidentes desta medida seriam um maior e melhor controlo dos documentos, com garantias de preservação da sua integridade, ao longo do tempo, bem como da sua qualidade e do seu uso efetivo, cumprindo-se assim os objetivos para os quais os mesmos são criados (ISO 18128:2014, p. 2).

⁵ Esta norma pode ser usada por qualquer organização pública, privada ou comunitária, associação, grupo ou indivíduo que produza e/ou detenha arquivos.

Porquanto, no contexto dos arquivos, a computação em Nuvem oferece tanto atrativos - inovação, desenvolvimento de serviços, eficiência, economia, escalabilidade e flexibilidade, entre outros - quanto riscos associados a aspetos tais como: segurança, privacidade, integridade, autenticidade, acessibilidade e preservação de informação. Quanto a esta última, refira-se a pouca transparência no que toca à continuidade comercial dos serviços prestados pelas empresas da área. Porquanto, para minimizar os riscos e maximizar os benefícios, devem adotar-se procedimentos standardizados e recomendados (Mckemmish, 2013, p.19). Nesse sentido, quer o cliente, quer o fornecedor dos serviços devem conciliar os seus interesses e requisitos, aquando da implementação de processos de arquivo em Nuvem (Stancic, Rajh, & Brzica, 2015, p214-216). Ainda, uma estrutura simples e eficaz, nesse processo, poderia ser a seguinte (figura 7):

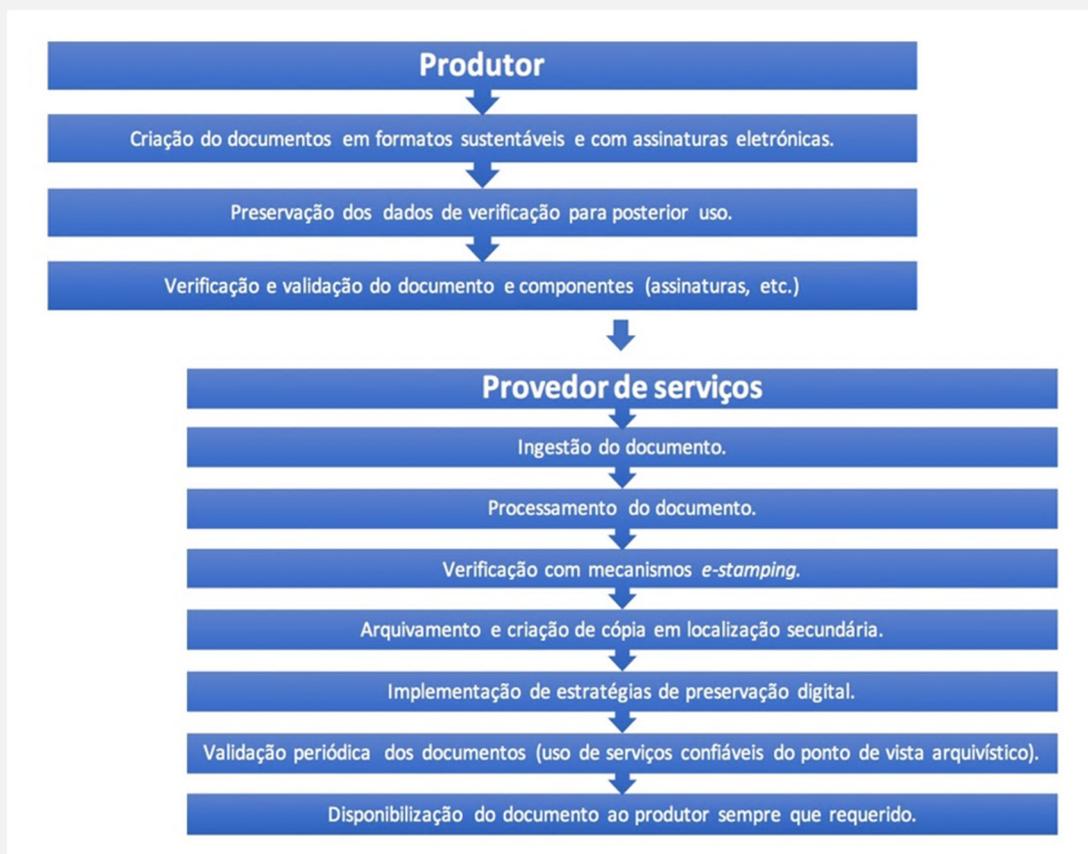


Figura 7. Processo de fornecimento de serviços em Nuvem para arquivos.

Fonte: Elaboração nossa, a partir de Stancic, Rajh e Brzica (2015, p.214-216).

Por seu turno, o Projeto *InterPARES Trust* (2017), de envergadura internacional e com um incontestável prestígio no meio arquivístico, e não só, torna-se num ponto de referência obrigatório para quem pretende desenvolver projetos de migração de arquivos para a Nuvem. Para além de desenvolver o conceito de confiança (*Trust*) e de elencar uma série muito ampla de requisitos necessários à sua obtenção, em níveis aceitáveis e nesses tipos de sistemas, o Website desse megaprojeto internacional recolhe, organiza, descreve e divulga, a toda a comunidade interessada, um vasto leque de projetos de pesquisa e de programas desenvolvidos internacionalmente e que se apresentam como casos de sucesso ou modelos a ser seguidos e/ou adaptados, porque levam consigo a aprovação do *InterPares Trust*. Vale dizer que muitas dos resultados e das recomendações desse projeto, que vem sendo desenvolvido há

vários anos (entre 2013 e 2018) encontram-se espelhados nas normas e noutros documentos citados por este estudo.

O CONTROLO DE RISCOS NA PERSPETIVA DAS ORIENTAÇÕES INTERNACIONAIS

Recomendações do NIST

Em 2011, O NIST elaborou as *Guidelines on security and privacy in public cloud computing*, identificando os desafios à segurança e à privacidade na computação em Nuvem pública. Esse documento é voltado às diversas categorias de profissionais: decisores, profissionais de segurança, auditores, administradores de sistemas e de redes, utilizadores de serviços, entre outros. Devido à natureza evolutiva do tema, o próprio NIST recomendou o uso combinado desse documento com recursos complementares.

Quanto às recomendações, elas podem ser sumariadas do seguinte modo (figura 8):

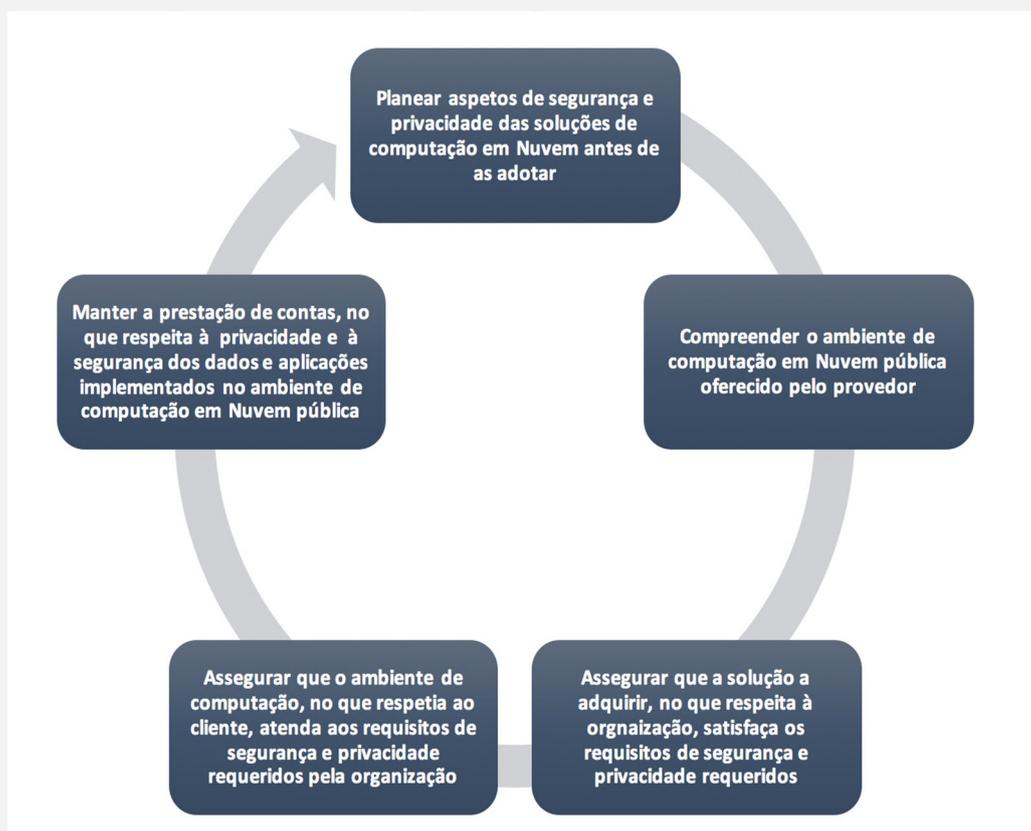


Figura 8. Recomendações do NIST (2011) para a segurança na computação em Nuvem.

Fonte: Elaboração nossa.

Recomendações da CSA

A CSA é uma organização mundial, sem fins lucrativos, que visa promover as boas práticas de segurança em Computação em Nuvem. Fundada em 2008, é uma organização que conta com profissionais da indústria com conhecimentos especializados em Nuvem, associações, governos e

membros corporativos e individuais, e visa oferecer pesquisas, treinamento, certificação, eventos e produtos relacionados especificamente à segurança na Nuvem (CSA, 2017).

As principais recomendações da CSA estão contidas no documento *CSA Security guidance for critical areas of focus in Cloud Computing*, publicado em 2009, e revisado em 2011 (Archer et al., 2011). Trata-se de uma ampla compilação de ideias de mais de setenta peritos internacionais da indústria, embora reconhecendo a abrangência do tema e a inviabilidade de incluir num único documento todos os possíveis cenários e variações.

O guia serve como sustentação na avaliação dos riscos na implementação da computação em Nuvem, incluindo decisões de segurança. Apresenta uma lista de verificação muito simples, em cinco etapas, usada para avaliar a tolerância de um cliente ao mover seus ativos para um dos vários modelos de computação em Nuvem, conforme se observa na figura 9:

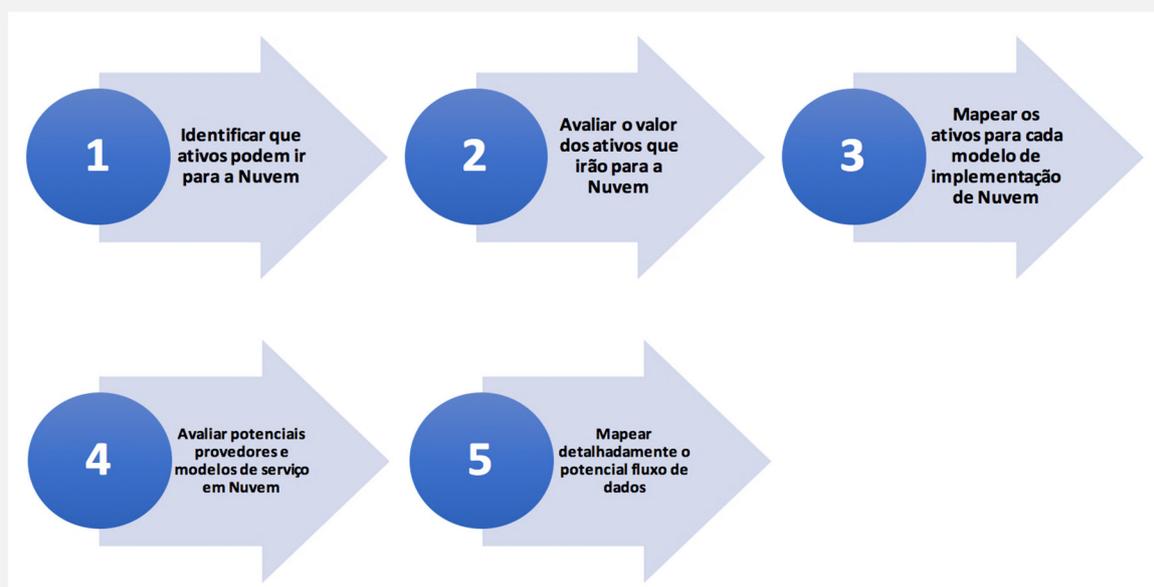


Figura 9. Recomendações da CSA (2011) para a segurança na computação em Nuvem.

Fonte: Elaboração nossa, a partir de (Archer et al., 2011).

DISCUSSÃO E CONCLUSÃO

Segundo Massingham (2010), mesmo quando não podemos eliminar o risco, devemos, pelo menos, tentar antecipá-lo e, assim, colocarmos em ação processos que possam reduzir o seu impacto negativo. Essa colocação se aplica à decisão das organizações na adoção ou não da Nuvem. Da mesma forma que Massingham (2010) menciona que o risco é uma escolha e não um destino, as organizações atualmente possuem muitas escolhas e diversas orientações sobre como devem implementar os seus sistemas computacionais. Atualmente, a adoção da Nuvem é uma excelente opção para as organizações, em geral, devido aos custos envolvidos, aos modelos de serviços e à sua simplicidade de utilização.

As organizações podem optar pela utilização da Nuvem Privada (dentro da própria organização), pela Nuvem Pública (fora da organização), ou por uma mistura delas; mas, como é bom de ver, dependendo da escolha adotada, os riscos serão diferentes, sendo necessário avaliá-los e medi-los.

Um dos principais pontos de discussão, em relação à adoção, ou não, de serviços em Nuvem, prende-se com a capacidade de implementação de um sistema paralelo de medição de riscos. Num tal sentido, são imprescindíveis as normas técnicas e as orientações internacionalmente reconhecidas, das quais trouxemos alguns muito bons exemplos para este estudo. Como vimos, elas oferecem princípios, regras e modelos simples, mas aplicáveis em variados casos, contextos e processos de negócio. Entretanto, e dado o seu caráter geral, ditos dispositivos não proporcionam métodos de análise específicos, sendo esta, naturalmente, uma sua limitação. Com isso, quer-se dizer que uma dada organização pode possuir diversos sistemas e aplicações, devendo, por isso, avaliar, para cada um deles, a viabilidade da adoção da Nuvem, e tomar as suas próprias decisões em função dos riscos observados.

Os processos e os SGD beneficiam-se de princípios e de instruções normativas específicas, dada sua importância para as organizações, quer no registo das suas atividades, quer na prestação de contas e na preservação da memória organizacional. Ditos sistemas fazem parte desse modelo de tecnologias de informação, possuindo, em muitos casos, versões específicas para a Nuvem, escolhidas pela fácil implementação, pelos custos acessíveis e pela escalabilidade. Apesar dessas vantagens, aqui também os riscos são inerentes ao processo e é preciso identificá-los, avaliá-los e tratá-los (pela mitigação ou pela erradicação), adotando medidas de segurança e de monitorização contínua desses sistemas, em tarefas de avaliação periódica.

As organizações devem, pois, empregar práticas de segurança e controlo apropriadas à computação em Nuvem, sem as quais não conseguem manter uma solução confiável de serviço. Avaliar e controlar os riscos nesses sistemas é, desse modo, um desafio e uma meta a atingir. Tanto os indicadores qualitativos quanto os quantitativos podem aplicar-se, devendo ser cuidadosamente ponderados em relação às garantias técnicas, gestionárias e operacionais, tomando-se as medidas necessárias para os reduzir a um nível aceitável.

Pelo exposto, conclui-se que a decisão pela adoção ou não da Nuvem como modelo é inerente a cada organização e deve estar diretamente relacionada com o seu modelo de negócios, para além de aspetos como: custos, administração, localização geográfica, políticas de *compliance*, entre outros aspetos.

As normas referenciadas nesse estudo ajudam na avaliação e análise de riscos trazendo recomendações que são fontes de grande importância para as organizações, porém, a decisão de seus gestores na adoção ou não de serviços de Nuvem deve ser baseada, não somente em fatores financeiros, mas ter em conta os seus objetivos de negócio e os riscos associados às suas próprias decisões.

REFERÊNCIAS BIBLIOGRÁFICAS

Ali, M., Khan, S. U., & Vasilakos, A. V. (2015). Security in cloud computing: Opportunities and challenges. *Information Sciences*, 305, 357–383. <https://doi.org/10.1016/j.ins.2015.01.025>

Antonio, R., Monteiro, A., Guimarães, B., Meneses, É., Ledesma, F., Nunes, H., ... Ferreira, S. (2014). A Gestão Documental na Governança da Informação. *APDSI*, 357. Retrieved from http://www.apdsi.pt/uploads/news/id844/Gestão Documental 2014_20141111.pdf

Archer, J., Boehme, A., Cullinane, D., Kurtz, P., Puhlmar, N., & Reavis, J. (2011). *CSA Security Guidance for Critical Areas of Focus in Cloud Computing V3.0*. Retrieved from <http://www.cloudsecurityalliance.org/guidance/csaguide.v3.0.pdf>

Badger, L., Patt-corner, R., & Voas, J. (2012). NIST Cloud Computing Synopsis and Recommendations Recommendations. *Nist Special Publication*, 800(146), 81. <https://doi.org/2012>

Borglund, E. A. M. (2015). What About Trust in the Cloud? Archivists' Views on Trust. *Canadian Journal of Information and Library Science-Revue Canadienne Des Sciences De L Information Et De Bibliotheconomie*, 39(2), 114–127.

Bushey, J., Demoulin, M., & McLelland, R. (2015). Cloud Service Contracts: An Issue of Trust / Les contrats de service d'informatique en nuage: Une question de confiance. *Canadian Journal of Information and Library Science*, 39(2), 128–153. Retrieved from https://muse.jhu.edu/journals/canadian_journal_of_information_and_library_science/v039/39.2.bushey.html

Chaves, S. (2011). *A Questão dos Riscos em Ambientes de Computação em Nuvem*. Universidade de São Paulo.

Chou, D. C. (2015). Cloud computing risk and audit issues. *Computer Standards & Interfaces*, 42, 137–142. <https://doi.org/10.1016/j.csi.2015.06.005>

CSA. (2017). About CSA. Retrieved January 23, 2017, from <https://cloudsecurityalliance.org/about/>

Dias, J. M. F., Rodrigues, R. de C. M. C., & Pires, D. F. (2012). A Segurança De Dados Na Computação Em Nuvens Nas Pequenas E Médias Empresas. *Revista Eletrônica de Sistemas de Informação E Gestão Tecnológica*, 56–69. Retrieved from <http://periodicos.unifacel.com.br/index.php/resiget/article/view/287>

Ferreira, O. (2012). *O Nivel de implementacao de Cloud computing nas empresas portuguesas*. Universidade Portucalense Infante D.Henrique.

INFORMS. (1997). Prof. Ramnath Chellappa. Retrieved January 23, 2017, from <http://www.bus.emory.edu/ram/>

InterPARES Trust. (n.d.). Retrieved July 16, 2017, from https://interparestrust.org/trust/about_research/studies

ISO. (2014). Are you in control of your records? Retrieved January 24, 2017, from <http://www.iso.org/iso/news.htm?refid=Ref1830>

ISO31000-2009. (2009). ISO31000:2009 - Risk management: Principles and guidelines. Retrieved from <http://www.iso.org/iso/home/standards/iso31000.htm>

ISO 18128-2014. (2014). ISO/TR18128:2014 - Information and documentation -- Risk assessment for records processes and systems Abstract. Retrieved from http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=61521

Lark, J. (2015). *ISO 31000 - Risk management - a practical guide for SMEs* (Vol. 22). <https://doi.org/10.1093/rpd/ncr142>

Luo, J., Wu, M., Gopukumar, D., & Zhao, Y. (2016). Big Data Application in Biomedical Research and Health Care: A Literature Review. *Biomedical Informatics Insights*, 1–10. <https://doi.org/10.4137/BII.S31559.TYPE>

- Machovec, G. (2014). Consortia and Next Generation Integrated Library Systems. *Journal of Library Administration*, 54(5), 435–443. <https://doi.org/10.1080/01930826.2014.946789>
- Madrid, M. M. (2013). A study of digital curator competences: A survey of experts. *International Information and Library Review*, 45(3–4), 149–156. <https://doi.org/10.1016/j.iilr.2013.09.001>
- Massingham, P. (2010). Knowledge risk management: a framework. *Journal of Knowledge Management*, 14(3), 464–485. <https://doi.org/10.1108/13673271011050166>
- Mckemmish, S. (2013). Recordkeeping and Archiving in the Cloud . Is There a Silver Lining ? *INFuture*, 17–29. Retrieved from [http://infoz.ffzg.hr/INFuture/2013/papers/1-02 McKemmish, Recordkeeping and Archiving in the Cloud.pdf](http://infoz.ffzg.hr/INFuture/2013/papers/1-02%20McKemish,%20Recordkeeping%20and%20Archiving%20in%20the%20Cloud.pdf)
- Meel, P., & Grance, T. (2011). *NIST Definition of Cloud Computing*. Retrieved from <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>
- Noh, Y. (2015). Imagining Library 4.0: Creating a Model for Future Libraries. *Journal of Academic Librarianship*, 41(6), 786–797. <https://doi.org/10.1016/j.acalib.2015.08.020>
- Ostrzenski, V. (2013). Cloud Computing and Risk : A look at the EU and the application of the Data Protection Directive to cloud computing. *Infopreneurship Journal*, 1(1), 29–38. Retrieved from <http://www.infopreneurship.net>
- Pires, I. (2016). *Os arquivos organizacionais e a normalização da gestão de documentos eletrónicos : análise de normas nacionais e internacionais (2001-2016)*. Universidade de Coimbra. Retrieved from [https://estudogeral.sib.uc.pt/bitstream/10316/32941/1/DISSERTACAO APROVADA 09-09 ISABEL .pdf](https://estudogeral.sib.uc.pt/bitstream/10316/32941/1/DISSERTACAO%20APROVADA%2009-09%20ISABEL.pdf)
- Popovic, K., & Hocenski, Z. (2010). *Cloud computing security issues and challenges. MIPRO, 2010 Proceedings of the 33rd International Convention*.
- Poulo, L. B. N. (2013). *Cloud Computing for Digital Libraries*. University of Cape Town.
- Regalado, A. (2011). Who Coined Cloud Computing. *MIT Technology Report*. Retrieved from <https://www.technologyreview.com/s/425970/who-coined-cloud-computing/>
- Shaw, A. K. (2013). Cloud Computing for Libraries: An Economic Strategy. *International Conference on Academic Libraries*, 1(2), 162–167. Retrieved from [http://eprints.rclis.org/19359/1/023_12_Amit Kumar Shaw_34.pdf%0A](http://eprints.rclis.org/19359/1/023_12_Amit%20Kumar%20Shaw_34.pdf%0A)
- Sosa-Sosa, V. J., & Hernandez-Ramirez, E. M. (2012). A file storage service on a cloud computing environment for digital libraries. *Information Technology and Libraries*, 31(4), 34–45. <https://doi.org/10.6017/ital.v31i4.1844>
- Stancic, H., Rajh, A., & Brzica, H. (2015). Archival Cloud Services: Portability, Continuity, and Sustainability Aspects of Long-term Preservation of Electronically Signed Records. *Les Services D'archivage Dans Un Nuage Informatique: Portabilité, Continuité et Durabilité: Aspects de La Conservation À Long Terme Des Documents Signés Électroniquement.*, 39(2), 210–227. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=lls&AN=109028640&site=ehost-live&scope=site>

Vurukonda, N., & Rao, B. T. (2016). A Study on Data Storage Security Issues in Cloud Computing. *Procedia Computer Science*, 92, 128–135. <https://doi.org/10.1016/j.procs.2016.07.335>

Wilson, K. (2012). Introducing the Next Generation of Library Management Systems. *Serials Review*, 38(2), 110–123. <https://doi.org/10.1016/j.serrev.2012.04.003>

Zissis, D., & Lekkas, D. (2012). Addressing cloud computing security issues. *Future Generation Computer Systems*, 28(3), 583–592. <https://doi.org/10.1016/j.future.2010.12.006>



A CATEGORIZAÇÃO COMO ORGANIZAÇÃO DO CONHECIMENTO NA PERSPECTIVA DE CHAIM ZINS

Guilherme Ataíde Dias¹, Josiana Florêncio Vieira Régis de Almeida²

¹Universidade Federal de Minas Gerais, Programa de Pós-graduação em Gestão e Organização do conhecimento, 0000-0001-6576-0017, guilhermeataide@gmail.com

²Universidade Federal de Minas Gerais, Programa de Pós-graduação em Gestão e Organização do conhecimento, 0000-0002-0386-8724, josianavieira@gmail.com

RESUMO Apresenta uma contextualização da categorização com a indexação como instrumentos de estruturação do conhecimento universal, propondo a utilização do modelo denominado “Os 10 pilares do conhecimento” elaborado por Chaim Zins. Utiliza a revisão de literatura e a pesquisa exploratória como instrumentos metodológicos desta pesquisa, explorando e contextualizando a organização do conhecimento e sua estruturação com os modelos de categorização. Menciona a Ciência da Informação (CI), como área de conhecimento para exemplificar a implicação teórica e prática do assunto. Conclui que a estruturação deste conhecimento em CI possibilita a visibilidade global da área de acordo com a categorização dos conteúdos armazenados e uma melhor classificação dentro das linhas gerais do conhecimento contemporâneo da sociedade. Esta categorização facilita a compreensão e recuperação dos domínios do conhecimento utilizados na área de CI, além de permitir uma divulgação e compartilhamento das relações existentes entre as partes destes domínios.

PALAVRAS-CHAVE *Categorização. Indexação. Organização do conhecimento. Pilares do conhecimento. Chaim Zins.*

ABSTRACT It presents a contextualization of categorization with indexation as tools for structuring universal knowledge, proposing the use of the model called "The 10 pillars of knowledge" elaborated by Chaim Zins. It uses literature review and exploratory research as methodological tools of this research, exploring and contextualizing the organization of knowledge and its structuring with categorization models. It mentions Information Science (CI) as an area of knowledge to exemplify the theoretical and practical implication of the subject. It concludes that the structuring of this knowledge in CI makes possible the global visibility of the area according to the categorization of the stored contents and a better classification within the general lines of the contemporary knowledge of the society. This categorization facilitates the understanding of the domains of knowledge used in the IC area, as well as the dissemination and sharing of the existing relationships between the parts of these domains.

KEYWORDS *Categorization. Indexing. Organization of knowledge. Pillars of knowledge. Chaim Zins.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A imensa quantidade de informações disponibilizadas pelos instrumentos da produção tecnológica precisa de uma reconfiguração e de uma estruturação do conhecimento existente na fragmentação destas informações. Esta estruturação facilita uma busca eficiente dos domínios do conhecimento e suas relações coerentes entre as partes envolvidas.

Deste modo, Zins (2007), complementa que os mapas, estruturas e modelos de conhecimento publicados afetam nossos mapas cognitivos e moldam a maneira como percebemos o mundo e agimos nele. Em outras palavras, a construção do conhecimento é permanente e não se pode fundar em definições fechadas e específicas. Uma das características destas redes de significações como imagem do conhecimento é a sua natural historicidade, ou seja, a contínua metamorfose dos conceitos. Uma outra característica é a heterogeneidade do conhecimento. A imagem da rede de conhecimento continuamente nos lembra que os nós/significados são naturalmente heterogêneos, no sentido de que envolvem relações pertencentes a múltiplos conteúdos, a diversas disciplinas. As noções, os conceitos realmente relevantes sempre terminam por ultrapassar as fronteiras disciplinares (Machado, 2004).

O conceito de “conhecimento humano” tem dois significados: domínio subjetivo e domínio universal. Conhecimento no domínio subjetivo representa pensamentos na mente de um conhecedor individual – pensamentos que podem ser caracterizados como crenças verdadeiras, justificadas. Conhecimento no domínio universal representa descobertas e pensamentos acumulados e coletivos da humanidade ao longo da história – pensamentos e descobertas que foram representados, gravados (escritos) e documentados em documentos textuais (Zins, 2007).

O objetivo deste estudo é apresentar uma contextualização da categorização e a indexação como instrumentos de estruturação do conhecimento universal, propondo a utilização do modelo denominado “Os 10 pilares do conhecimento” elaborado por Chaim Zins, o qual utiliza as categorias como um método de organização dos domínios conceituais de uma determinada área. A problematização consiste na ausência de uma estrutura de conhecimento que disponibilize o conhecimento científico de forma simples e unificada para facilitar a recuperação da informação no âmbito da ciência da informação, por exemplo. Existem categorizações baseadas em indexações temáticas, mas que não seguem uma representação global do conhecimento. É necessário pensar numa estrutura em que a informação e o conhecimento possam ser integrados adequadamente, dirigindo uma maior atenção aos problemas de organização do conhecimento. Utilizou-se a revisão de literatura e a pesquisa exploratória como instrumentos metodológicos desta pesquisa, explorando e contextualizando a organização do conhecimento e sua estruturação com os modelos de categorização. Este estudo utilizou como área de conhecimento, a Ciência da Informação para exemplificar a implicação teórica e prática do assunto.

REFERENCIAL TEÓRICO

Dahlberg (2014) descreve que o nosso conhecimento se condensa em conceitos por seu conteúdo informativo. Os conceitos são, portanto, unidades de conhecimento e formam os elementos de sistemas de conhecimento. Com isso, define-se a *Organização do Conhecimento* como o processo de modelagem do conhecimento que visa a construção de representações do conhecimento. Esse processo tem por base a análise do conceito e de suas características para o estabelecimento da posição que cada conceito

ocupa num determinado domínio, bem como das suas relações com os demais conceitos que compõem esse sistema nocional (Brascher & Café, 2008).

A partir dessa organização do conhecimento criam-se ferramentas que apresentam a interpretação organizada e estruturada do objeto, chamados de SOC (Sistemas de Organização do Conhecimento). No contexto da Ciência da Informação, os SOC ou esquemas de representação do conhecimento são instrumentos conceituais que fazem a tradução dos conteúdos dos documentos originais e completos, para um esquema estruturado sistematicamente, que representa esse conteúdo, com a finalidade principal de organizar a informação e o conhecimento e, conseqüentemente, facilitar a recuperação das informações contidas nos documentos (Carlan & Medeiros, 2011).

As unidades de conhecimento precisam ser identificadas através de instrumentos de análise. Esta ação de identificar e descrever um documento de acordo com seu assunto é chamada “indexação” (Naves, 1996). O processo de indexação compreende dois estágios: a determinação do assunto; e a representação de conceitos por termos de uma linguagem de indexação. Para isso, recomenda que a identificação de conceitos seja feita obedecendo a um esquema de categorias existente na área coberta pelo documento (Fujita, 2003). Segundo Vickery (1980), o processo de indexação comporta ainda mais um estágio, o de sumarização ou síntese, além do conhecimento de mundo (experiências) do indexador - interpretação/análise cognitiva (Novo & Miranda, 2015).

O uso das categorias para a organização de conceitos em uma determinada área de interesse foi introduzido por Ranganathan no âmbito da documentação, a partir de sua teoria da classificação facetada, na qual utiliza a noção de categoria para a análise dos assuntos contidos nos documentos (Campos & Gomes, 2006).

Categorias e hierarquias são a melhor maneira de organizar o conhecimento para a recuperação de informação e descoberta de conhecimento através do processo cognitivo, pela razão óbvia de que a informação estruturada é mais facilmente recuperada do que as informações desorganizadas (Lima & Raghavan, 2014).

Hessen (2000) aborda que as categorias permitem a obtenção de uma visão da estrutura essencial do mundo, dos princípios de toda a realidade, referem-se aos objetos e às determinações presentes nesses objetos e sua validade universal representa a posição que elas ocupam no topo de toda a ordenação. Segundo Machado (2004), a própria pesquisa científica carece muitas vezes de uma visão abrangente relativa aos valores envolvidos na fragmentação do saber. Nesta concepção, se o conhecimento humano é, como ensina *Aristóteles*, uma reprodução dos objetos; são determinados e dotados de forma, então os conceitos fundamentais do conhecimento, as categorias, apresentam as determinações mais gerais dos objetos, as estruturas objetivas do ser (Hessen, 2000).

A categorização é um processo cultural e social de construção da realidade, que organiza conceitos, parcialmente baseados na psicologia do pensamento. Na categorização, o reconhecimento das similaridades e diferenças leva à criação de um conhecimento novo, pelo agrupamento de entidades, de acordo com as similaridades e diferenças observadas (Lima, 2010). A atividade mental é um dos fatores que possibilitam o desenvolvimento das categorias de um documento, conforme ilustrado na figura 1:

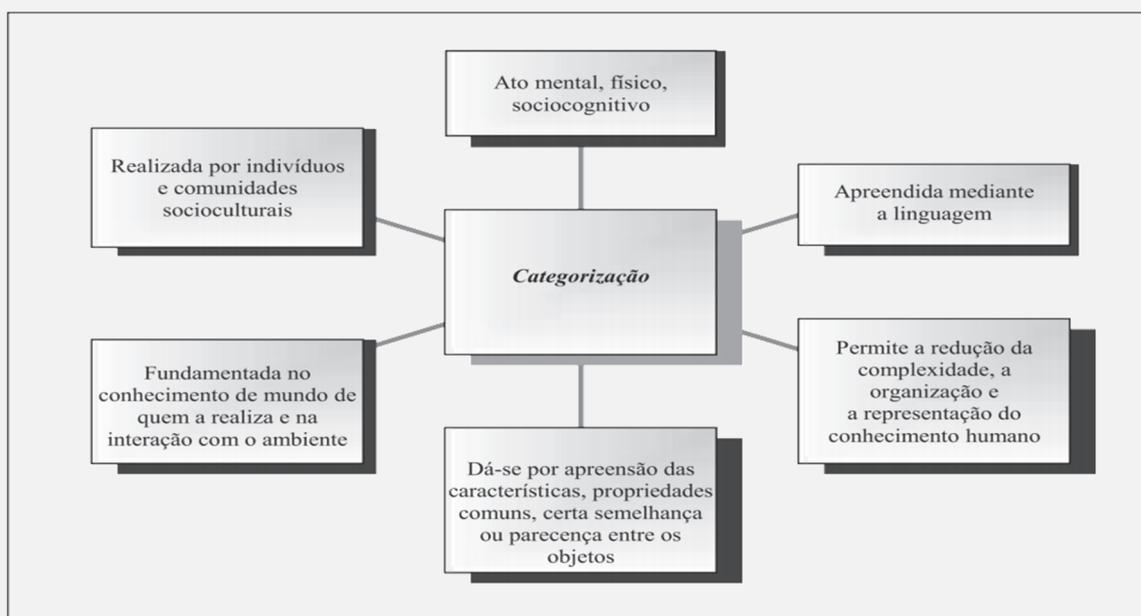


Figura 1: Elementos salientes nas definições de categorização

Fonte: Silva e Lima (2011)

As categorias, enquanto instrumentos de análise requerem um trabalho sistêmico e se instauram na própria compreensão de mundo dos sujeitos, transpostas ao universo dos registros do conhecimento. Para tanto, é imprescindível que se compreenda que uma categoria, de modo isolado, possui um alcance apenas fragmentário e incompleto, pois sua configuração é setorial e permite variados níveis de análise de um objeto – subcategorias. Sob essa perspectiva, cabe analisar se um conjunto de categorias eleitas pode propiciar um alcance total do objeto estudado, ou se existe uma categoria que comporte a dimensão do todo (Silva & Lima, 2011).

Corroborando com esta ideia, Jacob (2004) relata que um sistema de organização é composto por uma estrutura unificada que estabelece uma rede de relações entre as classes ou categorias que compõem o sistema. Essas relações são significativas informações de suporte, pois especificam conexões de princípios entre dois ou mais grupos dentro do mesmo sistema. Desta forma, apresenta-se uma síntese do processo de categorização, conforme a tabela 1:

Tabela 1: Processo de categorização

Categorização	
Processo	Síntese criativa de entidades com base no contexto ou semelhança percebida
Limites	A adesão em qualquer grupo não é obrigatória, as fronteiras são "difusas"
Associação	Flexível: A categoria associação é baseada no conhecimento generalizado e/ou contexto imediato
Crítérios de atribuição	Crítérios contexto-dependente e contexto-independente
Tipicidade	Os membros individuais podem ser classificados e ordenados pela tipicidade (estrutura graduada)
Estrutura	Agrupamentos de entidades; podem formar estrutura hierárquica

Fonte: Adaptado de Jacob (2004)

Observa-se que as categorias têm uma capacidade de estruturar/sistematizar todas as unidades do conhecimento de um domínio (Dahlberg, 1978). Consequentemente, aplicar a categorização é analisar o domínio a partir de recortes conceituais que permitem determinar a identidade dos conceitos

(categorias) que fazem parte deste domínio, servindo para orientar os profissionais no levantamento dos termos (Campos & Gomes, 2006).

Dentro da perspectiva da indexação e da categorização, pode-se considerar os *mapas do conhecimento* como um produto que fornece um panorama dos fundamentos e desenvolvimento da área e sugere um número de direções futuras de domínio e mapeamento. É uma poderosa ferramenta para esclarecer a terminologia básica e as várias concepções da CI. Sugere-se atualizar periodicamente nossos mapas do conhecimento (Zins, 2007).

Nesse contexto, Zins (2008; 2009) descreve a árvore denominada “Os 10 Pilares do Conhecimento” (Figura 2) como um mapa sistemático e exaustivo do conhecimento humano que apresenta, a estrutura do conhecimento e as relações coerentes entre os principais campos. Reflete a clareza, a simplicidade e o modo como o conhecimento é concebido no mundo. Chaim Zins dividiu o conhecimento humano em 10 pilares (Fundamentos, Sobrenatural, Matéria e energia, Espaço e terra, Organismos não-humanos, Corpo e mente, Sociedade, Pensamento e arte, Tecnologia e História) e 4 grupos (Conhecimento, Sobrenatural, Universo e Humanidade). O Mundo vivo corresponde aos Organismos não Humanos (pilar 5) e Corpo e Mente (pilar 6). Organismos não-Humanos é parte de universo, enquanto que Corpo e Mente é parte de humanidade, por isso estão inter-relacionados e são partes do mesmo fenômeno geral, o Mundo Vivo. Cada pilar é composto por categorias e sub-categorias que estabelecem conexões entre si.

As 10 categorias principais são divididas respectivamente em 32 subcategorias principais (2º nível). As 32 categorias são respectivamente divididas em 30 subcategorias (3º nível). Uma delas (categoria 7.3.1) é dividida em outras 4 subcategorias que formam o 4º nível de hierarquia. A estrutura de 4 níveis abrange 76 categorias e subcategorias. As 76 categorias sobrepõem-se parcialmente, formando 55 categorias de assunto únicas. As 55 categorias são mutuamente exclusivas e cobrem coletivamente o largo espectro do conhecimento humano no domínio universal (Santos & Zins, 2011).

Desta forma, elaborou-se uma síntese dos “10 pilares do conhecimento”, conforme apresentado na figura 2:

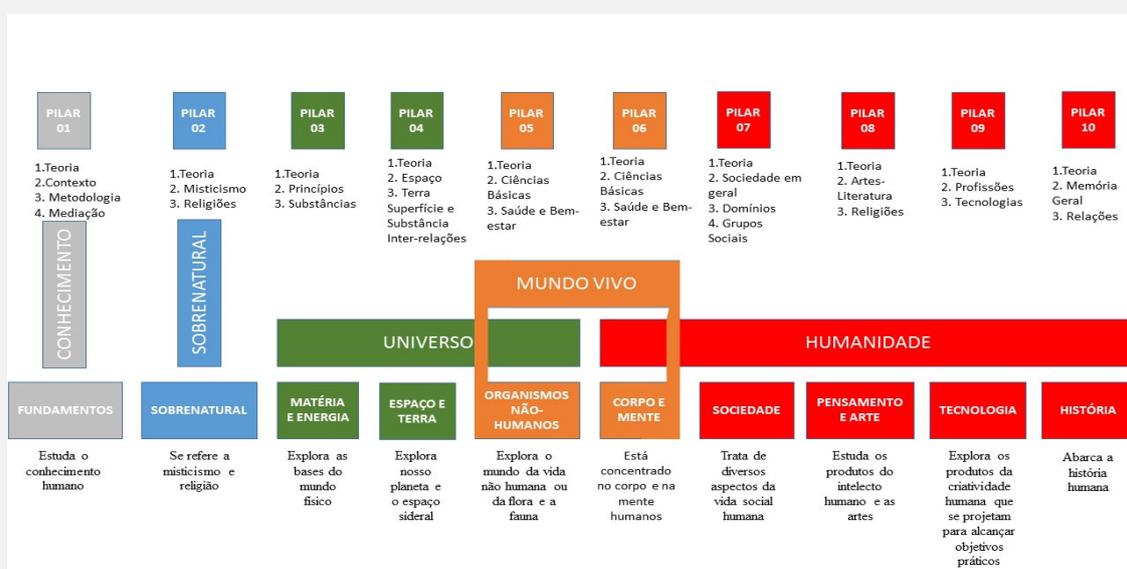


Figura 2: Síntese dos 10 pilares do conhecimento

Fonte: Adaptado de Zins (2008)

Observa-se que a categorização abre uma perspectiva de um rizoma para a horizontalidade da indexação. Assim, o conhecimento deve ser representado e organizado para além das relações arborescentes, pois as conexões cognitivas surgem naturalmente, conforme a necessidade do indivíduo em adquirir e produzir conhecimento, em determinados domínios, o que provoca relações rizomáticas (Miranda, 2006). Entende-se que o rizoma dentro da indexação dos conceitos remete à compreensão dos fluxos de conexão entre si ao abrir a possibilidade da inserção da subjetividade e singularidade dos domínios do conhecimento.

Além do processo rizomático, tem-se as árvores de conhecimentos como um instrumento de contribuição, condução e integração social dos indivíduos na sociedade. Visa construir uma rede que acompanhe o desenvolvimento e o registro das informações/conhecimentos e sua consequente troca entre os indivíduos (Levy & Authier, 2008).

O processo de estruturação por meio da categorização facilita a recuperação da informação pelos usuários. Zins (2000) relata que a recuperação da informação é um processo destinado a atender as necessidades de informação do pesquisador. Portanto, sua avaliação é principalmente com base em seu resultado. Uma análise fenomenológica da recuperação da informação revela em pelo menos cinco noções básicas que afetam os resultados da pesquisa: O pesquisador e suas capacidades, expertise e experiência em pesquisa; O processo cognitivo e o comportamento de busca; O domínio da informação; A tecnologia da informação; E as informações necessárias. Estes cinco fundamentos definem uma agenda para pesquisas em ciência da informação.

Zins (2007), faz alguns questionamentos acerca do campo da Ciência da Informação em seu artigo intitulado *Classification Schemes of Information Science: twenty-eight scholars Map the field*: Como os cientistas da informação estruturam o campo de Ciência da Informação? A literatura fornece milhares de Mapas de conhecimento mas nem todos são abrangentes e sistemáticos. Na verdade, a maioria deles é parcial, incompleto e inconsistente. Desta forma, observa-se a necessidade de obter uma estruturação que formule uma concepção sistemática da Ciência da Informação que deve basear-se em concepções dos conceitos constitutivos: dados, informações e conhecimento.

O processo de estruturação por meio da categorização facilita a recuperação da informação pelos usuários. Zins (2000) relata que a recuperação da informação é um processo destinado a atender as necessidades de informação do pesquisador. Portanto, sua avaliação é principalmente com base em seu resultado. Uma análise fenomenológica da recuperação da informação revela pelo menos cinco noções básicas que afetam os resultados da pesquisa: O pesquisador e suas capacidades, expertise e experiência em pesquisa; O processo cognitivo e o comportamento de busca; O domínio da informação; A tecnologia da informação; E as informações necessárias. Estes cinco fundamentos definem uma agenda para pesquisas em CI.

Zins (2007), faz um questionamento acerca do campo da CI em seu artigo intitulado *Classification Schemes of Information Science: twenty-eight scholars Map the field*: Como os cientistas da informação estruturam o campo de CI? A literatura fornece milhares de Mapas de conhecimento mas nem todos são abrangentes e sistemáticos. Na verdade, a maioria deles é parcial, incompleto e inconsistente. Desta forma, observa-se a necessidade de obter uma estruturação que formule uma concepção sistemática da CI que deve basear-se em concepções dos conceitos constitutivos: dados, informações e conhecimento.

Com base na distinção entre o subjetivo e os domínios universais de dados, informação e Conhecimento, a CI concentra-se no domínio universal. Está focada na perspectiva de meta-conhecimento do

conhecimento universal. A CI é o estudo das perspectivas mediadoras do conhecimento humano universal (ou seja, conhecimento humano no domínio universal). As perspectivas mediadoras incluem aspectos e condições cognitivas, sociais e tecnológicas, que facilitam a disseminação do conhecimento humano do originador para o usuário (Zins, 2007).

Os objetos podem ser facilmente representados por meio dos mecanismos de representação do conhecimento. Campos (2004) faz uma relação da representação do conhecimento nas três áreas do conhecimento: CI, Ciência da Computação e na Terminologia. No âmbito da ciência da computação, eles servem para auxiliar a implementação de estruturas computáveis. No âmbito da CI, possibilitam a elaboração de linguagens documentárias verbais e notacionais, visando a recuperação da informação e a organização dos conteúdos informacionais de documentos. No âmbito da terminologia, esses mesmos mecanismos permitem a sistematização dos conceitos e, conseqüentemente, a elaboração de definições consistentes.

É importante perceber de onde se deve partir para a elaboração de um modelo de representação. Alguns estudos esclarecem sobre as bases teóricas e metodológicas a respeito da modelização de domínios do conhecimento. Campos (2004) estabelece quatro princípios fundamentais para orientar o modelizador na construção de estruturas conceituais: o método de raciocínio; o objeto de representação; as relações entre os objetos e as formas de representação. É comum, inicialmente, o modelizador ter dúvidas sobre qual metodologia se deve utilizar para modelar um determinado domínio. Neste contexto, estes quatro princípios norteiam o modelizador no encontro de um modelo comum de conceitos no ato de modelar o conhecimento pretendido.

PERCURSO METODOLÓGICO

Trata-se de uma revisão de literatura e de um estudo exploratório existente sobre o assunto. Segundo Babbie (2015), grande parte da pesquisa exploratória é conduzida para explorar um tópico, ou seja, começar a familiarizar um pesquisador com esse tópico. Esta abordagem ocorre quando um investigador examina um novo interesse ou quando o assunto de estudo em si é relativamente novo. Um dos propósitos dos estudos exploratórios é satisfazer a curiosidade e desejo de uma melhor compreensão da realidade.

Utiliza-se a apresentação do modelo de estruturação do conhecimento proposto por Chaim Zins, denominado “Os 10 pilares do conhecimento humano” como uma forma de organização e categorização do conhecimento universal humano. Baseia-se na análise teórica e qualitativa a respeito da organização do conhecimento e da informação com ênfase na indexação. Abordou-se, primeiramente sobre a importância do conhecimento para a sociedade e em seguida, descreveu-se a organização do conhecimento sob a forma de categorização tendo-se em vista a perspectiva de Chaim Zins no tocante aos pilares do conhecimento registrado por ele. Além disso, registra-se a indexação como um instrumento de análise dos conteúdos. Esta metodologia de revisão de literatura permite-se conhecer e contextualizar a importância da organização e estruturação do conhecimento universal para obter-se um melhor acesso e visibilidade dos conteúdos registrados.

DISCUSSÕES

Com o considerável aumento no volume de informações disponíveis, torna-se necessário encontrar soluções para simplificar este fluxo informacional. Para tanto, a informação deve ser pensada e tratada como um bem comum a todos e precisa ser acessível sem necessitar de auxílio de um especialista. Esta nova forma de categorização proposta pelos *pilares do conhecimento*, possibilita uma maior abrangência dos campos do conhecimento humano e suas conexões, permitindo ao usuário obter a informação mais eficiente e eficaz no que diz respeito à classificação global da ciência em uma determinada área.

Pode-se inferir que o método de categorização e o a tarefa de análise conceitual dos documentos, são considerados os elementos norteadores para a organização do conhecimento em um dado domínio. Sendo assim, a representação do conhecimento registrado é realizada através de uma lista de assuntos que são categorizados em termos. Os termos sustentam conceitos e os combinam numa declaração de assuntos, de acordo com um padrão de relações. Isto é, cada assunto tem sua própria estrutura. As unidades, ou seja, conceitos ou ideias isoladas identificadas na descrição de um assunto particular, são pontos na estrutura ou redes de relações. A estruturação de conceitos componentes em um assunto desempenha um papel vital na comunicação, aprendizagem e memória (Vickery, 1980).

Conforme a revisão de literatura abordada neste estudo, entende-se que a categorização com base nos domínios do conhecimento permite uma recuperação da informação mais simples e de forma unificada para obtenção de uma resposta mais abrangente a qualquer área do conhecimento. Um dos pontos positivos deste mapa (10 pilares do conhecimento) é que ele não privilegia nenhuma área, pois engloba todos os níveis de conhecimento. Com este mapeamento, por exemplo, podemos observar em quais pilares a Ciência da informação tem obtido maiores produções científicas no Brasil e verificar também onde há maior carência destas produções. No âmbito teórico e prático, a utilização dos “10 pilares do conhecimento humano”, permite também, por exemplo, traçar o perfil da produção do conhecimento referente à área de CI diante da universalidade deste mapeamento, criando uma alternativa para a avaliação destes conhecimentos.

CONCLUSÕES

A estruturação deste conhecimento em CI possibilita a visibilidade global da área de acordo com a categorização dos conteúdos armazenados e uma melhor classificação dentro das linhas gerais do conhecimento contemporâneo da sociedade. Esta categorização facilita a compreensão dos domínios do conhecimento utilizados na área de CI, além de permitir uma divulgação e compartilhamento das relações existentes entre as partes destes domínios. A contextualização teórica e apresentação do mapa do conhecimento proposto por Chaim Zins com a categorização e indexação dos conteúdos possibilita uma nova exigência para classificação de qualquer área do conhecimento, visto que ocasiona uma oportunidade de obtenção de maior visibilidade e recuperação da informação registrada no ambiente digital.

Conclui-se que a ideia deste estudo apresenta uma possível reflexão sobre a forma como as estruturas do conhecimento humano são concebidos na CI. Assim, o mapa do conhecimento pode resolver os problemas de atinência, pertinência, precisão, relevância, cobertura, transparência e flexibilidade no ambiente virtual ao permitir a organização do conhecimento visando a recuperação da informação. A relevância prática deste estudo para a Ciência da informação infere-se na representação do conhecimento humano produzido no âmbito desta área, possibilitando aos docentes, discentes,

profissionais e pesquisadores, uma análise e classificação da realidade objetiva dos documentos e suas conexões.

REFERÊNCIAS BIBLIOGRÁFICAS

- Babbie, E. R. (2015). *The practice of social research* (14a ed.). Boston, MA: Cengage Learning US.
- Bräscher, M. & Café, L. (2008). Organização da informação ou organização do conhecimento. *Anais do Encontro Nacional de Pesquisa em Ciência da Informação*, 9 (pp. 1-14). São Paulo, SP.
- Campos, M. L. de Almeida. (2004). Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. *Ciência da Informação*, 33(1).
- Campos, M. L. A. & Gomes, H. E. (2006, setembro/dezembro). Metodologia de elaboração de tesauro conceitual: a categorização como princípio norteador. *Perspect. Ciênc. Inf.*, 11 (3), 348-359.
- Carlan, E. & Medeiros, M. B. B. (2011, agosto/dezembro). Sistemas de Organização do Conhecimento na visão da Ciência da Informação. *RICI: R.Ibero-amer. Ci. Inf.*, 4 (2), 53-73.
- Dahlberg, I. (1978, julho/dezembro). Teoria do conceito. *Ci. Inf.*, 7 (2), 101-107.
- Dahlberg, I. (2014, January). Brief Communication: What is Knowledge Organization?. *Knowl. Organ.*, 41 (1), 85-91.
- Fujita, M. S. L. (2003, julho/dezembro). A identificação de conceitos no processo de análise de assunto para indexação. *RDBCI: Rev. Digit. Bibliotecon. Cienc. Inf.*, 1 (1), 60-90.
- Hessen, J. (2000). *Teoria do conhecimento*. São Paulo: Martins Fontes.
- Jacob, E. K. (2004). Classification and categorization: a difference that makes a difference. *Library Trends*, 52 (3), 515-540.
- Lévy, P. & Authier, M. (2008). *As árvores de conhecimentos*. São Paulo: Editora Escuta.
- Lima, G. Â. B. O. (2010). Modelos de categorização: apresentando o modelo clássico e o modelo de protótipos. *Perspect. Ciênc. Inf.*, 15 (2), 108-122.
- Lima, G. Â. B. O. & Raghavan, K. S. (2014). Categories in Knowledge Organization. *Advances in Knowledge Organization*, 14, 88-95.
- Machado, N. J. (2004). *Conhecimento e valor*. São Paulo: Editora Moderna.
- Miranda, M. L. (2006). Organização e representação do conhecimento: fundamentos teóricos-metodológicos na busca e recuperação da informação em ambientes virtuais. *Anais do Encontro Nacional de Pesquisa em Ciência da Informação*, 7. Marília, SP.
- Naves, M. M. L. (1996, julho/dezembro). Análise de assunto: concepções. *Rev. Bibl. Brasília*, 20 (2), 215-226.

Novo, I. F. & Miranda, J. G. V. (2015, julho/dezembro). Organização do conhecimento na perspectiva do modelo dinâmico de análise conceitual. *Tend. Pesqui. Bras. Cienc. Inf.*, 8 (2), 37-58.

Silva, A. R. & Lima, G. Â. B. O. (2011). As categorias e a categorização cognitiva contemporânea: enfoque sobre os olhares da biblioteconomia e ciência da informação e das ciências cognitivas. *Rev. Interam. Bibliotecología*, 34 (3), 301-312.

Vickery, B. C. (1980). *Classificação e indexação nas ciências*. Rio de Janeiro: BNG, Brasilart.

Zins, C. (2000). Success, a structured search strategy: Rationale, principles, and implications. *Journal of the Association for Information Science and Technology*, 51(13), 1232-1247.

Zins, C. (2009). 10 pillars of knowledge: Map of human knowledge. In: G.M. Braga & L. V. R. Pinheiro. *Desafios do impresso ao digital: questões contemporâneas de informação e conhecimento* (pp.169-176). Brasília: IBICT.

Zins, C. (2008). *10 pillars of knowledge: map of human knowledge*. Recuperado em 02 de junho, 2016 de <http://www.success.co.il/knowledge/map/map.pdf>.

Zins, C. et al. (2007). Knowledge map of information science: implications for the future of the field. *Braz. J. Inf. Sci.*, 1 (1), 3-32.



INTERACCIÓN TECNOLÓGICA Y FUNDAMENTOS EPISTEMOLÓGICOS EN ORGANIZACIÓN DEL CONOCIMIENTO

Rosa San Segundo¹, Maria Adelina Codina-Canet²

¹Universidad Carlos III de Madrid, rosa.sansegundo@uc3m.es

²Universidad Carlos III de Madrid, mcodina@db.uc3m.es

RESUMEN En el mundo actual inmerso en las innovaciones tecnológicas, que afectan especialmente a los entornos informativos y comunicativos, emergen nuevos ecosistemas en los que los usos de la información ocupan un lugar destacado. Los actuales entornos de interacción tecnológica, conducen a una reflexión sobre los fundamentos epistemológicos, los procesos y los productos, así como sobre los modelos de formación y de desarrollo profesional. En el presente trabajo se aborda el tránsito que ha significado el cambio tecnológico y su incidencia en las nuevas formas de acceso a la cultura y al conocimiento. En la historia se han generado tres cambios por el modo de adquisición de conocimientos: La escritura, la imprenta, la digitalización. La primera se produjo con la implantación de la escritura, que permitió fijar con signos las informaciones que antes debían ser registradas en la mente. La segunda comenzó con la invención de la imprenta, que acercó el libro a un público extenso que antes no tenía acceso a este recurso y al mismo tiempo lo consolida como símbolo de conocimiento. Y la tercera viene producida por un cambio que comenzó a gestarse en la década de los años 80, con el inicio de la revolución electrónica y la difusión de nuevos medios de comunicación. Las tecnologías del conocimiento han generado nuevas denominaciones de *homo digitalis*, *homo connectatus* u *homo ciber* suponen ya un nuevo concepto y estadio en la era de la humanidad.

PALABRAS CLAVE *Interacción tecnológica, Epistemología, Organización del Conocimiento, Escritura, Lectura digital.*

ABSTRACT In today's world, immersed in technological innovations that especially affect informational and communicative environments, new ecosystems in which the uses of the information play a prominent role emerge. The current environments of technological interaction lead to a reflection on the epistemological foundations, processes and products, as well as on the training and professional development models. The present paper addressed the transition of these technological changes and their impact on the new forms of access to culture and knowledge. There are three main changes that have been generated by the way of acquiring knowledge in history: writing, the printing press, digitization. The first change occurred with the introduction of writing, which allowed to establish with signs the information that had to be recorded in the mind. The second change began with the invention of the printing press, which brought books closer to a wider audience that previously had no access to these resources and at the same time consolidated them as a symbol of knowledge. The third change began in the decade of the 80s, with the beginning of the electronic revolution and the dissemination of new media. The technologies of knowledge have generated the new denominations of *homo digitalis*, *homo connectatus*, or *homo ciber* that already mean a new concept and stage in the era of humanity.

KEY WORDS *Technological Interaction, Epistemology, Knowledge Organization of Knowledge, Writing, Digital Reading.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUCCIÓN:

La escritura, la imprenta y la digitalización son tres hitos que tienen en común, que en ellos se lee y se escribe. Así estas operaciones vinieron realizándose comúnmente durante siglos y fueron nuestra fuente de cultura y transmisión de conocimientos. Sin embargo, en los últimos años, el cambio es de mayor envergadura. En la actualidad el volumen de cosas que se conoce porque se ha leído es mucho menor que hace años, ya que han cobrado mayor protagonismo la televisión, Internet, diversos formatos multimedia, el cine, la radio, u otras. Este cambio está teniendo efectos profundos en el contenido de los conocimientos y en la forma en que están organizados. Estos conocimientos están menos articulados y son menos sutiles (Simone).

Se plantea si se ha producido una degradación cualitativa del saber, a pesar de que quizás sólo haya cambiado la forma material de su archivo. Es posible que con las nuevas modalidades de conocimiento se lleguen a activar nuevos módulos o funciones de la mente; al mismo tiempo, funciones que hemos tenido activadas durante siglos, dejarán de estar activadas. La nueva visión alfabética es la modalidad de la visión que permite adquirir informaciones y conocimientos a partir de una serie lineal de símbolos visuales, ordenados unos tras otros. Ésta se desarrolló a partir de la invención de la escritura y configuró la inteligencia secuencial, que opera en una sucesión de estímulos, que es diferente de la simultánea. Ésta está basada en imágenes y sonidos que presentan varios estímulos al mismo tiempo. Puede plantearse que la inteligencia secuencial es más evolucionada, ya que tiene que ser educada, entrenada y sirve de base a distintas formas de actividad mental que pueden llegar a ser muy complejas.

Sin embargo, en la era digital con la ingente cantidad de estímulos visuales y auditivos que recibimos cotidianamente ha perdido importancia esta visión alfabética y su soporte más usual, el texto. Así, renunciamos a la conquista evolutiva de esta visión alfabética y la inteligencia secuencial para regresar a unos medios más naturales, más complacientes más primitivos, en los últimos años se ha producido un cambio cultural y las tecnologías de la información han sido, en muchos sentidos, el motor de esos cambios. Ya que además de la estructura textual, como el orden del texto y la discontinuidad semántica también han afectado muchos otros aspectos, que inciden en nuestra forma de conocer el mundo, con gran incidencia en las formas de interacción con información y objetos digitales y lo que incide en que se está modificando nuestra forma de adquirir conocimientos y de generar conocimientos nuevos.

Esta tercera fase de los regímenes cognitivos está en sus inicios, en la que aún convivimos con los universos de la escritura y de la imprenta y al mismo tiempo nos distanciamos de ellos (Piscitelli, 2005). Ya es adecuado hablar del *homo videns*, que pertenece que este tercer estadio, y comparte en común mucho de la Edad Media. La era digital a menudo se compara con la Edad Media, como vislumbró Umberto Eco, con el advenimiento de un universo más opaco, más ininteligible y más vulnerable con la tecnología. Así se hace necesaria una crítica de la razón electrónica, el mundo digital modifica nuestra forma de conocer. Siguiendo la propuesta anticipadora de McLuhan, en tanto que fue precursor de los análisis actuales acerca nuestras mentes, afirma que nuestra realidad está transida por la linealidad del alfabetismo, será necesario un aprendizaje nuevo para pensar y percibir de manera no lineal (Kerckhove,

1999b). La nueva realidad de las tecnologías de la información hará que todos nuestros intercambios e interacciones estén mediadas por lo virtual, nuevos valores puramente simbólicos conforman ya la realidad, a modo de ejemplo en el contexto de lo económico, basado en un concepto material del dinero, opera donde ya no es necesaria su presencia.

La especie humana siempre ha tenido gran fascinación por el cambio tecnológico, lo que además ha supuesto un cambio y avance de gran envergadura, El homo digitalis, homo connectatus, homo ciber suponen ya un nuevo concepto de estadio de la humanidad, de era. El cambio tecnológico conlleva control sobre las actividades comunicativas, produciendo una socialización transversal. Entramos en una sociedad que explota las capacidades tecnológicas con cierta disolución de la ciudadanía sustituida por una identidad con subjetividades con perfil de usuario. Se ha pasado de la vertebración social de la ciudadanía a un nuevo sujeto social político, situado en el espacio publico, pero el espacio social no va a ser lo mismo que espacio político. El entorno del ciberespacio va acaparando e incluyendo todas las experiencias de vida, se presenta más accesible, y además se completa con la traslación de las anteriores relaciones personales, ligadas a numerosos rituales, ceremonias, fiestas y otros, a la sustitución por las relaciones virtuales. El cambio tecnológico conlleva control sobre las actividades comunicativas, produciendo una socialización transversal. Se están creando nuevos resortes para una nueva civilización que supone una ruptura con el pasado. El vehículo de estos cambios son los nuevos medios tecnológicos de información y comunicación, las nuevas tecnologías de la información digital están determinando el paso a una nueva era de la historia de la humanidad.

Estamos conformando una mente tecnocrática donde hay un culto a la información, donde las tecnologías de la información en nuestra sociedad están inmensas en una mitificación, lo que puede comportar un deterioro para el pensamiento, para la vida privada, y puede articular los fundamentos de un estado con deficiencias democráticas, pues existen, existen limitaciones de la tecnologías digitales. (Roszak, 2005). Las grandes teorías y hallazgos científicos no pueden ser fruto sólo de un ordenador, considera que los ordenadores son herramientas magníficas que también pueden convertirse en un instrumentos que generen el divorcio entre la razón y la imaginación, y además generando prosélitos de un culto exagerado a la tecnología.

El nativo digital se nos presenta como una persona dependiente de su hipnotizador digital e indiferente respecto a la diversidad cultural más extensa que agoniza por efecto de la digitalidad. Se hace necesario el análisis y la crítica tecnológica, frente a al inmersión de máquinas y aplicaciones cada vez más totalizadoras y adictivas que impregnan el mundo laboral, educativo, político, comunicativo y, también, el de las culturas, imaginarios y creencias. Se ciernen así, sobre el planeta, el desmantelamiento remoto de prácticas culturales como nunca antes conoció la historia, un sometimiento de la diversidad simbólica por parte del poder económico, financiero, político, militar, mediático (García Gutiérrez, 2017), por lo que se hace necesario su análisis.

Con el advenimiento de la civilización de la pantalla, que incluye televisión y ordenadores, puede venir el inevitable colapso de la alfabetización libresco, que puede cuestionar distintos logros de la cultura letrada, en tanto que se ignora que existe cierta violencia implícita que toda tecnología del conocimiento lleva consigo en los planos cognitivo y socioeconómico. Ya que la palabra está dejando de ser el vehículo transmisor de información y conocimiento que había sido durante milenios, es importante considerar como ofrecer sustitutos de aprendizaje efectivo y analítico a través de las tecnologías de la información, y no solo utilizar los recursos informáticos para producir un aprendizaje ligero y superficial. A partir de las dos vertientes de la tradicional inteligencia secuencial y actual simultánea se

han de desarrollar interfaces, podemos encontrar un justo medio que explote ambos tipos de aprendizaje y de acceso a la información. Al mismo tiempo, no se puede responsabilizar, en exclusiva, a los medios audiovisuales y en especial a la televisión de estas formas de imposición (Piscitelli, 1998), pues no es viable proclamar de forma aislada el poder apolíneo de la escritura y de la lectura, en contraposición de la dimensión dionisiaca de la imagen y de los medios audiovisuales.

Además las nuevas tecnologías digitales son tecnología de la participación. La web no solo requiere nuevas formas de aprender y de conocer sino de participar. Y ha aparecido una cultura participativa con blogs, wikis, redes sociales. La web social, se caracteriza por el desarrollo de tecnologías orientadas a la participación y colaboración entre comunidades virtuales. Algunas de las herramientas desarrolladas han permitido establecer redes sociales que conforman comunidades en donde los usuarios interactúan

Estamos conformando una inteligencia digital colectiva, nos encontramos ante multitudes con inteligencias vinculadas, con gran acceso a ingentes cantidades de información. Esta multitud inteligente emerge con las tecnologías de la comunicación que vincula los talentos humanos en cooperación. Las tecnologías que articulan la multitud inteligente son los teléfonos móviles, tables , ordenadores, automoviles ya interconectados y toda una gran información que se trasmite a través de aparatos de uso cotidiano.

La interacción mundial de inteligencias individuales en conexión conlleva s consecuencias, esta impresionante conexión mundial avanza a convertirse en una red global de "inteligencias individuales en conexión". El empleo de nuevas capacidades cognitivas y la interacción global contempla la posibilidad de que en el presente estamos experimentando uno de los mayores saltos en la evolución de nuestra especie. La conexión digital de inteligencias bien podría ser el paso próximo en la evolución de la inteligencia humana. Interacción de multitud inteligente. En el siglo XXI, el agente del cambio socioeconómico, disciplina científica por excelencia será la tecnología y la información digital, esto implica cierto determinismo tecnológico como ley universal.

En la interconexión global de información semántica se puede producir un patrón exponencial de crecimiento tecnológico, como expuso Raymond Kurzweil. Raymond Kurzweil estudió la comprensión de los ritmos y direcciones de los desarrollos tecnológicos, sugiriendo un patrón exponencial de progreso tecnológico que persiste a través de la historia humana. esta predicción ha hecho la tasa de innovación en tecnologías de la computación que crece de un modo no lineal, sino exponencial. Este crecimiento exponencial de las capacidades de los ordenadores, supuestamente, dará lugar a nuevas tecnologías fantásticas, que podrían estar disponibles antes de lo imaginado con las actuales expectativas de crecimiento tecnológico lineal, este desarrollo tecnológico viene generado en tanto que la ciencia y la tecnología se convierten en información. De igual forma que los avances tecnológicos en el mundo de los ordenadores darán lugar a máquinas cada vez más potentes, numerosas y baratas. A partir de lo cual propone Kurzweil su teoría de la singularidad tecnológica (Kurzweil, 2012). Dicho evento se ha nombrado así por analogía con la singularidad gravitacional observada en los agujeros negros, donde existe un punto en el que las reglas de la física dejan de ser válidas, y donde la convergencia hacia valores infinitos hace imposible el definir una función.

Teoría que plantea la aparición en el futuro de una inteligencia superior a la humana como consecuencia del acelerado desarrollo de la tecnológico, esta proliferación de la tecnología se ha convertido en un motor del rápido cambio tecnológico Según Kurzweil, plantea un patrón que culmina en un progreso tecnológico inimaginable en el Siglo XXI, el cual conduce a la Singularidad. Lo describe en su obra *La singularidad está cerca: una historia real basada en el futuro*, donde incluye su concepto de

crecimiento exponencial, su perspectiva de expansión radical de la vida y como trascenderemos nuestra biología

Segun Kurzweil, una vez llegado al punto en que se cree una inteligencia superior a la humana, se entraría en una etapa post humana que probablemente conduzca a la extinción de la humanidad o a su subordinación a esos nuevos entes inteligentes. Kurzweil predice que un ordenador pasará el test de Turing hacia el 2029, demostrando tener una mente (inteligencia, consciencia de sí mismo, riqueza emocional...) indistinguible de un ser humano. Esse momento es el denominado singularidad tecnológica. Kurzweil prevé que la primera inteligencia artificial será construida alrededor de una simulación por ordenador de un cerebro humano (Kurzweil, 2012).. Usando el concepto de singularidad de Kurzweil como base, la universidad podrá planear y proporcionar a su alumnado las destrezas y herramientas para guiar el proceso de la singularidad "para el beneficio de la humanidad y su entorno".

Los ordenadores alcanzarán la capacidad del ser humano y serán insertados en nuestro cuerpo y en prácticamente todo lo que nos rodea (cámaras, grabadoras y otros aparatos diminutos estarán por todas partes). Se generalizarán los "implantes neuronales" que mejorarán la visión, la memoria y el razonamiento, y la nanotecnología permitirá introducir en el torrente sanguíneo máquinas que permitirán mayor control sobre las enfermedades a través de la descarga directa de software. Más tarde, la separación entre realidad virtual y "real" se hará cada vez menos clara y nuestras mentes podrán "copiarse", llegando el día, allá por el 2099, en que "nuestros cerebros serán mayoritariamente no-biológicos", tal y como sostiene Kurzweil: "seremos capaces de saber todo conectando nuestros cerebros a –qué casualidad– Google". Ya en plena singularidad, la fusión entre nuestras mentes con inteligencias artificiales a través de implantes neuronales hará que los humanos y las máquinas se mezclen en el mundo físico y mental. La mayoría de los seres conscientes carecerán de una forma física permanente.

La última publicación de Ray Kurzweil, *Cómo crear una mente. El secreto del pensamiento humano*, incluye una serie de experimentos mentales para comprender mejor cómo pensamos los humanos. También presenta un modelo del neocórtex y plantea su Teoría de la Mente basada en el Reconocimiento de Patrones, con un análisis de las diferentes partes biológicas del cerebro y su evolución, para entonces discutir cómo sería un neocórtex digital, creado gracias al crecimiento acelerado de la tecnología.

En Cómo crear una mente defiende que la mente es una "propiedad emergente" del cerebro, de manera que la creación de cerebros digitales resultará en la creación de mentes digitales. De hecho, el cerebro, actual sustrato biológico de la mente humana, puede ser sustancialmente mejorado gracias a sustratos no biológicos cuidadosamente diseñados y mucho más avanzados. Kurzweil no sólo considera que la mente es una consecuencia directa del cerebro, sino que además las inteligencias artificiales tendrán consciencia, libre albedrío y hasta identidad propia. Aborda las nuevas tecnologías y el cerebro humano, la estructura más compleja del universo conocido.

Los humanos, en una etapa pos humana o bien en una etapa anterior nos constituimos ya en ciborg, en tanto que organismo cibernético (Haraway, 1991), híbrido de máquina y organismo, una criatura de realidad social, que incluso hoy se constituye ya con la tecnología de la información, de la tablet, el ordenador, el automóvil, las múltiples prótesis, el smartphone. El cyborg es, a la vez, epistemología y ontología, una propuesta de sujeto y de abordar el análisis de la realidad, y también de conocimiento. El débil cuerpo humano y su pequeña memoria se sobredimensionan con múltiples tecnologías. El discurso

del cuerpo, del conocimiento y de la realidad ya se han desarrollado mediante un nuevo conjunto de tecnologías. La red es inevitablemente el nuevo espacio epistemológico y existencial del ser.

El nuevo modelo totalizante de interacción con la tecnología determina las nuevas epistemologías. En la actualidad, se ha llegado a un modelo totalizante tecnológico de interacción que determina e inunda el sustrato de las epistemologías. Nuestro cerebro se adapta a lo digital. La estructura de la información que ha estado mediada por la oralidad y por la escritura y ahora lo va a estar por la información digital, la interconexión total, la incorporación en todos los ámbitos de tecnología. El medio conforma ya no sólo el mensaje, sino que parece conformar la estructura, e incluso conforma ya la realidad.

Lo digital constituye nuestro espacio cotidiano e identitario, Internet se inicia como tecnología, también como discurso, y más tarde como imaginario, como espacio identitario, como comunicación, también para el ámbito comercial, empresarial, económico, militar, universitario y otros muchos. Los paradigmas iniciales válidos para explicar los fenómenos emergentes que surgían de la interacción fueron muy útiles en un primer momento, hoy ya son insuficientes. Internet es ya parte constitutiva de nuestra interacción, memoria, comunicación y cognición.

METODOLOGÍA

Se plantea una metodología analítica para abordar nuevas formas de acceso a los conocimientos, mediante la lectura y la información digitales, fundamentando su naturaleza, formas de proceder y características más esenciales para explicar, hacer analogías, comprender y, finalmente, establecer perspectivas de las implicaciones de la futura lectura e interacción digitales.

La hipótesis de partida evoca que las distintas tecnologías siempre nos han hecho desarrollarnos y evolucionar en muchos aspectos, inclusive como especie. Esta tecnología en pos de la información parece que ya contempla la posibilidad de que estemos iniciando y experimentando uno de los nuevos y mayores saltos en la evolución de nuestra especie. La relevancia del tema es notoria, en tanto que el nuevo acceso, la virtualización, la interacción, y la conexión digital global de inteligencias ya apunta a convertirse en el siguiente estadio en la evolución de la inteligencia humana, incluso conformándose el organismo cibernético, cibor.

RESULTADOS

La palabra está dejando de ser el vehículo transmisor de información y de conocimiento que había sido durante milenios, por lo que es importante plantear sustitutos de aprendizaje efectivo y analítico a través de las tecnologías de la información, y no solo utilizar los recursos informáticos para producir un aprendizaje que pueda ser más ligero y superficial. A partir de las dos vertientes de la inteligencia secuencial y simultánea se han de desarrollar interfaces, que abarque ambos tipos de aprendizaje y de acceso a la información (Simone, 2001).

CONCLUSIONES

Las nuevas tecnologías de la información no son comparables con los instrumentos tecnológicos anteriores, ni con la escritura ni con la imprenta. La era y el futuro digital ya han comenzado y ello incide en los diferentes pasos evolutivos de los seres humanos en relación con la lectura y el procesamiento de información, el homo digitalis ya está conformado. El nuevo espacio tecnológico articula las capacidades cognitivas, establece el ordenamiento simbólico, si el lenguaje inventó un orden, la tecnología digital lo ha determinado y estructurado aun más, vertebrando una ciudadanía digital y conformará a los humanos cibor. Las tecnologías del conocimiento se han de articular en un desarrollo y utilización de la manera más eficiente, sin vulneración de los derechos fundamentales, socialmente responsable, y asunción del progreso epistemológico.

REFERENCIAS BIBLIOGRÁFICAS

BELKIN, Nicholas J. (1990). The cognitive viewpoint in information science. *Journal of Information Science*, 16: 11-15.

BOYD, Danah. (2014). *It's complicated: the social lives of networked teens*. New Haven: Yale University Press

BREA, José Luis. (2007). *Cultura RAM. Mutaciones de la cultura en la era de su distribución electrónica*. Barcelona: Gedisa. http://www.joseluisbrea.net/ediciones_cc/c_ram.pdf

CARR, Nicholas. (2011). *Superficiales. ¿Qué está haciendo internet con nuestras mentes?*. México DF: Taurus.

COBO ROMANÍ, Cristóbal; PARDO KUKLINSKY, Hugo. (2004). *Planeta 2.0. Inteligencia colectiva o medios fastfood*. Grup de Recerca d'Interaccions Digitals, Universitat de Vic. Flacso México. Barcelona / México DF, 2007. E-book de acceso gratuito. Versión 0.1 / Septiembre de 2007. <http://www.planetawebdospuntocero.net>

COHEN, Elizabeth. (2011). *Does life online give you 'popcorn brain'?* <http://www.cnn.com/2011/HEALTH/06/23/tech.popcorn.brain.ep/index.html>

DEHAENE, Stanislas. (2014). *El cerebro lector*. Buenos aires: Siglo XXI.

GARCIA, Felipe; PORTILLO, Javier, ROMO, Jesús; BENITO, Manuel. (2005). *Nativos digitales y modelos de aprendizaje*. Universidad de País Vasco / Euskal Herriko Unibertsitatea. <http://spdece07.ehu.es/actas/Garcia.pdf>

GARCÍA CANCLINI, Néctor. (2007). *Lectores, espectadores e internautas*. Barcelona: Gedisa.

GARCÍA GUTIÉRREZ, Antonio Luís. (2002). *La Memoria Subrogada: Mediación, Cultura y Conciencia en la Red Digital*. Granada. Universidad de Granada.

GARCÍA GUTIÉRREZ, Antonio Luís. (2017). *Frentes digitales. Totalitarismo tecnológico y transcultural*. Comunicación Social

HARAWAY, Donna.(1991). *Ciencia, cyborgs y mujeres. La reinención de la naturaleza. Madrid: Cátedra.*

KERCKHOVE, Derrick de. (1999^a). *Inteligencias en conexión. Hacia una sociedad de la web.* Barcelona, Gedisa.

KERCKHOVE, Derrick de. (1999b). *La piel de la cultura. Investigando la nueva realidad electrónica.* Barcelona: Gedisa

KURZWEIL, Ray. (2012). *La singularidad está cerca.* Lola Books.

KURZWEIL, Ray. (2013). *Cómo crear una mente. El secreto del pensamiento humano.* Lola Books, <http://es.scribd.com/doc/182862491/Como-Crear-Una-Mente>

LESSIG, Lawrence. (2001). *El código y otras leyes del ciberespacio.* Madrid, Taurus.

PARRA, Sergio. (2010). *¿Cómo nació la lectura?* <http://www.xatakaciencia.com/biologia/como-nacio-la-lectura>

PISCITELLI, Alejandro. (1998). *Post-Televisión. Ecología de los medios en la era de Internet.* Buenos Aires: Paidós.

PISCITELLI, Alejandro. (2002). *Ciberculturas 2.0: en la era de las máquinas inteligentes.* Buenos Aires: Paidós.

PISCITELLI, Alejandro.2005. *Internet: la imprenta del siglo XXI.* Barcelona: Gedisa. <http://terras.edu.ar/aula/tecnicatura/6/biblio/6PISCITELLI-Alejandro-Ecologia-de-la-Red.pdf>

QUEAU, Philippe.(1995). *Lo virtual. Virtudes y vértigos.* Barcelona: Paidós.

REIG, Dolors; VÍLCHEZ, Luis F. (2013). *Los jóvenes en la era de la hiperconectividad: tendencias, claves y miradas.* Madrid: Fundación Telefónica.

RHEINGOLD, Howard. (1994). *Realidad virtual. Los mundos generados por ordenador que modificarán nuestras vidas.* Barcelona: Gedisa.

RHEINGOLD, Howard, (2004). *Multitudes inteligentes. La próxima revolución social.* Barcelona, Gedisa.

ROSZAK, Theodore. (2005). *El culto a la información. Un tratado sobre alta tecnología, inteligencia artificial y el verdadero arte de pensar.* Barcelona: Gedisa. <http://es.scribd.com/doc/181189928/Roszak-Theodore-El-culto-a-la-informacion-1986-pdf>

SAN SEGUNDO, Rosa. *A New Conception of Representation of Knowledge.* En: KNOWLEDGE Organization. International Journal. , Vol 31, 2004, n.2, p. 106-111, ISSN 0943-7444 http://74.125.155.132/scholar?q=cache:Mjw7iAM0dHYJ:scholar.google.com/&hl=es&as_sdt=0&as_vis=1

SAN SEGUNDO, Rosa. *Nuevas formas de lectura, conocimiento y ciudadanía en el entorno digital.* En: María de la Paz Pando Ballesteros, Alicia Muñoz Ramírez y Pedro Garrido Rodríguez (eds. y dirs.) *Pasado y presente de los derechos humanos. Mirando al futuro.* Madrid, Catarata, 2016

SAN SEGUNDO, Rosa ; MARTÍNEZ ÁVILA, Daniel *New Conceptual Structures for the Digital Environment: From KOS to the Semantic Interconnection*. En: *Categories, Contexts and Relations in Knowledge Organization*. 13, pp. 206 - 211. Ergon-Verlag, 2012.

SARTORI, Giovanni. 2012. *Homo videns: La sociedad teledirigida*. Madrid: Taurus
https://docs.google.com/document/d/1y5d7aTr_Shmj41UzHpB2Be5_uDxReMygErHrfHhSDro/edit?pli=1

SCOLARI, Carlos. 2009. Desfasados Formas de conocimiento que estamos perdiendo, recuperando y ganando. *Versión*, 22: 163-185.<http://es.scribd.com/doc/79777096/Desfasados-Las-formas-de-conocimiento-que-estamos-perdiendo-recuperando-y-ganando>

SIMONE, Raffaele. 2001. *La tercera fase: formas de saber que estamos perdiendo*. Madrid: Taurus.

SPARROW, Betsy; LIU, Jenny, WEGNER, Daniel M. 2011. Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, 33 (6043):776-778
STINSON, Liz. 2013. *Shadow: A Beautiful App That Tracks Your Dreams*. <http://www.wired.com/2013/09/shadow-a-beautiful-app-that-helps-you-remember-your-dreams/>

STRICKLAND, Eliza. 2014. *Building Mind-Controlled Gadgets Just Got Easier*.
<http://spectrum.ieee.org/biomedical/devices/building-mindcontrolled-gadgets-just-got-easier>

SUROWIECKI, James. 2004. *The wisdom of crowds : why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*. New York: Doubleday.

VIRILIO, Paul. 1999. *La bomba informática*. Madrid, Cátedra.



MAPAS CONCEPTUALES PARA MODELAJE CONCEPTUAL DE ONTOLOGIAS DE DOMINIO

Gracy Kelli Martins¹, Denysson Axel Ribeiro Mota²

¹Universidade Federal do Cariri 0000-0002-1805-9292, gracy.martins@ufca.edu.br

²Universidade Federal do Cariri, 0000-0003-2069-134X, denysson.mota@ufca.edu.br

RESUMEN Presenta el rescate de literatura cuanto a la utilización de teorías que auxilien en la construcción de un mapa conceptual, que oriente el modelaje conceptual en el desarrollo de ontologías de dominio. El proceso de mapeo se apoyó en el análisis de dominio, realizando levantamientos que involucraron cuestiones que van desde la determinación de la delimitación adecuada, hasta la captura de los términos y conceptos contenidos en las fuentes de información sobre el dominio analizado. Mucho más de carácter teórico, esta pesquisa tensionó describir como el levantamiento de determinado dominio debe ser conducido a partir de su aspecto de representación del conocimiento, y como sucede el establecimiento de las relaciones entre los términos y los conceptos, etapa preliminar en el desarrollo de ontologías de dominio y objetivo central de este trabajo. No se buscó desarrollar una ontología en el ámbito de la inteligencia artificial, a través de la lógica computacional, y tampoco desarrollar una nueva metodología, pero el recate y aplicación de los métodos a mucho trabajados por la Ciencia de la Información (CI) y hoy necesarios para las tecnologías de representación y recuperación de la información. Fundamentada a partir de la fuerte característica interdisciplinar de la CI, la propuesta tiene como base los Principios Fundamentales para Modelaje de Dominios (CAMPOS, 2004) y representaciones del conocimiento de forma estructurada en la representación gráfica de los mapas conceptuales de Novak (1995).

PALABRAS CLAVE Mapa Conceptual, Modelaje Conceptual, Organización y Representación del Conocimiento, Ontologías de Dominio.

ABSTRACT This work presents a rescue of literature related to the use of theories that help in the construction of a conceptual map, which guides conceptual modeling in the development of domain ontologies. The mapping process was based on the domain analysis, carrying out surveys that involved situations ranging from the determination of the appropriate delimitation, to the capture of the terms and concepts contained in the sources of information about the domain under analysis. Much more of a theoretical nature, this research stressed how the survey of a given domain should be driven from its aspect of knowledge representation, and how happens the establishment of relations between terms and concepts, preliminary stage in the development of domain ontologies and central objective of this work. It was not sought in this paper to develop an ontology in the field of artificial intelligence, through computational logic, nor to develop a new methodology, but the retrieval and application of methods to widely used by Information Science (IS) and today required for information representation and retrieval technologies. Based on the strong interdisciplinary character of the IS, the proposal is based on the Fundamental Principles for Domain's Modeling (CAMPOS, 2004) and representations of knowledge in a structured way in the graphical representation of Novak's concept maps (NOVAK, 1995).

KEYWORDS Concept Maps, Conceptual Modelling, Knowledge Organization and Representation, Domain Ontologies.

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUCCIÓN

Las conexiones entre la Ciencia de la Información (CI), Ciencia de la Computación (CC) y los estudios cognitivos han estrechado lazos en la busca por herramientas y procesos que mejoren los sistemas de información, principalmente los dirigidos a la recuperación de información. Ontologías de dominio es un área de estudios explorada por la CC, desarrollada a partir de la inteligencia artificial. La CI, por otro lado, ya posee significativa realización de investigaciones sobre las ontologías de dominio, como también su producción científica se ha intensificado al redor de este asunto.

En gran parte, los estudios describen las etapas para el desarrollo de ontologías de dominio, y como estas deben ser hechas. Ese proceso es constituido de las etapas de especificación, adquisición de conocimiento, conceptualización, formalización, integración, implementación y mantenimiento de la ontología creada (ALMEIDA, 2006, p.118). De esas etapas, es unanimidad en la literatura que la parte de mapeo de dominio, que consiste en la adquisición de conocimiento, no presenta detalles para que se comprenda como es realizado dicho procedimiento (SALES, 2006; SILVA; SOUZA; ALMEIDA, 2008; OYOLA; ALVARENGA, 2009, SALES; SAYÃO, 2012; CAMPOS, 2015).

En este trabajo, por ser de naturaleza más teórica, no se buscó desarrollar una ontología en el ámbito de la inteligencia artificial, por medio de la lógica computacional, mas proponer como el levantamiento de dominio puede ser conducido a partir de su aspecto de representación de conocimiento. Fundamentado a partir de la fuerte característica interdisciplinar de la Ciencia de la Información, se buscó una base metodológica a través de los criterios de modelaje conceptual de dominios, propuesto por Campos (2004, p. 25-31) y la representación gráfica a partir de un mapa conceptual, con la metodología de Novak (1995; 2000).

Así, el objetivo central de este trabajo fue realizar un estudio teórico y metodológico de la representación de conocimiento, con el objetivo de obtener subsidios sobre la construcción de un mapa conceptual de una ontología de dominio, para exponer el proceso de levantamiento de dominio y representarlo gráficamente. El mapeo envolvió cuestiones que van desde la determinación del recorte adecuado, hasta la captura de los términos y conceptos constantes en las fuentes de información y las relaciones entre los conceptos generales del dominio.

ONTOLOGIAS DE DOMINIO Y MODELAJE CONCEPTUAL

La internet es hoy el mayor repositorio de informaciones de todos los tiempos, con constante avance, posibilitando que cualquier individuo, de cualquier parte del mundo, tenga acceso a los más diversificados contenidos, estando conectado a través de un computador (LIMA-MARQUES, 2006, p.13). Debido a ese gran volumen de informaciones, identificase un número cada vez mayor de problemas de acceso a la información. Algunos indicadores de esos problemas pueden ser apuntados por el crecimiento acentuado de fuentes de información digitalizadas, y la creciente búsqueda por áreas del conocimiento en hacer disponible información, observando la volatilidad de esos datos, una vez que

la actualización es uno de los principios garantidos por ese medio de comunicación (FEITOSA, 2006, p.14).

Desde el surgimiento de la *web*, los procesos son realizados en un contexto sintáctico. En la *Web*, los computadores hacen apenas la presentación de la información, pero el proceso de interpretación es hecha por seres humanos (BREITMAN, 2005, p. 2, traducción libre). No que eso haya cambiado, pero hoy los esfuerzos están direccionados para una mayor integración entre la generación, acceso y uso de las informaciones vinculadas a la red, con el intento de que esas interpretaciones sean realizadas, en parte, por las maquinas.

La constitución de la *Web* como espacio de comunicación, creación de informaciones e intercambio de conocimientos proporcionó el surgimiento de grupos sociales que se auto organizan, y desencadenan un proceso de comunicación a través de la reproducción de elementos compositores y transformadores de información dentro del sistema de organización social (STROCKINGER, 2001, p. 109).

Delante de las inúmeras posibilidades de utilización de la información a través de la *Web*, y con la necesidad cada vez más acentuada por una organización y recuperación relevante de esas informaciones, investigadores de la inteligencia artificial vienen proponiendo una serie de modelos para categorizar la información de modo estandarizado (BRITMAN, 2005, p.6). La búsqueda por dicha estandarización motivó comunidades académicas, que han desarrollado prototipos variados, buscando soluciones informacionales y tecnológicas. Entre diversos estudios en ese ámbito, se destacan las ontologías de dominio.

Las ontologías surgieron en el área de la Inteligencia Artificial, en la década de 1990, en un contexto llamado de relaciones semánticas. Esta área tiene por objetivo embutir el conocimiento en la máquina, a través de la intervención humana. Mejor definiendo, lo que la inteligencia artificial busca es capacitar el computador para la comprensión de actividades que solamente un ser humano sería capaz de realizar (SALES, 2006, p. 19, traducción libre).

Hoy es posible observar en la literatura una extensa producción de investigaciones que proponen metodologías para la construcción de ontologías. Sin embargo, no hay estandarización para las etapas del proceso de desarrollo de ontologías, que debe obedecer la presencia de componentes básicos como clases, relaciones, funciones, axiomas e instancias. Feitosa (2006, p. 75) hace uso de las colocaciones de Gómez-Pérez (1999) para esclarecer que “com a utilização desses componentes, as ontologias fornecem um vocabulário comum para uma área e definem o significado dos termos de tal vocabulário, bem como as relações entre eles”. Reiterando tales colocaciones, Guarino (1995) resalta que existe una preocupación más direccionada para los sistemas inteligentes como esquemas de buscas, que precisamente para el modelaje conceptual desarrollada para ellos.

El desarrollo y la pesquisa en el área de la ontología viene adquiriendo gran importancia y substancial interés, con objetivos básicos, buscando evitar minimizar la sobrecarga de informaciones y organizarlas para que sean recuperadas y ofrezcan acceso adecuado, atendiendo al interés de los usuarios (LIMA-MARQUES, 2006, p. 23). Guarino (1995, p. 4) muestra que los estudios sobre ontologías y modelaje conceptual deben necesariamente ser estudiados bajo una perspectiva interdisciplinar, para más allá de las herramientas, de la lógica y de la CC, buscado también estudios en áreas como Filosofía y aquellas que tratan de las relaciones entre las lenguajes naturales y sentido común.

Para la construcción de ontologías, la industria de la Inteligencia Artificial cuenta con métodos y dispositivos computacionales que simulan la capacidad humana de resolver problemas, y de esa forma ordenan el dominio escogido para la aplicación de la ontología, a través del uso del computador y de sus software, responsables por coordinar tareas de búsqueda, comparación y negociación en la red, reduciendo enormemente el esfuerzo realizado por el usuario (BREITMAN, 2005, p. 154, traducción libre).

Las pesquisas direccionadas para las ciencias cognitivas, que unen disciplinas como Psicología, Lingüística, Inteligencia Artificial, Filosofía, Educación y Ciencia de la Información, muestran que el interés por los modelos cognitivos ha sido cada vez mayor. Problemas relacionados a la representación, información, comunicación y conocimiento son fundamentales para todas estas disciplinas. La Ciencia de la Información, teniendo como objeto de estudio la información, ha utilizado recursos metodológicos y los cuadros conceptuales de esas mismas disciplinas, creando un cuerpo de conocimiento multidisciplinar. Sayão (2001, p. 87) afirma, también, que

Desde que a matéria-prima dos processos cognitivos são “objetos mentais”, tais como conceitos, idéias e conhecimento, muitos dos trabalhos nestas áreas consideram como as pessoas organizam conhecimento, como os conceitos são formados na mente humana, como as pessoas agrupam objetos em suas mentes, ou seja, como as pessoas os categorizam, quais são as teorias que as pessoas têm sobre como o mundo funciona.

Entre los modelos existentes en la ciencia cognitiva y la Ciencia de la Información, los modelos conceptuales de datos conocidos, fueron creados como herramientas de representación, que funcionan en ambientes específicos, y, así siendo, la literatura no reconoce ningún modelo generalizado (SAYÃO, 2002, p. 89 traducción libre).

Campos (2004, p. 23) presenta, a partir de una investigación sistemática de los métodos y teorías para modelos conceptuales, cuatro principios fundamentales que pueden ser utilizados en el modelaje de dominios de conocimiento y que se presentan como una propuesta para estructuración de modelajes conceptuales:

- El primer principio es relacionado al método de raciocinio utilizado para la organización del conocimiento dentro de un dominio;
- El segundo analiza cómo está definido el objeto de representación, o sea, cual es la unidad de conocimiento que se va representar;
- El tercero es direccionado a la relación entre los objetos, con el objetivo de verificar las posibilidades de ligación/separación semántica entre los conceptos de un determinado dominio;
- El cuarto muestra las formas de representación gráfica que un modelo puede adoptar.

En este contexto, la ontología de dominio hace parte de la construcción de mecanismos de representación y recuperación de información a través de modelos conceptuales, que presente mayor precisión en las respuestas para las solicitudes informacionales. Lo que los actuales sistemas de recuperación de la información realizan es la búsqueda con base en una orden sintáctica, donde no tienen la capacidad de diferenciar sinonimias y atribuir precisión semántica en el intercambio de informaciones con diferentes bases virtuales, como afirma Breitman (2005, p. 2-3, énfasis del autor):

A grande verdade é que a Internet se desenvolveu mais rapidamente como meio para troca de documentos entre **pessoas**, em vez de um meio que fomentasse a troca de dados e informações que

pudessem ser **processadas automaticamente**. Como consequência, o conteúdo semântico das páginas, ou seja, seu significado é codificado de uma maneira acessível para seres humanos apenas.

De este modo, pensar el modelaje conceptual de una ontología de dominio utilizando herramientas como los mapas conceptuales, permiten que este modelaje refleje la estructura cognitiva humana y sus formas de ordenación y la aprehensión de informaciones a partir de la organización de conceptos y sus relaciones.

LAS RELACIONES ENTRE CONCEPTOS Y MAPAS CONCEPTUALES

La teoría de los mapas conceptuales de Joseph D. Novak (1995) tiene como base la Teoría del Aprendizaje Significativa de David Ausubel (1968), que tiene como principio guía que la aprehensión de informaciones se da a partir de la organización de conceptos y sus relaciones de modo jerárquico, partiendo de los más generales para los más específicos (NOVAK, 1995, p. 113).

Novak (1995, p. 51) llama la atención para las variaciones en los significados formados a través de los mapas, que pueden surgir de sus ligaciones, dando origen a nuevas relaciones y significados entre los conceptos. Presenta como ejemplo la conexión entre los conceptos de agua y hielo que pueden estar conectados por expresiones como: *puede ser, transformase en, algunas veces es*, y generan significados parecidos, pero no idénticos. Podrán ocurrir variaciones más notorias si, de acuerdo con el ejemplo del autor, acrecentar al mapa el concepto de *molécula*, donde podrá originar, de esta forma, nuevas relaciones y nuevos significados que incluyan el hielo, el agua y las moléculas. “Assim, os mapas conceituais são instrumentos poderosos para observar as alterações de significado que um estudante dá aos conceitos que estão incluídos no mapa” (NOVAK, 1995, p. 51).

De acuerdo con las exposiciones de Novak (1995), un mapa conceptual es una estructura modelada a partir de conceptos generales y específicos que son relacionados por preposiciones, para representar un determinado dominio, siendo entendido como un modelo de representación visual de las redes de conceptos y sus significados.

Para el desarrollo de la metodología de los mapas conceptuales Novak (1995, p.113, traducción libre) hizo uso de tres de las principales ideas de la Teoría del Aprendizaje Significativo: la estructura cognitiva es *organizada jerárquicamente*, con los conceptos más específicos subordinados a los conceptos más generales y amplios; los conceptos de la estructura cognitiva están sujetos a una *diferenciación progresiva*, acompañada del reconocimiento de mayor cobertura y especificidad en la proporción de los objetos, y nuevas relaciones con otros conceptos; la *reconciliación integradora* ocurre cuando dos o más conceptos son relacionados en términos de nuevos significados cuando se resuelven conflictos de significados entre conceptos.

Así el mapa desarrollado deberá presentar un recorte bien establecido para el uso del elemento gráfico permitiendo la comprensión de la organización del conocimiento y modelaje conceptual. Organizar el conocimiento, en este sentido, en lo que se refiere al ámbito de la CI, consiste en dividir y agrupar las representaciones de la información, de acuerdo con principios y reglas prefijadas, dotadas de significados dentro de una jerarquía construida a partir de un sistema de clasificación adecuado para parte del conocimiento y realidad institucional, donde será realizada la acción de ordenar, representar y clasificar.

Considerando que el estudio tiene como foco el trabajo direccionado a los conceptos, la identificación de los conceptos, así como sus relaciones, las teorías adoptadas deben tener direccionamiento para coleta, garantido que esta hará uso de garantía literaria, valiéndose de las ocurrencias en la literatura y pesquisas direccionadas.

Para Novak (1995, p. 97-98, traducción libre) un buen mapa conceptual debe presentar tres aspectos importantes: 1) exponer los conceptos y las preposiciones (frases de ligación) fundamentales en lenguaje explícito y conciso; 2) presentar las relaciones entre las ideas principales de forma simple y vistosa, aprovechando la notable capacidad humana de representación visual; 3) realzar tanto las relaciones jerárquicas entre los conceptos y las proposiciones como los enlaces cruzados entre grupos de conceptos y proposiciones.

En el ámbito de las Ontologías, para el proceso de ordenación conceptual y, consecuentemente, modelaje conceptual, estas etapas se centran en “conversão da linguagem natural para a linguagem artificial, garantindo consistência na estrutura terminológica e no estabelecimento de inferências, e garantindo a adoção de algoritmos consistentes para a interpretação pela máquina” (SALES; SAYÃO, 2012, p. 4). Para atender esta estructura, la CI cuenta con teorías y metodológicas consolidadas para construcción de instrumentos de representación de la información y del conocimiento. Y, en la esfera del modelaje conceptual, identificar los tipos de relaciones presentes en un dominio puede ser considerada una etapa del modelaje conceptual de este dominio, pues son las relaciones que enlazan los conceptos a otros conceptos, permitiendo evidenciar la abstracción de una realidad.

Considerando los principios de Campos (2004) y analizando la teoría de mapas conceptuales en la perspectiva de Novak (1995), inferimos que estos no son apenas instrumentos, pero toda una fundación teórica capaz de reflejar modelos conceptuales, ya que parten de la idea de que el conocimiento puede ser representado por intermedio de mapas comunicantes que, cuando desdoblados, representan la estructura interna de un dominio.

Esos tres enfoques, a saber Principios para Modelaje de Dominios del Conocimiento (CAMPOS, 2004), Mapas Conceptuales (NOVAK, 1995) y Aprendizaje Significativa (AUSUBEL, 1968 apud NOVAK, 1995), se alían para mejor orientación en el proceso de modelaje conceptual, como podemos observar en la Tabla 1.

Tabla 1. Teorías para Desarrollo de Modelaje Conceptual

CAMPOS	AUSBEL	NOVAK
El primer principio es relacionado al método de raciocinio utilizado para la organización del conocimiento dentro de un dominio.	La estructura cognitiva es organizada jerárquicamente, con los conceptos más específicos, subordinados a los conceptos más generales y amplios.	Exponer los conceptos y las preposiciones (frases de ligación) fundamentales en lenguaje explícita y concisa.
El segundo analiza cómo está definido el objeto de representación, o sea, cual es la unidad de conocimiento que se va representar.	Los conceptos de la estructura cognitiva están sujetos a una diferenciación progresiva, acompañada del reconocimiento de mayor cobertura y especificidad en la proporción de los objetos, y nuevas relaciones con otros conceptos.	Presentar las relaciones entre las ideas principales de forma simple y vistosa, “aprovechando la notable capacidad humana de representación visual”.
El tercero es direccionado a la relación entre los objetos, con el	A reconciliación integradora ocurre cuando dos o más conceptos	Resaltar tanto las relaciones jerárquicas entre los conceptos y

objetivo de verificar las posibilidades de ligación/separación entre los conceptos de un determinado dominio.	son relacionados en términos de nuevos significados cuando se resuelven conflictos de significados entre conceptos.	las proposiciones como los enlaces cruzados entre grupos de conceptos y proposiciones.
El cuarto muestra las formas de representación gráfica que un modelo puede adoptar.		

De este modo, la CI contribuye significativamente con su tradición en investigaciones en Organización y Representación del Conocimiento que aliada a los mapas conceptuales, como teoría para ordenación de ideas y representación de las relaciones entre los conceptos de la estructura cognitiva, posibilitan mejor orientación para la construcción del modelaje conceptual, siendo esta etapa primordial en el desarrollo de ontologías de dominio.

CONSIDERACIONES FINALES

El cuestionamiento que movió este estudio provenía del interés en identificar teorías y técnicas capaces de orientar la construcción de modelajes conceptuales en el ámbito de las ontologías de dominio e CI. Mucho ha sido publicado sobre las metodologías para el desarrollo de ontologías, pero entre diversos estudios, evidenció como la teoría de los mapas conceptuales se muestra como fuerte aliada para optimizar este proceso. El desarrollo de formas representacionales de los dominios del conocimiento exige, además de la adopción de técnicas especializadas, recursos informacionales y una relación interdisciplinar para contemplar cuestiones actuales de recuperación de información.

La CI presente contribuciones para la construcción de ontologías con fuertes bases teóricas, direccionadas para la estructuración y relación conceptual. Fue posible percibir la clara pertinencia de la adopción de criterios para modelaje conceptual de dominios y los mapas conceptuales para la selección y relación semántica del dominio, proceso a priori simbólico y de inferencia humana, proporcionando un trabajo analítico-sintético en el intento de representación cognitiva para las relaciones conceptuales. Los sistemas computadorizados perfeccionan las técnicas de representación y recuperación informacional, pero para el desarrollo de un lenguaje lógico, dentro de la inteligencia artificial, es necesario el tratamiento de los contenidos informacionales a través de modelos que intenten reproducir la ordenación de conceptos, semejantes a las órdenes del pensamiento.

Se constató que las relaciones conceptuales son pasible de aplicación en el modelaje de ontologías y estimulan la continuidad de los estudios y pesquisas, pues la tecnología avanza a grandes pasos y demuestra clara necesidad de adecuación de metodologías direccionadas para la organización de informaciones en todos los niveles de representación del conocimiento. A nivel de representación gráfica, el modelaje conceptual posibilita la profundización de nuevos conceptos, generando el enlace con diferentes redes identificadas a través de sus preposiciones (conexiones), que describen la jerarquía conceptual, y sus facetas de todo-parte, en las relaciones partitivas (conceptos dentro de los conceptos).

No se buscó aquí el desarrollo de nuevas metodologías, pero el rescate y aplicación de métodos hace mucho trabajados por la CI y hoy necesarios para las tecnologías de la CC. Sin embargo, lo que se tensionó fue un estudio para la identificación de teorías para el desarrollo de modelos conceptuales en el proceso de la construcción de ontologías. Como afirmado anteriormente, no se intentó proponer algo nuevo, pero contribuir con los métodos de conocimientos construidos hasta el momento. Las

exposiciones teóricas relacionadas al modelaje conceptual aquí descritas visan contribuir para las discusiones del tratamiento conceptual y sus relaciones semánticas, pero reconociendo que en algunos aspectos dejan de ser abordados, no por ser considerados menos importantes, pero por la extensión y necesidad de delimitación del tema en foco. Esperase con este paso dado en el intento de aplicación de las teorías mencionadas, tener contribuido para la validación de los métodos existentes en la CI, y aplicables en ontologías de dominio, considerando que muchos caminos han sido trillados para alcanzar la unificación por parte de las metodologías ya existentes y futuras, pero aún hay mucho por recorrer.

REFERÊNCIAS BIBLIOGRÁFICAS

Almeida, M. B. (2006). *Um modelo baseado em ontologias para representação da memória organizacional*. Belo Horizonte: Universidade Federal de Minas Gerais. Retrieved from http://www.enancib.ppgci.ufba.br/premio/UFMG_Almeida.pdf

Breitman, K. (2005). *Web semântica: a internet do futuro*. LCT: Rio de Janeiro: LTC.

Campos, M.C. (2015). Ontologia para o relato de experimentos em um domínio científico: questões e aplicações. . *Encontro Nacional de Pesquisa em Ciência da Informação*, João Pessoa, PB, Brasil, 16.

Campos, M. L. A. (2001). *A organização de unidades do conhecimento em hiperdocumentos: o modelo conceitual como um espaço comunicacional para realização da autoria*. (Tese de doutorado) Universidade Federal do Rio de Janeiro, Rio de Janeiro.

Campos, M. L. A. (2004) *Modelização de domínios de conhecimento: uma investigação de princípios fundamentais*. *Revista Ciência da Informação*, Rio de Janeiro, 33 (1). Retrieved from <http://revista.ibict.br/index.php/ciinf/article/viewArticle/56/52>

Guarino, N. (1995). *Formal ontology, conceptual analysis and knowledge representation*. *International Journal of Human and Computer Studies*, 43 (5-6), pp. 625-640.

Lima, V. M. A. (2004) *Da classificação do conhecimento científico aos sistemas de recuperação de informação: enunciação de codificação e enunciação de decodificação da informação documentária*. (Tese de Doutorado). Universidade de São Paulo, São Paulo. Retrieved from <http://www.teses.usp.br/teses/disponiveis/27/27143/tde-06032006-150120/>

Lima-Marques, M. (2006) *Ontologias: Da filosofia à representação do conhecimento*. Brasília: Thesaurus.

Novak, J. D.& Gowin, D. Bob. (1995). *Aprender a aprender*. Lisboa: Plátano Edições Técnicas.

Novak, J. D.& Gowin, D. Bob. (2000). *Aprender, criar e utilizar o conhecimento. Mapas conceituais como ferramentas de facilitação nas escolas e empresas*. Lisboa: Plátano Edições Técnicas.

Oyola, A. V. & Alvarenga, Lídia. (2009) Mapeamento e representação do conhecimento e modelagem ontológica: Uma experiência na área da cultura do sorgo. *Encontro Nacional de Pesquisa em Ciência da Informação*, João Pessoa, PB, Brasil, 10.

Sales, L. F. (2006) *Ontologia de domínio: um estudo das relações conceituais e suas aplicações*. 2006. (Dissertação de Mestrado em Ciência da Informação). Universidade Federal Fluminense, Niterói.

Sales, L.F. & Sayão, L.F. (2012) Modelo triádico de relações: um protótipo de modelagem conceitual para a área nuclear. *Encontro Nacional de Pesquisa em Ciência da Informação*, João Pessoa, PB, Brasil, 13.

Silva, D.L, Souza, R. R., Almeida, M. B. (2008). *Ontologias e vocabulários controlados: comparação de metodologias para construção*. *Ciência da Informação*, Brasília, 37 (3), pp. 60-75.

Stockinger, G. (2001). *Para uma teoria sociológica da comunicação*. Salvador: UFBA. Retrieved from <http://www.bocc.ubi.pt/pag/stockinger-gottfriedteoria-sociologica-comunicacao.pdf>



PERSPECTIVAS DE INVESTIGAÇÃO EM REPRESENTAÇÃO E ORGANIZAÇÃO DO CONHECIMENTO: ATUALIDADE E TENDÊNCIAS

ONTOLOGÍA DE FENÓMENOS NATURALES: PLANEACIÓN Y DISEÑO

Adriana Suárez Sánchez¹, Ariel Alejandro Rodríguez García²

¹Posgrado en Bibliotecología y Estudios de la Información, Universidad Nacional Autónoma de México, 0000-0003-0215-337X, nizadriss@comunidad.unam.mx

²Instituto de Investigaciones Bibliotecológicas y de la Información, Universidad Nacional Autónoma de México, 0000-0001-5779-037-1, rgarciaa@unam.mx

RESUMEN se expone un proyecto que consiste en la planeación y el diseño de una ontología de "Fenómenos Naturales". Los objetivos establecidos para la ontología fueron: (1) representar el dominio y (2) convertirse en un instrumento indizador de un conjunto de recursos sobre el tema. La metodología seguida para la elaboración de la ontología se formó de las siguientes etapas: definición y delimitación del dominio, extracción y normalización conceptual, estructuración categorial, definición de atributos, establecimiento de relaciones y asignación de axiomas. Se propone que las ontologías pueden tener diversas aplicaciones en la organización temática llevada a cabo por la bibliotecología. Se determinó que, en tareas de representación, permiten la explicitación del dominio a través de la definición de sus conceptos, atributos y relaciones, así como la generación de mapas tópicos que permiten encontrar, identificar, seleccionar y explorar temas de interés. Mientras que, en la parte organizadora, los conceptos establecidos pueden ser aplicados en el arreglo temático de entidades de información. Se concluye que las ontologías son sistemas para la organización del conocimiento con alto potencial en la representación y organización temática de dominios de información vinculados a contextos web.

PALABRAS CLAVE *Ontologías, Sistemas para la Organización de Conocimiento, Organización del Conocimiento, Representación del Conocimiento, Fenómenos Naturales.*

ABSTRACT the project is related to the planning and design of an ontology about "Natural Phenomena". The objectives established for the ontology were: (1) represent the domain and (2) become an instrument for indexing a set of resources on the topic. The methodology used for the ontology construction consisted in the following stages: definition and delimitation of the domain, extraction and conceptual normalization, categorical structuring, definition of attributes, establishment of relations and assignment of axioms. The premise is that ontologies ought to have multiple applications in subject organization used by librarianship. It was determined that in representation tasks they allow the domain to be explained through the definition of its concepts, attributes and relationships, as well as the generation of topical maps that allow finding, identifying, selecting and exploring interest topics. Also, in the organizing part, they established concepts that can be applied in thematic arrangement of information entities. The conclusion is that ontologies are systems for the organization of knowledge with high potential for the representation and subject organization of information domains linked to web contexts.

KEYWORDS *Ontologies, Knowledge Organization Systems. Knowledge Organization, Knowledge Representation, Natural Phenomena.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUCCIÓN

Los fenómenos naturales son cambios que se producen en la naturaleza. Su tipología incorpora acontecimientos como arcoíris, terremotos, aurora boreal, nubes estratosféricas, ciclones, huracanes, aluviones, arcos circuncenitales, brumas, burbujas inflamables, etc. En nuestro planeta la ocurrencia de los fenómenos naturales es constante, día a día nuestro planeta registra diversos acontecimientos importantes de identificar, registrar y estudiar debido a las implicaciones que pueden tener en la vida humana (Debroise & Seinandre, 2003)

En México existen varios centros de investigación sobre fenómenos naturales entre los que es posible mencionar el Servicio Sismológico Nacional (SSM), el Centro Nacional para la Prevención de Desastres (CENAPRED) y la División de Ciencias de la Tierra (DCT) de la Universidad Nacional Autónoma de México (UNAM) cuyas tareas son monitorear, documentar e investigar los fenómenos naturales que se presentan en el país y en otras regiones del planeta. Tales instituciones además del monitoreo de los eventos producen un conjunto de recursos de información sobre el tema que deben ser almacenados y organizados considerando las necesidades de la comunidad especialista.

Para la descripción física de tales recursos se cuenta con esquemas de metadatos que satisfacen las necesidades ordenadoras pero, en lo relativo a la descripción temática se observan deficiencias en los sistemas organizadores del conocimiento (SOC en español, KOS en inglés). En México no se cuenta con un tesoro sobre fenómenos naturales, la herramienta más cercana al tema es el *Tesoro SPINES: un vocabulario controlado y estructurado para el tratamiento de información sobre ciencia y tecnología para el desarrollo* (UNESCO, 1984) que apenas registra el descriptor “Fenómenos naturales” y ofrece algunos términos relacionados. De igual modo, los encabezamientos de materia más usados en el país -*Library of Congress Subject Headings* (Library of Congress, 2007) y *Lista de Encabezamientos de Materias para Bibliotecas* (Biblioteca Luis Ángel Arango, 1998)- no cuentan con la representación temática adecuada, así, por ejemplo, no incluyen el descriptor “Fenómenos naturales”, registrando sólo algunos acontecimientos esenciales como “Tormentas” “Lluvia” y “Huracanes”.

Otro problema que se detecta en los sistemas para la organización del conocimiento es que establecen una división contundente entre los fenómenos que ocasionan desastres naturales (ciclones, terremotos, huracanes) y aquellos que no los ocasionan (aurora polar, arcoíris lunar, pilares de luz) (Debroise & Seinandre, 2003). Aunque tal división es lógica, motiva que los fenómenos naturales asociados a desastres naturales estén más definidos en comparación con los fenómenos naturales inofensivos que, prácticamente, son desestimados en las herramientas indizadoras, sin considerar que hay un conjunto de literatura y especialistas sobre el tema.

La investigación de los fenómenos naturales en México es una disciplina emergente, productora de un conjunto de recursos de información que deben ser organizados de manera adecuada. La descripción física de tales recursos es un asunto resuelto pues se cuenta con diversos esquemas de metadatos; sin embargo, la descripción temática se encuentra bastante limitada debido a los pocos descriptores que incluyen los principales sistemas organizadores del conocimiento. Considerando lo anterior, el presente proyecto se planteó la construcción de un sistema organizador del conocimiento, propio del español mexicano, que tratara adecuadamente el dominio “Fenómenos naturales”. Luego, como producto de tal planteamiento, surgió, una interrogante fundamental: ¿qué tipo de sistema?

Históricamente, la bibliotecología ha desarrollado diversos sistemas para la organización del conocimiento entre los que es posible mencionar listados de temas, encabezamientos de materia, taxonomías y tesauros. Tales herramientas se vislumbraron como una posibilidad; sin embargo, las necesidades organizadoras del dominio exigían que tal instrumento explicitara adecuadamente las relaciones entre los conceptos y pudiera agregarse al espacio web mediante entradas temáticas que vincularan a entidades de información digitales. Considerando las necesidades representadoras y organizadoras señaladas, se determinó la creación de una ontología de dominio.

Las ontologías son “una especificación de una conceptualización” (Gruber, 1995) fundamentadas en técnicas representadoras con las cuales es posible modelar un dominio de conocimiento (Gruber T. , 2008). Por técnicas se entienden actividades dirigidas al establecimiento de elementos modeladores: clases (o categorías), atributos (o propiedades), relaciones y reglas o axiomas que forman la estructura de un dominio. Jiménez (2008) explicita su función modeladora de conocimiento cuando menciona que son:

Una representación formal de un determinado dominio o área de conocimiento. Esta representación consiste en la extracción de los conceptos más importantes del dominio y de las relaciones que se dan entre ellos, creándose una estructura de conceptos relacionados que proporciona un vocabulario común para el conocimiento modelado por la ontología y constituye una importante forma de compartir información (Jiménez Mavillard, 2008, pág. 1).

En la bibliotecología, en general, y en el ordenamiento temático de recursos de información en espacios digitales, en particular, las ontologías resultan esenciales debido a sus posibilidades representadoras y organizadoras. Al respecto, Marcia Zeng (2008) señala que, en comparación con otros sistemas más tradicionales de la bibliotecología, éstas, además de presentar los términos de un área temática, los definen y asocian mediante el establecimiento de atributos, relaciones y axiomas que permiten instaurar vinculaciones y extensiones de vocabulario. De modo que:

Una ontología abarca la estructura clasificatoria usada por las taxonomías y los tesauros. Pero su principal característica es la presentación de propiedades para cada clase dentro de la estructura clasificatoria. Con una taxonomía completa y propiedades exhaustivas, una ontología funciona de dos modos: como un vocabulario conceptual y como una plantilla de trabajo que permite el almacenamiento, la búsqueda y el razonamiento, basado en instancias y reglas (Zeng, 2008, pág. 176).

Bajo la premisa de que las ontologías incluyen y relacionan conceptos de un dominio a la vez que son una plantilla al momento de recuperar información por temas. El presente proyecto consideró su la construcción de una Ontología de Fenómenos Naturales (OFN). Dicha herramienta, tuvo tres factores generadores:

- La necesidad de una comunidad de expertos en ciencias de la tierra que necesitan contar con un sistema para la organización del conocimiento adecuado para el dominio “Fenómenos naturales”.
- La necesidad de organizar un conjunto de recursos digitales producidos por un gremio altamente especializado e interesado en la investigación sobre el tema.
- El interés de explorar un sistema para la organización del conocimiento reciente que posee alta definición conceptual, esquematización del dominio en una estructura jerárquica precisa, atributos y relaciones.

Los objetivos establecidos para la OFN fueron:

- Representar el dominio “Fenómenos naturales” bajo una adecuada estructuración del área sustentada por conocimiento experto.
- Servir como herramienta indizadora de un conjunto de recursos de información, considerando la herramienta un sistema para la organización del conocimiento con terminología altamente especializada y con garantía de uso entre los especialistas del área.

METODOLOGÍA

En la planeación y diseño de la OFN se consultaron algunas metodologías para la construcción de ontologías, entre las que es posible mencionar Noy & McGuinness (2007), King & Reinold (2008) y Almeida & P., (2003). Aunque se consideraron ciertos postulados de tales metodologías, se optó por un proceso propio sustentado en dos momentos:

- 1) Establecimiento de los elementos formantes de la ontología
- 2) Etapas para su construcción

Circunscritas en la organización de la información, las ontologías poseen diversas características (formales, semi formales, informales, terminológicas, de procesos, de tareas, etc.). Frente a estas posibilidades la OFN es una ontología de dominio con cuatro características: terminológica, en tanto que incluye los conceptos de un dominio; modeladora de conocimiento, dado que busca representar adecuadamente el campo tratado; formal, puesto que incluye conceptos y nociones relacionales de una comunidad especializada, y aplicada, porque servirá como sistema para la organización temática de recursos de información.

A partir de las características antes señaladas, sus elementos formantes, fueron definidos como sigue (Fig.1):

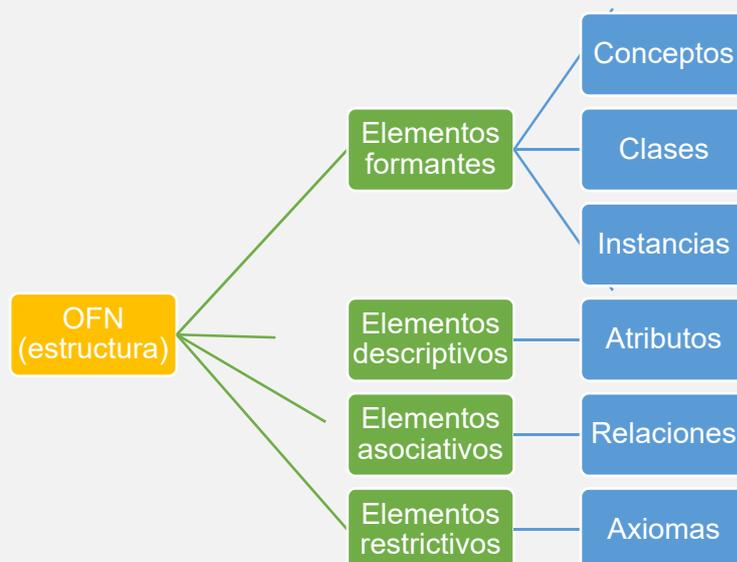


Fig.1. Elementos formantes de la OFN

Fuente: elaboración propia

- **Conceptos:** son unidades de pensamiento abstraídos de una multitud de objetos mediante el análisis de las propiedades en común de esos objetos (DIN: 2342: Vocabulary of Terminology, 2011). Refiere a una unidad de conocimiento que sintetiza en una forma lingüística los predicados necesarios y verificables sobre un objeto escogido (Dahlberg, 1985). Constituyen la base de las ontologías debido a su naturaleza representadora de entidades.
- **Clases:** consisten en el orden que le es asignado a las personas, los objetos o los conceptos a partir de sus características, atributos, cualidades, rasgos, etc. (Reitz, 2004). Se fundamenta en la premisa de que en la naturaleza existen cosas con características comunes que posibilitan agruparlas. Vista así, la clasificación opera por comparación entre las características de las cosas, estableciendo un principio de orden que obedece a un conjunto complejo y dinámico de inferencias (Kobashi & Francelin, 2011).
- **Instancias:** son cada uno de los individuos que ocupan las clases más bajas en las jerarquías de la ontología. En sentido estricto, son los objetos específicos de una clase (García Jiménez, (2004); Lucas da Silva, (2008)). Contienen elementos o datos que describen o ejemplifican una entidad de modo inequívoco mediante propiedades que no se recogen en las clases o subclases que le anteceden.
- **Atributos:** constituyen características de una entidad o variables que las definen, posibles de establecer a partir de la observación y estudio de éstas ya sea mediante procesos enfocados o temporales. Son funciones que asignan a cada miembro de una serie de individuos un valor específico (Delgado Gómez, 2002) y ayudan a definir las características de las clases (Lamarca Lapuente, 2008). FRBR (IFLA, 2004) indica que la descripción de los atributos posibilita que “las tareas del usuario (encontrar, identificar, seleccionar, obtener, relacionar y navegar) sean cumplidas a partir de una búsqueda ya sea en un catálogo o sistema recuperador de la información” (Tillet, 2003, pág. 5).
- **Relaciones:** son una asociación a través de varias cosas, de modo que esa asociación tenga una significación particular. Están presentes en la mayoría de sistemas para la organización de la información. No obstante, en los instrumentos insertos en contextos web -como es el caso de las ontologías- se hacen más necesarios y útiles porque forman una red de aspectos legibles para los agentes artificiales que, en términos prácticos, posibilitan mayores posibilidades de recuperación ante solicitudes de búsqueda planteadas al sistema recuperador.
- **Axiomas:** son proposiciones lógicas y matemáticas donde cumplen la función de una premisa, considerada evidente, que sirve para la conclusión de un segundo elemento. En el ámbito de las ontologías, son verdades que se declaran sobre relaciones que deben cumplir los elementos de una ontología, es decir, un “predicado verdadero/falso en forma lógica. Todos los axiomas que integran la ontología proveen una abstracción total del universo que describe.” (Ríos Alvarado, 2013, pág. 13)

Una vez establecidos los elementos formantes, se determinó el proceso para la construcción de la OFN, constituido de seis etapas:

- Definición y delimitación del dominio
- Extracción y normalización conceptual

- Categorización
- Determinación de atributos o propiedades
- Establecimiento de relaciones
- Designación de axiomas

1. DEFINICIÓN Y DELIMITACIÓN DEL DOMINIO

Un dominio es un espacio de información/conocimiento con características específicas. Al respecto, Smiraglia establece que:

[...] un dominio es cualquier grupo que es útil para la construcción de un sistema para la organización del conocimiento. Esto nos lleva naturalmente a las condiciones que el grupo muestra: conocimiento (una base ontológica), objetivos (una teleología subyacente), los métodos de investigación (hipótesis y metodologías) y un sistema funcional para la comunicación (la semántica social). Este desembalaje nos da mucha libertad de acción. Un dominio, entonces, puede ser un grupo de personas que trabajan juntos si comparten conocimientos, objetivos, métodos de operación y la comunicación. Puede ser una comunidad de aficionados, una disciplina académica, un académico, un departamento y así (Smiraglia, 2015, pág. 86).

En la construcción de ontologías la definición y delimitación del dominio es fundamental porque entre las disciplinas existen secciones difusas. Considerando que el dominio de la OFN es “Fenómenos naturales”, en su definición y delimitación se revisaron fuentes de consulta y sistemas para la organización de conocimiento con miras a establecer sus alcances y límites frente a otros dominios como “Desastres naturales”, “Eventos naturales”, “Ciencias de la tierra”, etc.

Tras el trabajo de investigación, el dominio quedó definido del modo siguiente (Fig. 2):

Fenómenos naturales: son procesos permanentes de movimientos y transformaciones que sufre la naturaleza. Frecuentemente un fenómeno natural aparece casi como sinónimo de acontecimiento inusual o sorprendente. Un ejemplo de fenómeno natural es el arcoíris, de igual modo que los ciclones. De modo erróneo se ha hecho una equiparación entre fenómenos naturales y desastres naturales, sin que exista correspondencia. El abanico de los fenómenos naturales incorpora tanto acontecimientos que no ocasionan daños en la vida humana (aurora polar, olas) como aquellos que pueden generar desastres naturales (ciclones, tormentas eléctricas). (Debroise & Seinandre, 2003)

Fig. 2. Definición del dominio “Fenómenos naturales”

Una vez que el dominio ha sido definido, la delimitación implica trazar límites al área de acción. Así, los fenómenos naturales están circunscritos por un término más amplio (BT) que es “Ciencias de la Tierra” y tienen como términos reducidos (NT) “Terremotos”, “Nubes”, “Ciclones”, etc.

2. EXTRACCIÓN Y NORMALIZACIÓN CONCEPTUAL

Los dominios constituyen su universo de existencias mediante conceptos que representan, diferencian y definen las entidades de un campo de conocimiento, posibilitando, por un lado, su existencia desde el

punto de vista de la organización del conocimiento y, por otro, la comunicación intrínseca y extrínseca de la disciplina (Moreiro González J. , 2006).

Para la extracción de conceptos de un dominio se pueden aplicar diversas técnicas emanadas de teorías lingüísticas, computacionales y documentales. Los procesos de extracción conceptual aplicados en la construcción de la OFN fueron: reuso de vocabularios previos (*LCSH*, *LEMB*, *Glosario del Servicio Meteorológico Nacional*), extracción conceptual manual de recursos sobre el dominio y entrevistas con expertos (Fig. 3).

Una vez extraídos los conceptos es importante normalizarlos con miras su presentación. La estandarización permite unificar el uso de singulares, plurales, siglas, mayúsculas, minúsculas, etc. En la normalización conceptual de la OFN se consideraron instrumentos como *ISO 2788:1986. Guidelines for the establishment and Development of Monolingual thesauri* (ISO, 1986) e *ISO 25964-1:2011. Thesauri for Information Retrieval* (ISO, 2011) que orientan en la construcción de vocabularios controlados.



Fig.3. Conceptos en la OFN
Fuente: elaboración propia

3. CATEGORIZACIÓN

Las categorías son agrupamientos de entidades con características comunes o bien:

[...] metaconceptos que nombran agrupamientos de términos de una determinada área. En este sentido, se llama categoría a los agrupamientos que resultan de la división del universo de conocimiento, de acuerdo con las características intrínsecas o fundamentales de cada concepto. (Cavalvanti, 1978, pág. 15)

Tanto en la realidad como en la organización del conocimiento las categorías cumplen una función muy importante en tanto que:

[...] permiten relacionar unas cosas con otras en términos de similitud o diferencias y están involucradas en la manera como percibimos, analizamos, predecimos o clasificamos. Sin las categorías, percibiríamos el mundo como un desorden difuminado de cosas sin relaciones comprensibles o sobresalientes. (Glushko, 2013, pág. 237)

En la construcción de las categorías se aplicaron principios jerárquicos, partiendo de la clase más alta “Fenómenos naturales”, sus posteriores subdivisiones, por ejemplo, “Fenómenos naturales hidrológicos”, “Fenómenos naturales meteorológicos”, etc., hasta llegar a las clases más bajas, por ejemplo, “Tromba marina”, “Arcoíris de fuego”, etc., (fig. 4).

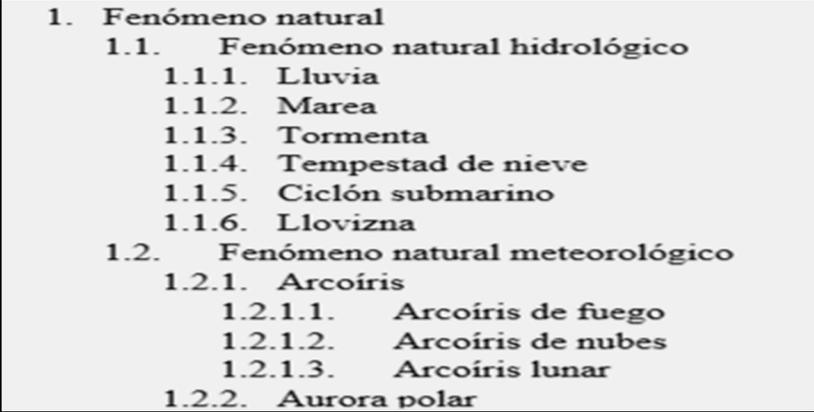
- 
1. Fenómeno natural
- 1.1. Fenómeno natural hidrológico
 - 1.1.1. Lluvia
 - 1.1.2. Marea
 - 1.1.3. Tormenta
 - 1.1.4. Tempestad de nieve
 - 1.1.5. Ciclón submarino
 - 1.1.6. Llovizna
 - 1.2. Fenómeno natural meteorológico
 - 1.2.1. Arcoíris
 - 1.2.1.1. Arcoíris de fuego
 - 1.2.1.2. Arcoíris de nubes
 - 1.2.1.3. Arcoíris lunar
 - 1.2.2. Aurora polar

Fig. 4. Categorización en la OFN

Fuente: elaboración propia

4. DETERMINACIÓN DE ATRIBUTOS

La noción de propiedades o atributos, entendidos éstos como elementos descriptivos, ha estado presente en la representación que la bibliotecología ha hecho de los recursos. Sin embargo, ha cobrado mayor importancia a partir de las declaraciones de (IFLA, 2004) bajo el modelo conceptual del universo bibliográfico, que consiste en diversas entidades que se relacionan entre sí y pueden ser descritas mediante elementos de información (Tillet, 2003).

En las ontologías temáticas los atributos están estrechamente relacionados con los conceptos que esclarecen el término, por lo que en la OFN sólo fueron asignados en los casos donde pudiera existir traslape semántico (Fig. 5).



Fig.5. Determinación de atributos en la OFN

Fuente: elaboración propia

5. ESTABLECIMIENTO DE RELACIONES

Las relaciones en las ontologías responde a tres funciones: (1) explicitar la naturaleza y comportamiento de los conceptos para reconstruir el conocimiento, (2) presentar modelos organizativos de la información que sean significativos y (3) capturar procesos o dominios de información. Las ontologías construyen un entramado relacional que ofrece a los usuarios un instrumento de acceso por temas más flexible a partir de la navegación desde la visión conceptual, jerárquica y asociativa; explorar las relaciones entre temas/ o sus denominaciones, por ejemplo, explorar las relaciones con el objetivo de entender la estructura de un dominio de información y su terminología y establecer una gama de atributos y relaciones que posibilita definir con precisión temas de interés.

En la OFN se retomaron las relaciones establecidas por el modelo *Requisitos funcionales para datos de autoridad de materia* (FRSAD) (IFLA, 2013), caracterizado por ser un modelo entidad-relación para la organización temática de recursos de información pero de todas las relaciones mencionadas por el modelo sólo se detectaron cinco tipos: equivalencia, dependencia jerárquica, todo/parte, causa/efecto y espacial. De estas, las dominantes en el dominio fueron las de equivalencia (Fig.6).

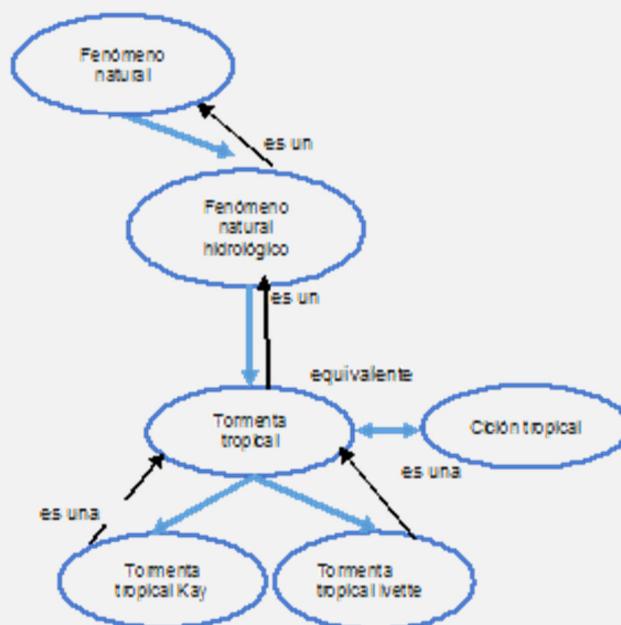


Fig. 6. Relaciones de equivalencia en la OFN

Fuente: elaboración propia

6. DESIGNACIÓN DE AXIOMAS

La construcción de la ontología se realizó con el programa de fuente abierta *Protégé 4.3* y se consideraron sólo dos axiomas que tal programa posibilita: “Join” y “Disjoint”, así cada clase hereda o limita propiedades a las clases inferiores que se le vinculan.

ANÁLISIS DE RESULTADOS

Los aspectos esenciales de la OFN fueron los siguientes:

- Incluye 167 conceptos asociados a “Fenómenos naturales”
- Incorpora 6 niveles categóricos o jerárquicos
- Establece atributos para conceptos con semántica difusa
- Presenta 5 tipos de relaciones entre conceptos: equivalencia, dependencia jerárquica, todo/parte, causa/efecto y espacial.

Acorde a los objetivos establecidos, como herramienta representadora la OFN refleja adecuadamente la estructura del dominio ya que posee un número considerable de términos que ofrecen mayor profundidad conceptual que la registrada en LCSH y LEMB. También en lo tocante a la representación, el establecimiento de relaciones entre los conceptos del dominio permite visualizar un entramado asociativo de gran utilidad tanto para los usuarios expertos como para aquellos no expertos, pero interesados en el tema. Como herramienta organizadora, la ontología aporta descriptores precisos para la indización de un conjunto de recursos que anteriormente eran colocados en categorías generales.

Llevadas al contexto digital, las ontologías, en comparación con los encabezamientos de materia y los tesauros, posibilitan observar de modo jerárquico los conceptos que incluyen. En consecuencia, si la ontología se pone a disposición de los usuarios, brindará a los usuarios la posibilidad de explorar la estructura del dominio mediante navegación, es decir, el usuario puede partir de un término amplio del dominio, por ejemplo, “Fenómenos naturales” y éste se desplegará en categorías como “Fenómeno natural hidrológico” o “Fenómeno natural meteorológico”, a su vez, tales clases se pueden subdividir en categorías como “Marea”, “Ciclón submarino”, “Centellas”, etc.

Tal aspecto de las ontologías, las convierte en herramientas a través de las cuales los usuarios están en posibilidad de delimitar sus intereses de búsqueda temática a la vez que generan conocimiento respecto al dominio (Fig. 7).

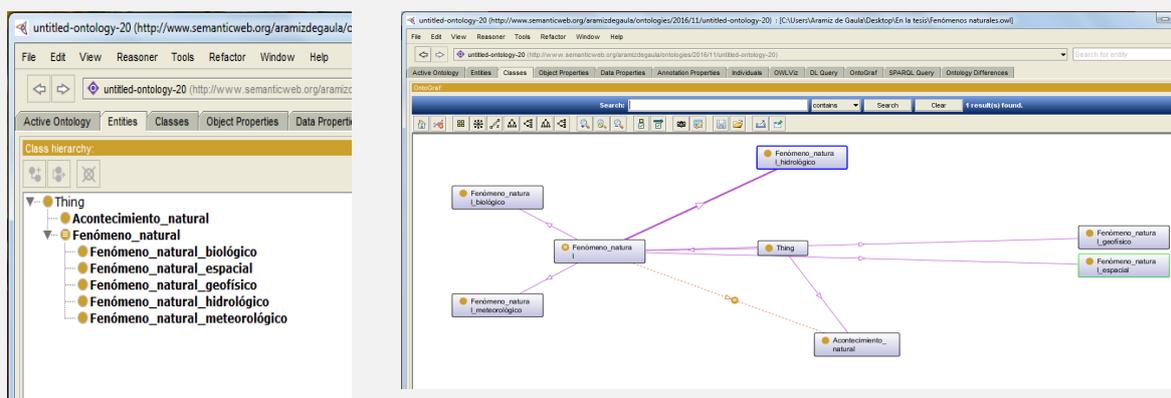


Fig. 7. OFN en Protégé 4.3

Fuente: elaboración propia

CONCLUSIONES

Organizar información desde la perspectiva temática ha sido una de las actividades fundamentales de la bibliotecología y, para ello, se han construido diversos sistemas para la organización del conocimiento. Durante décadas, las clasificaciones y los vocabularios controlados han sido las herramientas más empleadas para organizar temáticamente dominios y recursos de información. Sin embargo, en los últimos años el panorama de los sistemas para la organización del conocimiento se ha enriquecido con nuevos sistemas, emergidos de las tecnologías de la información, entre los que es posible mencionar los tesauros y taxonomías digitales y las ontologías.

Las ontologías son un sistema para la organización de dominios de información especializados, por sus características posibilitan una cobertura temática adecuada del área, así como el establecimiento de propiedades y relaciones entre los conceptos existentes. A partir del proceso de planeación y construcción de la OFN, la presente investigación concluye que:

- Las ontologías tienen alto potencial representador y organizador en diversos ámbitos.
- Existen numerosas metodologías para ontologías, pero si éstas se quieren insertar en funciones bibliotecológicas, deben analizarse detalladamente ya que en su mayoría están dirigidas a ámbitos computacionales.
- Las ontologías son una opción viable para representar y organizar dominios de información de comunidades altamente especializadas que requieren ordenamientos

precisos, más allá de las posibilidades que brindan los encabezamientos de materias y los tesauros.

- Sus ventajas representadoras y organizadoras están especialmente asociadas a contexto de organización de información en el espacio web.
- El proceso de creación de la OFN fue una tarea multidisciplinaria en la que fueron necesarios conocimientos: lingüísticos, computacionales, bibliotecológicos y expertos de dominio.
- La OFN es una herramienta satisfactoria que entrará en una etapa de implementación mediante su vinculación con un sitio web que albergue recursos de información bajo las entradas temáticas establecidas.

REFERENCIAS

Almeida, M. B., & P., B. M. (2003). Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. *Ciencias da Informação*, 32(3), 7-20.

Biblioteca Luis Ángel Arango. (1998). *Lista de encabezamientos de materia para bibliotecas*. Bogotá: Eberhard.

Cavalvanti, C. (1978). *Indexação y Tesouro: metodologia y técnicas*. Brasília : Associação de Bibliotecários do Distrito Federal.

Dahlberg, I. (1985). Begriffsbeziehungen und Definitionstheorie. *Terminologie und Benachbarte Gebiete*, 137-148.

Debroise, A., & Seinandre, E. (2003). *Fenómenos naturales: un planeta activo*. Barcelona : Vox.

Delgado Gómez, A. (2002). Organización de la información mediante el uso de lenguajes de modelado. *Scire: representación y organización del conocimiento*, 8(1), 55-86.

DIN: 2342: Vocabulary of Terminology. (2011). Berlín: Swedish Standards Institute.

García Jiménez, A. (2004). Instrumentos de representación del conocimiento: tesauros versus ontologías. *Anales de documentación*(7), 79-95.

Glushko, R. J. (Ed.). (2013). *The Discipline of Organizing*. Cambridge: MIT Press .

Gruber, T. (1995). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal of Human and Computer Studies*(43), 907-928.

_____. (2008). *Ontology*. Recuperado el 17 de Febrero de 2016, de <http://tomgruber.org/writing/ontology-definition-2007.htm>.

IFLA. (2004). *Requisitos funcionales de los registros bibliográficos*. Madrid: Ministerio de Cultura.

_____. (2013). *Requisitos funcionales para datos de autoridad de materia*. Recuperado el 1 de Julio de 2017 de: <https://www.ifla.org/files/assets/cataloguing/frsad/frsad-final-report-es.pdf>.

ISO. (1986). *ISO 2788:1986.Guidelines for the establishment and Development of Monolingual thesauri*. Ginebra, Suiza: ISO.

_____. (2011). *ISO 25964-1:2011. Thesauri for Information Retrieval*. Ginebra, Suiza: ISO.

Jiménez Mavillard, A. (2008). *Razonamiento con ontologías*. Recuperado el 8 de Marzo de 2016, de <http://www.cs.us.es/cursos/ia2-2008/trabajos/razonamiento-con-ontologias.pdf>.

King, B. E., & Reinold, K. (2008). *Findign the Concept, not Just the Word: A Librarians Guide to Ontologies and Semantics*. Oxford: Chandos Publishing.

Kobashi, N. Y., & Francelin, M. M. (2011). Conceitos, Categorias e Organizaçao do Conhecimento. *Information & Information, 16*(3), 1-24.

Lamarca Lapuente, M. J. (2008). *Ontologías*. Recuperado el 30 de Mayo de 2008, de Hipertexto: el nuevo concepto de documento en la cultura de la imagen . Recuperado el 1 de Julio de 2017 de: <http://www.hipertexto.info/documentos/ontologias.html>.

Library of Congress. (2007). *Library of Congress Subject Headings* (30 ed.). Washington: Library of Congress.

Lucas da Silva, D., Rocha Souza, R., & Barcellos Almeida, M. (2008). Ontologias e vocabulários controlados: comparação de metodologías para su construção. *Ciencia da informação, 37*(3), 60-75.

Moreiro González, J. (2006). Evolución paralela de los lenguajes documentales y la terminología. En B. Rodriguez Bravo, & M. Alvite Díez (Ed.), *La interdisciplinariedad y la transdisciplinariedad en la organización del conocimiento* (págs. 27-42). León, España: Universidad de León.

Noy, N. F., & McGuinness, D. L. (2007). *Ontology Development 101: A Guide to Creating Your First Ontology*. Recuperado el 31 de Marzo de 2017 de: http://protege.stanford.edu/publications/ontology_development/ontology101.pdf

Reitz, J. M. (2004). *Dictionary for Library and Information Science*. Westport, C.T.: Libraries Unlimited.

Ríos Alvarado, A. B. (2013). *Obtención de axiomas en el aprendizaje de ontologías*. Tamaulipas, México: Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (Tesis).

Smiraglia, R. (2015). *Domain Analysis for Knowledge Organization*. Nueva York: Chandos .

Tillet, B. (2003). *El modelo FRBR: requisitos funcionales para registros bibliograficos*. Recuperado el 18 de Noviembre de 2015 de: <http://www.loc.gov/catdir/cpso/frbrspan.pdf>.

UNESCO. (1984). *Tesaurus SPINES: un vocabulario controlado y estructurado para el tratamiento de la información sobre ciencia y tecnología para el desarrollo*. Madrid: Instituto de Información y Documentación en Ciencia y Tecnología : Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura.

Zeng, M. L. (2008). Knowledge Organization Systems (KOS). *Knowledge Organization* , 35 (2-3), 160-182.



TAXONOMIA COMO INSTRUMENTO AUXILIAR PARA ESCOLHA DE PROCEDIMENTOS DE AUDITORIA CONTABIL

José Humberto da Cruz Cunha

Universidade de Brasília (UnB), 0000-0003-0210-9854, jhccruzeiro@yahoo.com.br

RESUMO Estuda o uso da taxonomia como instrumento auxiliar para a escolha de procedimentos no processo de auditoria. Há inicialmente a construção da taxonomia com o objetivo de classificar distorções contábeis numa estrutura documentária representativa do processo de auditoria, utilizando a classificação facetada na intenção de formar trilhas objetivas nessa estrutura classificatória. Esse processo decisório inicia-se com a análise de risco de distorções contábeis e termina na escolha de procedimentos pelo auditor. As relações lógicas entre termos de diferentes categorias representam o processo de auditoria e proporcionam fundamento teórico para decisões do auditor quanto à escolha de seus procedimentos, tendo como base os aspectos dos riscos de distorções das demonstrações contábeis. No sentido prático, houve a análise de conteúdo nos papéis de trabalho de uma auditoria executada pelo Tribunal de Contas da União, órgão de auditoria das contas públicas do Governo Federal brasileiro. É um estudo exploratório que tem como procedimento técnico o levantamento bibliográfico e a pesquisa documental, os quais subsidiam as análises qualitativas. Os resultados obtidos indicam essa taxonomia como instrumento auxiliar às decisões dos auditores para a escolha de procedimentos. Há ainda indício de redução de subjetividade à medida que as escolhas do auditor passam por processos fundamentados em relações lógicas da Teoria do Conceito, considerando que a sequência lógica do processo de auditoria está representada na estrutura de classificação documentária, onde tais etapas evidenciam sequências objetivas no processo de análise de risco de distorções até a escolha de procedimentos.

PALAVRAS-CHAVE *Taxonomia, auditoria, classificação facetada, procedimentos de auditoria, redução de risco.*

ABSTRACT This is a study about taxonomy as an auxiliary device used by decision making for an accounting audit process. First, there is a taxonomy developed from financial distortion risks, classified based on a documentary structure that represents the audit process steps. A faceted classification was used considering its flexibility, and interested to make logical relations on audit trail. That decision making process starts by financial risk analysis and ends on audit procedures choice. The logical relationships between terms on different categories represent the audit process and provide a theory support for the auditor. After a theoretical taxonomy development, there is a documentary research combined with a content analysis on audit paper works from The Federation Audit Board of Brazil (Tribunal de Contas da União/TCU) to apply that model with a real case. That is an exploratory study using a bibliographical and documental research to support the qualitative approach. As results, it indicates the taxonomy as an auxiliary device to auditor's decision making to pick procedures based on risk analysis on a objective way. It indicates a subjective aspects reduction by that process based on logical relationships under Conceptual Theory. Also, it concerns that the logical steps on audit trail is represented on documentary classification structure, and it shows an objective steps on process for choosing procedures stand on risk analysis.

KEY WORDS *Taxonomy, audit, faceted classification, audit procedures, risk reduction.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

O processo de auditoria contábil possui subjetividade nas decisões do auditor, o que é normalmente conhecido como risco de detecção, e diretamente influencia o risco de auditoria (Boynton & Johnson & Kell, 2002; Longo, 2011). Por outro lado, tem-se a taxonomia como estrutura para organização e recuperação de informações em linguagem documentária, com potencial de objetividade (Campos & Gomes, 2008; Baptista & Araujo Junior & Carlan, 2010; Sousa & Araujo Junior, 2013). Diante dessa realidade e interessado em reduzir a subjetividade nas decisões do auditor, busca-se estudar a utilização da taxonomia para essa finalidade. Assim, tem-se a seguinte pergunta de pesquisa: *Como a taxonomia pode auxiliar o auditor na escolha de procedimentos?*

Na literatura contábil, as metodologias de auditoria não possuem um padrão e cada empresa tem sua forma de executar, desde que obedeçam a legislação aplicada (Yoshitake, 2009; Boynton & Johnson & Kell, 2002). Essa flexibilidade e a diferença entre níveis de conhecimento dos auditores deixam mais evidente o risco de detecção na auditoria contábil. Em geral, os procedimentos são escolhidos após a análise de riscos de distorções nas demonstrações financeiras, provenientes de erros ou fraudes. Especificamente, a possibilidade de escolha, aplicação e interpretação inadequada dos procedimentos, dependente de aspectos subjetivos intrínsecos ao auditor (Almeida, 2010; Boynton & Johnson & Kell, 2002; Yoshitake, 2009; Whittington, Delaney, 2011; Beasley et al, 2015).

Todavia, a taxonomia é um sistema de classificação que apoia o acesso à informação permitindo classificar, alocar, recuperar e comunicar informações em um sistema de maneira lógica, destinando-se a inúmeras finalidades, às quais todas objetivam classificar e organizar (Baptista & Araujo Junior & Carlan, 2010; Sousa & Araujo Junior, 2013). Sendo assim, o estudo da sua inserção no processo de auditoria tem a finalidade de organizar, classificando erros e fraudes em uma linguagem documentária, de maneira que auxilie o auditor a percorrer passos mais precisos (objetivos) na escolha dos procedimentos a serem aplicados (reduzindo espaços para subjetividade).

A decisão sobre qual procedimento utilizar para determinado risco é de natureza tática. Os impactos do planejamento das decisões tática causam um impacto menor que um trabalho gerado por atitudes não pensadas (Aziza & Kaplan & Fitts & Norton, 2013; Parnell, Bresnick & Tani & Johnson, 2013). Na visão dos autores, as decisões táticas direcionam diretamente as operacionais para atingir aos objetivos máximos da estratégia da organização. E, de acordo com Coleman (2012, p. 67) “gerenciamento de risco requer a execução de decisões táticas e estratégicas para controlar os riscos que deveriam ser controlados e explorar os riscos que deveriam ser explorados”. A auditoria é essencialmente executada a partir da identificação e gerenciamento de riscos que afetem as informações contábeis.

Utilizando a taxonomia proposta, as decisões táticas na escolha e aplicação de procedimentos de auditoria são formadas em um processo com lastro documental, uma trilha. Isso acontece a partir do uso de uma estrutura com regras e padrões definidos. No caso da taxonomia, esse caminho possui um padrão que permite o processamento da linguagem natural para então tornar-se padronizada. Essa estrutura mapeia trilhas conceituais na auditoria e focaliza na indicação de procedimentos com base nos riscos classificados, numa rota que colabora para redução da subjetividade natural do processo de

auditoria. “A classificação é uma função importante para a transparência e o compartilhamento de informações, que são caminhos seguros para a tomada de decisão” (Sousa, 2003, p.240).

Ao falar de distorções contábeis provenientes de erros ou fraudes, é entendido que são lições aprendidas pelos auditores e são relatadas em papéis de trabalho e na literatura da área. A expressão Lições Aprendidas “refere-se a aprendizagem de projetos, caminhos do conhecimento que tem (a) mérito (qualidade), (b) merecimento (valor), ou significado (importância)” (Thomas, 2015, p. 5). O valor do aprendizado está em evitar a repetição de experiências ruins e em repetir experiências de sucesso. No caso das Lições Aprendidas pela taxonomia proposta, trata-se da tentativa de evitar novamente erros e fraudes (Milton, 2010).

No contexto de aplicação da taxonomia na auditoria, assumimos aqui que a taxonomia é uma forma de representação do conhecimento do auditor aplicada à organização e recuperação da informação, no caso lições aprendidas. Segundo Bräscher e Carlan (2010, p. 159), “no domínio das representações do conhecimento, as taxonomias são instrumentos que organizam logicamente os conteúdos informacionais”. Assim, a taxonomia de lições aprendidas, uma ferramenta de linguagem documentária que armazena riscos de distorções contábeis possui um conjunto de informações que acessa o âmbito do conhecimento no momento em que o auditor toma decisões a partir do uso dessas informações.

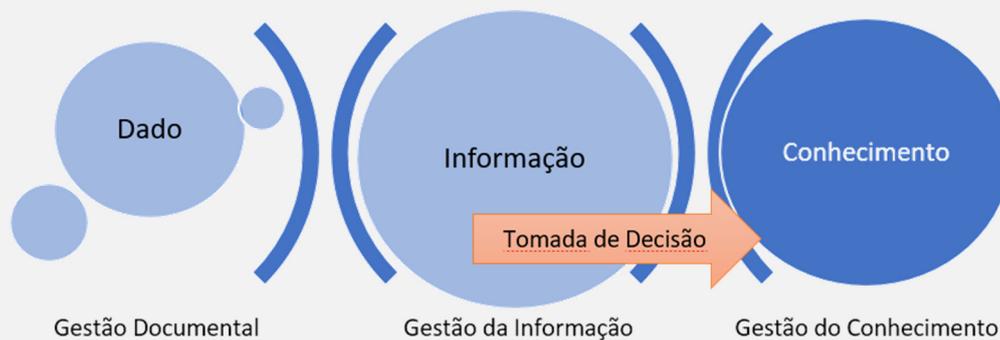


Figura 1: Acesso ao conhecimento através da tomada de decisão utilizando linguagem documentária.

Referência: elaboração própria.

Em relação aos estudos que tratam da taxonomia com aplicação na área contábil, a abordagem em gestão de risco na auditoria contábil assim como estrutura proposta são aspectos inovadores na Ciência da Informação. A taxonomia com a finalidade de auxiliar na identificação de problemas já acontece em outras áreas do conhecimento foram levantados. Na Aviação Civil e na Prevenção de Acidentes de Trânsito, por exemplo, foi observado que há o uso da taxonomia no sentido de auxiliar a auditoria quanto à validação e checagem de manutenção, auxiliando na gestão de riscos de acidentes (Hsiao, 2011). Entretanto essas são taxonomias hierárquicas ou de problemas ou de procedimentos. No caso dessa pesquisa tem-se a proposta de uma taxonomia única que classifica distorções (problemas) e na mesma estrutura permite identificar os procedimentos. E, diante dessa complexidade utiliza-se a classificação facetada.

A classificação facetada vai além da hierárquica permitindo uma relação analítico-sintética, apresentada por Ranganathan (1960) como caracterizada pela combinação de números em diferentes unidades de arranjos designando permutações e combinações, onde números de classes para toda possibilidade de assuntos podem ser construídos. Em seguida, Vickery (1966, p.9) apresenta o sistema de classificação facetada fundamentado no princípio da coordenação de conceito, onde há a combinação de dois ou mais

símbolos, onde cada qual possui seu grupo conceitual independente. Tais relações focam um determinado conteúdo onde as categorias e os termos relacionados possuem relações lógicas.

Além disso, a construção dos conceitos de cada faceta e dos seus respectivos termos são moldados pela Teoria do Conceito explanada por Dahlberg (1978), onde o conjunto de características de determinado termo formam seu conceito. E, no âmbito da faceta, o conjunto de características dos termos compõem seus conceitos em sentido lógico.

De maneira geral, segundo Baptista, Araujo Junior e Carlan (2010), as taxonomias adquirem importância no contexto das organizações, como instrumentos auxiliares à gestão do conhecimento e na organização e recuperação da informação. Elas garantem que todas as aplicações da instituição utilizem a mesma linguagem para organizar, armazenar e apresentar a informação, facilitando e permitindo o acesso e a recuperação da informação. De igual maneira, numa estrutura classificatória a relação entre termos de diferentes facetas é fundamentada com tais relações lógicas conceituais.

Sendo assim, o objetivo aqui é estudar o uso da taxonomia como instrumento auxiliar para a escolha de procedimentos no processo de auditoria.

Interessado em atingir o objetivo geral, fez-se necessária:

- 1) a análise conceitual para a construção de uma metodologia de uma taxonomia atrelada ao processo de auditoria, a partir da aplicação como modelo de representação de lições aprendidas, observando como variáveis: (a) aos aspectos conceituais para composição das categorias da taxonomia; (b) as relações lógicas entre atributos e conceitos; e, (c) as relações de equivalência, hierárquica e de associação entre os termos.
- 2) a partir de caso prático, a verificação o risco de detecção na auditoria mediante as trilhas conceituais aqui analisadas, observando como variáveis: (a) os Riscos de Distorções, Afirmarções e Procedimentos; e, (b) pontos divergentes e convergentes entre as escolhas dos auditores e a sugestão do modelo proposto.

Quanto à metodologia, a pesquisa é qualitativa e exploratória, a qual acomete a interpretação para os problemas objetos da investigação e visa proporcionar maior familiaridade com o problema a fim de torna-lo explícito (Matias-Pereira, 2007; Flick, 2009). A construção da taxonomia deu-se mediante a pesquisa bibliográfica tendo em vista o levantamento e a análise dos aspectos conceituais do processo de auditoria, observando as questões intrínsecas à análise de risco e escolha de procedimentos, assim como na Ciência da Informação a respeito da estrutura da taxonomia enquanto linguagem documentária. A coleta de dados se dá a partir da pesquisa documental em papéis de trabalho de uma auditoria do Tribunal de Contas da União (TCU), órgão responsável pela auditoria das contas do Governo Federal brasileiro. A análise de conteúdo é utilizada para análise técnica desses dados coletados, visando manter a objetividade e a consistência nos resultados da pesquisa (Elo & Kyngas, 2007; Rossi & Serralvo & Joao, 2014).

DESENVOLVIMENTO E APLICAÇÃO DA TAXONOMIA

O processo de auditoria contábil foi descrito a partir de uma pesquisa bibliográfica, onde ficou entendido que tal processo inicia-se com a (1) análise de riscos de distorções relevantes – possibilidades de erros ou fraudes e, partir de então (2) identifica o objetivo do auditor referente à verificação se

determinado risco de fato existe e, por fim (3) escolhe e aplica o procedimento julgado adequado pelo auditor. Tais etapas antecedem a interpretação dos testes e elaboração do relatório (Yoshitake, 2009; Boynton & Johnson & Kell, 2002). Nesse sentido a essência da taxonomia seguiu o caminho lógico utilizado pelos auditores contábeis, respeitando essas três etapas como fundamentais na estrutura de classificação, sendo nomeadas como (1) Riscos, (2) Afirmações e (3) Procedimentos. Etapas com natureza conceituais distintas, porém inter-relacionadas no sentido lógico do processo de auditoria.

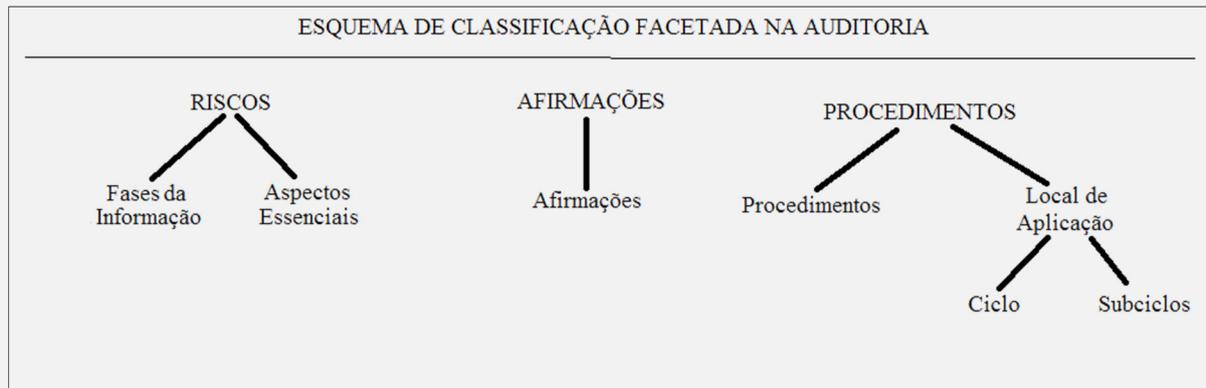


Figura 2. Esquema de Classificação Facetada na Auditoria.

Referência: elaboração própria.

Diante de naturezas distintas que transcende relações hierárquicas, nesse estudo foi desenvolvida a taxonomia nos moldes estruturais da classificação facetada exposta por Ranganathan (1960) e Vickery (1966), combinada com as relações lógicas proposta por Dahlberg (1978), formando assim uma trilha de auditoria que permite a organização das informações de riscos de distorções contábeis, mantendo a lógica processual da auditoria para permitir a identificação do procedimento recomendado para cada caso.

As relações apresentadas por Dahlberg (1978) são dadas pelas características em comum entre os conceitos, onde as relações lógicas atrelam os termos com o conceito da sua respectiva categoria e as relações semânticas entre as diferentes categorias de classificação. Interpretando essas relações lógicas à estrutura taxonômica facetada, as relações lógicas agrupam conceitos que compõem uma mesma faceta e, as relações semânticas conectam conceitos construindo o percurso de navegação.

As três etapas da figura 2 são relacionadas de maneira processual (percurso da navegação), onde originaram as seguintes facetas:

- **Fases da informação:** informações contábeis relacionadas às fases de reconhecimento, mensuração e evidenciação;
- **Aspectos essenciais:** características comuns que representam a essência de todas as distorções levantadas;
- **Afirmações:** aspectos associados às informações contábeis testadas pelos auditores;
- **Procedimentos:** testes substantivos aplicados pelo auditor para coleta e análise de riscos de distorções;
- **Ciclos de recursos:** categoria que classifica as informações em entradas e saídas econômicas e/ou financeiras; e
- **Subciclos de recursos:** subcategorias dos ciclos de recursos utilizadas em uma estrutura contábil abrangendo as informações contábeis.

Na relação de equivalência devem ser usadas notações que caracterizem a relação de equivalência entre os termos da taxonomia. As notações empregadas são:

- USE/NÃO USE: notação empregada para indicar o “termo” autorizado para uso; e
- UP: notação empregada para indicar “utilizado para”. Tem como finalidade indicar o “termo” não autorizado e apontar para o “termo” autorizado. As relações de equivalência representadas por USE/NÃO USE e UP são chamadas de remissivas;

Os nomes e conceitos de cada faceta, assim como os termos que as compõe são retirados do arcabouço conceitual contábil. O conjunto de termos de cada faceta contem aspectos que justificam o conceito da respectiva faceta. Os termos de uma faceta se relaciona com os das facetas subsequentes de maneira lógica através de uma relação entre os aspectos de cada termo. Isso respeita da Teoria do Conceito (Dahlberg, 1978) e preenche a observação das variáveis do objetivo específico 1: (a) aspectos conceituais para composição das categorias da taxonomia; (b) as relações lógicas entre atributos e conceitos; e, (c) as relações de equivalência, hierárquica e de associação entre os termos.

Tais relações lógicas são observadas entre as diferentes facetas, onde as relações entre elas correspondem a uma trilha na taxonomia (estrutura de linguagem documentária) que permite a classificação dos riscos identificados pelo auditor e, a partir da análise de seus aspectos conceituais continuar a classificação na sequência da estrutura até que haja a identificação e aplicação dos procedimentos de auditoria.

Após o desenvolvimento teórico, houve a validação técnica da taxonomia, onde riscos de distorções contábeis foram levantados na literatura contábil (num total de 75) para serem classificados. A medida que os riscos coletados são encaixados a cada termo nas respectivas facetas, a estrutura conceitual foi validada. O processo classificatório para validação consistiu classificar os riscos coletados a partir da análise de seus aspectos identificados na linguagem natural escrita na literatura contábil. Na classificação, para cada risco foram identificados seus procedimentos correspondentes, validando assim o processo entre análise do risco e a escolha dos procedimentos.

As informações para serem alocadas à trilha da estrutura taxonômica devem ser interpretadas observando suas características para a classe de entrada Fases da Informação. Nesse ponto, as informações em linguagem natural podem sofrer alterações de redação antes de serem inseridas no modelo. Algumas descrições de riscos mostram-se ambíguas quanto à classificação na faceta de entrada. Isso se dá pelo fato da linguagem natural conter ambiguidades e conflitos conceituais em uma mesma descrição. E ressalta a necessidade da linguagem documentária para reduzir a subjetividade na escolha e interpretação dos procedimentos.

Uma vez inseridos na taxonomia, devem seguir o restante da rota onde essas informações serão vinculadas às demais categorias posicionando-se para escolha dos procedimentos e seu identificação do respectivo ciclo de aplicação.

Após o desenvolvimento e validação técnica, houve a pesquisa documental para aplicação da taxonomia à um caso prático nos papéis de trabalho do Tribunal de Contas da União (TCU) referentes à determinada auditoria financeira. A análise de conteúdo foi utilizada para análise dos dados coletados (Mueller, 2007). Essa análise consistiu no exame de um objeto de comunicação, relatórios provenientes de auditoria, onde foi efetuada a (a) análise prévia para conhecimento do conteúdo, (b) separação as palavras e frases relacionadas à busca para pesquisa e (c) análise técnica para aplicação na taxonomia.

Segundo Bardin (1977), essa técnica contém como etapas a pré-análise, a exploração do material e o tratamento dos resultados, inferência e interpretação. Essa análise de conteúdo considerou a análise dos três pontos fundamentais que deram origem à taxonomia (Riscos, Afirmações e Procedimentos), permitindo assim sua validação prática.

No ano de 2016 o Tribunal de Contas da União (TCU) iniciou a aplicação das normas internacionais de auditoria para demonstrações contábeis. Incluindo o TCU, os órgãos públicos brasileiros não praticavam auditoria financeira das demonstrações contábeis de acordo com as exigências internacionais. O foco da auditoria no Brasil para esse setor era o caráter legal processual e para questões operacionais de uso do dinheiro público. Contudo, com o processo de harmonização das normas brasileiras às internacionais, o TCU iniciou a aplicação da auditoria focando as demonstrações contábeis. O processo TC-031.235-2015-8 cedido pelo TCU para essa pesquisa é um dos dois primeiros que o tribunal executou, iniciado em 2016 focando as informações contábeis.

Durante a análise de conteúdo fez-se uma leitura preliminar. Definiu como foco dois documentos no conjunto completo do processo de auditoria do TCU:

- 1) Matriz de Avaliação de Riscos - Fundo do Regime Geral de Previdência Social (FRGPS)
- 2) Matriz de Planejamento dos Procedimentos

Dentro do processo, três elementos foram observados para a coleta de dados, quais sejam (1) atributos relacionados à Análise dos Riscos de Distorções Contábeis, provenientes de erros ou fraudes, (2) Afirmações de auditoria e (3) Procedimentos, assim como as áreas contábeis para sua aplicação. Esses elementos são basilares na estrutura conceitual da taxonomia e, estão dentro de processos de auditoria que seguem as normas internacionais.

Em 77 casos observados no relatório do TCU, apenas 13 (16,88%) foram divergentes com as sugestões de classificação de acordo com o modelo conceitual da taxonomia proposta. Segundo as análises, essas divergências identificadas correspondem à generalização nas descrições dos riscos nos relatórios de auditoria, o que indica a subjetividade constatada na literatura.

Diante das divergências identificadas, concluiu-se que não apresentam comprometimento conceitual da taxonomia. Cada divergência foi analisada e constatou-se que exibem uma falta de fundamento conceitual do próprio relatório, assim como na literatura com ambiguidade e conflitos conceituais nas descrições dos riscos (linguagem natural). Essa constatação é reforçada com a informação nos papéis de trabalho de que essa auditoria necessita ser melhorada em função de ser a primeira vez que é aplicada nas regras internacionais. Uma suposta inexperiência nesse contexto, favorecendo a taxonomia desenvolvida.

Os riscos correspondentes às divergências tiveram suas características analisadas e observou-se que se reescritos de acordo com a estrutura da taxonomia, na faceta inicial, permitem a readequação na estrutura e então identificação do procedimento correspondente. Em outras palavras, isso é percebido ao verificar que a linguagem documentária força o desmembramento de algumas descrições de riscos coletadas em linguagem natural, aumentando a precisão de detalhes e classificando-as numa estrutura padrão, base para recuperação de informações sobre lições aprendidas.

Por sua vez, as convergências (83,12% dos casos no relatório do TCU) indicam a possibilidade de uso da taxonomia por auditores para decisões de quais procedimentos utilizar, já que a estrutura conceitual

desenvolvida está alinhada às decisões práticas dos auditores para os Riscos, as Afirmações e os Procedimentos.

Nesse sentido, essa estrutura taxonômica possibilita a interpretação dos riscos de distorção em linguagem natural, adequação à linguagem documentária para então identificação dos procedimentos de auditoria. Processo decisório do auditor a respeito dos procedimentos. Comporta ainda lições aprendidas sobre esses riscos, que podem ser alimentadas constantemente formando um banco de dados. As informações de lições aprendidas são então armazenadas e podem ser recuperadas por usuários dessas informações.

CONCLUSÕES

Ao estudar o uso da taxonomia como instrumento auxiliar para a escolha de procedimentos no processo de auditoria, é observado que suas bases conceituais são da gestão de documentos e sua estrutura como linguagem documentária torna-se nesse estudo um suporte para a gestão de risco na auditoria contábil. Assim, na ótica contábil combinada com a Ciência da Informação, tem-se a taxonomia como instrumento de gestão de documentos proporcionando subsídio conceitual, tendo em vista a gestão da informação na auditoria contábil e a tomada de decisões táticas na perspectiva da gestão do conhecimento.

Os resultados da aplicação demonstram que a estrutura taxonômica reflete o processo de decisão do auditoria, organizando as informações entre a análise de risco e escolha de procedimentos de auditoria. Processo no qual os aspectos conceituais compõem as categorias e as relações lógicas entre atributos e conceitos que as fundamentam, enquanto as relações de equivalência, hierárquica e de associação permitem consolidar a função processual de maneira que haja um lastro nesse processo decisório de escolha dos procedimentos. A linguagem padrão para confecção dessa taxonomia relaciona os termos de maneira que a identificação dos riscos de distorções e os aspectos determinantes para a escolha dos procedimentos formam uma trilha para recuperação de lições aprendidas. A padronização e o lastro são os fatores que permitem reduzir a subjetividade das escolhas (destacando o aspecto documental). Os resultados são limitados aos casos estudados, apesar de sua consistência sugerir sua replicação.

Duas funções para a taxonomia foram observadas. Primeiro, função decisória, a partir da coleta dos riscos de distorção o auditor pode seguir uma trilha e apoiar-se em bases teóricas para fundamentar a aplicação dos procedimentos, mesmo sem conhecimento sobre a instituição, mas a partir da linguagem utilizada para descrever o risco. Segundo, função de banco de dados, pois quando já estiver preenchida com lições aprendidas sua estrutura padronizada permite fazer com que os auditores recuperem essas informações, onde erros anteriores servirão de lição para os casos seguintes. Nesse sentido, a estrutura da taxonomia proposta tem um intuito maior em contribuir como bases conceituais para pesquisas futuras.

Quanto sua implicação prática, tem-se que a taxonomia transcende o aspecto documentário e aflora sua característica de suporte na organização da informação. A taxonomia de lições aprendidas pode ser utilizada durante a auditoria pelo próprio auditor que queira analisar novamente um determinado risco e seus procedimentos aplicáveis. Outra situação é quanto à supervisão, onde um auditor em posição superior revisa os papéis de trabalho e, com o auxílio da taxonomia pode verificar o caminho decisório

utilizado. Ademais, na supervisão de órgãos de controle e, a linguagem documentária orienta a análise da descrição de riscos auxiliando na organização das informações coletadas em linguagem natural.

REFERÊNCIAS BIBLIOGRÁFICAS

- Almeida, M. C. (2010). *Auditoria: um curso moderno e completo* (2ª ed.). São Paulo: Atlas.
- Aziza, B. & Kaplan, R. S. & Fitts, J. & Norton, D. P. (2013). *Drive business performance: enabling a culture of intelligent execution*. Microsoft Executive Leadership Series, v. 15. Hoboken, N.J.: Wiley. Acesso em 31/12/2016: <<http://rbdigital.oneclickdigital.com>>.
- Bardin, L. (1977). *Análise de conteúdo*. Edições 70: Lisboa.
- Baptista, D. M. & Araujo Junior, R. H. De & Carlan, E. (2010). O escopo da análise da informação. In Robredo, J. & Bräscher, M. *Passeios pelo bosque da informação: estudos sobre representação e organização da informação e do conhecimento EROIC* (pp. 62-80). IBICT: Brasília.
- Beasley, M. S. & Buckless, F. A. & Glover, S. M. & Prawitt, D. F. (2015). *Auditing Cases: an interactive learning approach* (6.ed.) Pearson.
- Boynton, W. C. & Johnson, R. N. & Kell, W. G. (2002). *Auditoria* (7.ed.) São Paulo: Atlas.
- Bräscher, M. & Carlan, E. (2010). Sistemas de organização do conhecimento: Antigas e novas linguagens. In Robredo, J. & Bräscher, M. *Passeios pelo bosque da informação: estudos sobre representação e organização da informação e do conhecimento EROIC* (pp. 147-176). IBICT: Brasília.
- Campos, M. L. A. & Gomes, H. E. (2008, Agosto) *Taxonomia e classificação: o princípio de categorização*. Data Grama Zero, v. 9, n. 4.
- Coleman, T. S. (2012). *Quantitative risk management: a practical guide to financial risk*. Wiley finance series, 669. John Wiley & Inc.. Hoboken. New Jersey. <<http://public.eblib.com/choice/publicfullrecord.aspx?p=817972>> Acesso em 31 dez 2016.
- Dahlberg, I. (1978). *Teoria do Conceito*. Ciência da Informação, Rio de Janeiro, 7 (2): 101-107.
- Elo, S. & Kyngas, H. (2008). *The Qualitative Content Analysis Process*. Journal of Advanced Nursing 62(1), 107–115. DOI: 10.1111/j.1365-2648.2007.04569.
- Flick, U. (2009). *Introdução à Pesquisa Qualitativa* (3 ed.). Porto Alegre: Artmed.
- Hsiao, Y. (2011). *An audit-based prediction model for aviation maintenance safety*. Dissertation (Doctor of Philosophy) - Faculty of the Graduate School of the University at Buffalo. New York.
- Longo, C. G. (2011). *Manual de auditoria e revisão das demonstrações financeiras: novas normas brasileiras e internacionais de auditoria*. São Paulo: Atlas.
- Matias-Pereira, J. (2007). *Manual de metodologia da pesquisa científica*. São Paulo: Atlas.
- Milton, N. (2010). *The Lessons Learned Handbook: Practical Approaches to Learning from Experience*. Woodhead Publishing Limited: Oxford.

- Mueller, S. O. M. (2007). *Métodos para pesquisa em Ciência da Informação*. Brasília: Thesaurus.
- Parnell, G. S. & Bresnick, T. & Tani, S. N. & Johnson, E. R. (2013). *Handbook of Decision Analysis. Wiley Handbooks in Operations Research and Management Science*, v. 4. Hoboken: Wiley. <http://www.123library.org/book_details/?id=62684> Acesso em: 28 dez 2016.
- Ranganathan, S. R. (1960). *Colon Classification* (6ª ed.). ISBN 81-7000-423-3 e 978-81-7000-423-3. Delhi.
- Rossi, G. B. & Serralvo, F. A. & Joao, B. N. (2014, Setembro). *Análise de Conteúdo*. Universidade Nove de Julho, São Paulo. 13.4: p39. DOI: <http://dx.doi.org/10.5585/remark.v13i4.2701>.
- Sousa, R. T. B. (2003). Os princípios arquivísticos e o conceito de classificação. In Rogrigues, G. M. & Lopes, Iza L. *Organização e representação do conhecimento*. Brasília: Thesaurus. pp.240-271.
- Sousa, R. T. B. & Araujo Junior, R. H. de. (2013, Janeiro). *A classificação e a taxonomia com instrumentos efetivos para a recuperação da informação arquivística*. Ciência da Informação, Brasília, DF, v. 41, n. 1, pp.148-160.
- Thomas, W. H. (2015). *The Basics of Project Management Evaluation and Lessons Learned* (2ª ed.) CRC Press: New York.
- Vickery, B. C. (1966). *Faceted Classification Schemes*. Graduate School of Library Service. Vol. V. New Jersey.
- Whittington, O. R. & Delaney, P. R. W. (2011). *CPA Exam Review 2012: Auditing and Attestation*. Hoboken, NJ: Wiley.
- Yoshitake, M. (2009). *Auditoria contábil: metodologia de processo de auditoria*. Curitiba: Juruá.



GESTÃO DO CONHECIMENTO E SISTEMAS DE GESTÃO DO CONHECIMENTO NAS MICROEMPRESAS: FATORES CRÍTICOS, REQUISITOS E RECOMENDAÇÕES

Paulo Madeira Gonçalves¹, Maria Cristina Vieira de Freitas²

¹Universidade de Coimbra, 0000-0002-1576-5749, paulo.goncalves@student.fl.uc.pt

²Universidade de Coimbra, 0000-0002-8849-8792, cristina.freitas@fl.uc.pt

RESUMO As pequenas e médias empresas representam uma grande parte do universo empresarial, dentro e fora de Portugal. A sua competitividade e sustentabilidade são estratégica, económica e socialmente relevantes para o desenvolvimento dos países e para a criação e a manutenção de emprego. Este estudo de revisão bibliográfica, de âmbito exploratório e descritivo, debruça-se sobre textos recentemente publicados e tem como objetivo identificar fatores críticos e medidas a considerar na implementação de processos de Gestão do Conhecimento, bem como requisitos e recomendações essenciais na adoção de Sistemas de Gestão do Conhecimento pelas microempresas. Os resultados obtidos apontam para a identificação de 11 fatores críticos de sucesso envolvidos nos processos de Gestão do Conhecimento, enumerados em diferentes graus de relevância, associados a cinco requisitos essenciais à adoção dos Sistemas de Gestão do Conhecimento. Como conclusão, refere-se a existência de um número residual de literatura científica visível no período consultado e relativa ao modo como as microempresas, especialmente em Portugal, usam e gerem o conhecimento, bem como aos seus impactos. Verifica-se, contudo, nos estudos analisados, uma relação entre os processos e as estratégias de gestão do conhecimento e as melhorias na *performance* global, na inovação, na sustentabilidade e na vantagem competitiva dessas empresas, pelo que se recomenda a sua adoção.

PALAVRAS-CHAVE *Gestão do Conhecimento, Sistemas de Gestão do Conhecimento, microempresas, fatores críticos de sucesso, requisito dos sistemas.*

ABSTRACT Small and medium-sized enterprises represent a large part of existing companies, inside and outside of Portugal. Their competitiveness and sustainability are strategically, economically and socially relevant for the country development and for employment creation and maintenance. This article aims to contribute in reviewing the scientific literature recently published about knowledge processes, strategies and knowledge management systems acquired and implemented by micro companies, through the identification of critical factors, measures and requirements relevant for that. It consists of a systematic data collection, throughout bibliographic databases and other publications available in Open Access. The results point out to the identification of 11 critical success factors involved in knowledge management processes and strategies, categorized in different levels of relevance, linked with five requirements needed in evaluating the adoption of knowledge management systems by the micro enterprises. The conclusions suggest that there are few studies visible, especially in Portugal, about the way micro enterprises use and manage knowledge and about the potential impacts caused in its overall performance, innovation and sustainability. However, the reviewed literature suggests a positive relationship between knowledge management processes and strategies and an improvement in global performance, innovation, competitive advantage and sustainability of micro enterprises, so that's why we recommend its adoption.

KEYWORDS *Knowledge Management Systems; Knowledge Management; micro enterprises; success critical factors; systems requirements.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Na classificação das empresas, segundo a dimensão, Portugal segue a Recomendação 2003/361/CE, emanada da Comissão Europeia. No artigo 2º desse dispositivo, a categoria das Pequenas e Médias Empresas (PME) é constituída pelas empresas que empregam menos de 250 trabalhadores e possuem um volume de negócio anual que não excede os 50 milhões de euros ou apresentam um balanço total anual que não excede os 43 milhões de euros. Enquadradas nessa categoria genérica, encontram-se as seguintes subcategorias: i) microempresas: com um máximo de 9 trabalhadores e um volume de negócios anual ou um balanço total anual igual ou inferior a dois milhões de euros¹; ii) pequenas empresas: com um máximo de 49 trabalhadores e um volume de negócios anual ou um balanço total anual igual ou inferior a 10 milhões de euros; iii) médias empresas: com um máximo de 249 efetivos e um volume de negócios anual igual ou inferior a cinquenta milhões de euros ou cujo balanço total seja igual ou inferior a quarenta e três milhões de euros (Comissão Europeia. Recomendação 2003/361/CE, 2003) (figura 1).

Classificação	Nº de Trabalhadores	Volume anual de Negócios	Balanço Total Anual
Média	< 250	≤ 50 milhões de euros	≤ 43 milhões de euros
Pequena	< 50	≤ 10 milhões de euros	≤ 10 milhões de euros
Micro	< 10	≤ 2 milhões de euros	≤ 2 milhões de euros

Figura 1. Indicadores de caracterização e diferenciação das PME.

Fonte – Elaboração nossa, a partir da Recomendação 2003/361/CE (Comissão Europeia, 2003).

Diante de tais indicadores, percebe-se que o tecido empresarial nacional e o internacional configuram-se pela predominância esmagadora dessa categoria de empresas. Trata-se de um grupo caracterizado pela sua relevante prestação econômica e social, e com um não menos considerável potencial de empregabilidade (Mateus et al., 2015; Antunes, 2009).

Mateus et al. (2015) confirmam essa realidade, recorrendo a dados de 2013, que mostravam que quase 800 mil PME empregavam nada menos do que oito, em cada dez trabalhadores, e contribuíam com o significativo valor de dois terços da riqueza gerada pelo setor não financeiro europeu. Nesse rol, relevavam, segundo esses mesmos dados, um tipo muito particular de PME: as microempresas (figura 1).

¹ Obedecendo às diretivas europeias, tratou o Decreto-Lei n.º 372/2007, alterado pelo Decreto-Lei n.º 143/2009, de regular que a essa categoria pertencem as empresas que empregam menos de 10 pessoas e que apresentam um volume de negócio ou balanço total anual igual ou inferior a dois milhões de euros.

Essa classe de instituições era a responsável, à época, por indicadores com valores igualmente significativos, tais como: 95% das empresas, 43% do emprego e 22% do valor acrescentado.

No contexto português, dados recentemente revelados pelo PORDATA (2016) dão conhecimento de resultados semelhantes. De acordo com essa instituição, as microempresas portuguesas representam 96,2% do universo das empresas não financeiras. Outros dados, recolhidos numa outra escala e divulgados mais recentemente, revelam que nos restantes países membros da Organização para a Cooperação e o Desenvolvimento Económico (OCDE, 2017), as microempresas constituem a grande maioria das empresas comerciais.

Do ponto de vista da dinâmica empresarial, contrariamente ao que ocorre com as empresas de grande porte, as microempresas caracterizam-se, de um modo geral: i) pelos escassos recursos que possuem para fazer face às necessidades; ii) pela predominância de estruturas horizontais e informais de direção; iii) pela formalização de poucas regras. Segundo Durst e Edvardsson (2012), esses fatores, aparentemente desvantajosos, quando bem equacionados, proporcionam um ambiente favorável ao empreendedorismo e à inovação, na medida em que potenciam o diálogo e a proximidade entre as pessoas envolvidas nos processos, tornando-as mais colaborativas e participativas. A agregar valor a essa afirmação, Nonaka, Toyama e Konno (2000) ponderaram, há quase duas décadas, que a inovação e o conhecimento traduziam-se em valores decisivos e em vantagem competitiva em contextos afetados pelas rápidas mudanças, tais como: mercados, produtos, tecnologias, regras e sociedades.

Com efeito, nos ecossistemas em que funcionam as microempresas, a gestão, as tomadas de decisão e as operações tendem a concentrar-se nas mãos de poucas pessoas, já que o número de trabalhadores é reduzido, em comparação com o de outras subcategorias das PME.

A contrastar com esses dados, que corroboram a relevância dessas organizações para o tecido económico e social dos vários países, identifica-se uma ausência de estudos, no contexto português, a indicar o modo como elas criam, usam e partilham o conhecimento e mesmo a analisar os impactos que a adoção e a implementação de processos e de estratégias de Gestão do Conhecimento (GC) poderão causar na sua *performance* global, nos seus processos de inovação e na sua sustentabilidade.

Sendo um valor, o conhecimento é definido como uma crença pessoal justificada, que aumenta a capacidade de um indivíduo para atuar de forma eficaz (Huber, 1991 cit. por Alavi & Leidner, 1999). Desta forma, está intimamente ligado à ação, sendo um processo dinâmico e contínuo, criado através de interações entre pessoas e organizações envolvidas num determinado contexto (Nonaka, Toyama & Konno, 2000).

A GC contribui para o sucesso das organizações, independentemente da sua dimensão e localização (Okunoye & Karsten cit. por Wang & Yang, 2016). Tomando como parâmetro a dimensão epistemológica, Nonaka e Takeuchi (1995) associam a criação do conhecimento e a produção de inovação nas organizações às interações entre o conhecimento tácito e o explícito. Também referem que, do ponto de vista ontológico, esta relação dá-se entre os três níveis de partilha e de difusão do conhecimento, sendo eles: individual, grupal e organizacional. A seu ver, a criação e a partilha de conhecimento são fatores diferenciadores, que potenciam a criação de novos modelos de negócio, novas formas de trabalho e de comunicação. O conhecimento, por sua vez, sendo um ativo valioso, deve ser gerido, de modo a poder ser utilizado e reutilizado, de forma efetiva, e assim gerar novo conhecimento, num ciclo contínuo, o que vem favorecer a criação e a inovação tão perseguidas pelas empresas, independentemente da sua dimensão.

Não obstante o facto de o número pequeno de colaboradores poder ser considerado, se bem aproveitado, uma vantagem competitiva, especialmente no desencadeamento de processos de GC – o que em parte se explica pelas relações de proximidade que se potenciam nesses ambientes –, quando esse número se traduz em real escassez, essa realidade se transforma. Ambientes com uma carga relativamente alta de contenção de recursos condicionam as oportunidades e o tempo investido na gestão de assuntos estratégicos, favorecendo o predomínio do conhecimento tácito sobre o explícito, que é inibido na presença desse tipo de constrangimentos (Durst & Edvardsson, 2012; I-Ching, Seidel, Shahbazpour & Howell, 2013). Nessas circunstâncias, os canais de comunicação passam a orientar-se mais para o meio externo do que para o interno, sendo este mais favorável às trocas e às partilhas, contrariamente ao que geralmente se observa nas empresas de maior dimensão, que necessitam de desenvolver um enfoque sistemático também na sua comunicação interna (Durst & Edvardsson, 2012). Por isso, e em face de uma tal conjuntura, as empresas de pequeno porte tendem a efetuar uma gestão pouco formalizada e sistematizada do seu próprio conhecimento, recorrendo a soluções mais pragmáticas e menos planeadas. Some-se o facto de a sua heterogeneidade, que se verifica em termos de setor de atividade e de recursos, dificultar a concretização de uma abordagem individualizada de GC, condicionando, nos casos em que esta exista, a adoção de modelos genéricos e, por isso mesmo, não raro, pouco eficazes, porque inadaptados à sua realidade.

Finalmente, nessa categoria de empresas, quer pela necessidade de desenvolver algum tipo de mecanismo de compensação das suas desvantagens, quer pela opção consciente pela adoção de medidas que visem o aumento da eficiência e da produtividade, nota-se que há um estímulo ao investimento na qualificação dos funcionários e na criação de parcerias, sendo estas algumas das formas frequentemente utilizadas para fazer face à escassez de recursos e à competição com empresas de maior dimensão, contribuindo essas medidas para a sua diferenciação² e, conseqüentemente, sustentabilidade.

Partindo desse enquadramento e dada a escassez de análises dessa natureza, no contexto português, pretende-se, com este estudo, de carácter exploratório e bibliográfico, identificar e analisar, do ponto de vista da literatura consultada, fatores críticos e medidas positivas a considerar na implementação de processos de Gestão do Conhecimento (GC) nessas empresas, bem como, dadas as suas particularidades, apontar alguns dos requisitos e recomendações essenciais na adoção de Sistemas de Gestão do Conhecimento (SGC), trazendo à luz alguns dos seus possíveis impactos.

METODOLOGIA

Este é um estudo de revisão, com um carácter exploratório, prescritivo e tendencial, na medida em que se debruça sobre textos mais recentemente publicados sobre as matérias de interesse.

A pesquisa bibliográfica que o ampara foi realizada de forma sistemática, buscando-se textos publicados, especialmente, na última década e recolhidos por meio de uma amostragem não

² Refira-se que o número de parcerias é também um importante indicador a usar na categorização das PME em subcategorias (cf. Antunes, 2009).

probabilística e intencional, em consultas bibliográficas efetuadas em textos publicados em Acesso Aberto e também nas seguintes bases de dados: i) Portal B-On³, ii) *Web of Science* e iii) *Google Scholar*⁴.

Os termos de consulta usados na pesquisa simples ou avançada, e efetuada em mais de um idioma, foram os seguintes: Conhecimento Organizacional; Gestão do Conhecimento Organizacional; Sistemas de Gestão do Conhecimento; Sistemas de Informação; Microempresas; Pequenas e Médias Empresas; Fatores Críticos de Sucesso; Portugal. Os filtros utilizados foram: Biblioteconomia e Ciência da Informação (área); Inglês e Português (idioma); Revistas acadêmicas, texto integral e revisto por especialistas (fontes). Como complemento, e em função da necessidade de efetuar o devido enquadramento das PME ao contexto português, realizou-se uma pesquisa em fontes documentais, sensivelmente, na legislação atual e pertinente (figura 2).

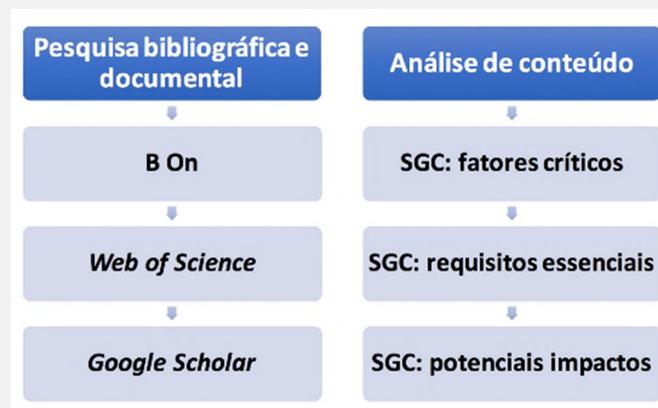


Figura 2. Matriz de recolha e de análise de dados do estudo.

Fonte - elaboração nossa.

Feitas essas diligências, foram recuperados 164 textos, dos quais e com base na leitura de secções específicas - resumo, introdução e conclusão - foram selecionados 31 itens para compor a revisão, tendo sido usados para compor a bibliografia apenas 21 textos.

Terminada a fase de recolha de dados, passou-se à fase de análise dos conteúdos (figura 2), que, neste estudo, limitou-se à identificação e reflexão sobre as ideias principais contidas nos textos pesquisados, que ajudaram a responder aos objetivos de pesquisa delineados, gerando-se, assim, os tópicos incluídos nos resultados, discussão e conclusão.

RESULTADOS E DISCUSSÃO

OS PROCESSOS DE GC E A SUA IMPLEMENTAÇÃO PELAS MICROEMPRESAS: FATORES CRÍTICOS E MEDIDAS POSITIVAS

³ Recuperado de: <http://www.b-on.pt>.

⁴ Esse foi consultado com o fim de verificar a análise de citações e de relevância dos textos recuperados. Recuperado de: <https://scholar.google.com>.

As Tecnologias de Informação e Comunicação, vulgo TIC, permitem o acesso a um conjunto alargado de meios usados na criação, no armazenamento, na transferência e na partilha de conhecimento (Alavi & Leidner, 2001). A escolha e o investimento nesses equipamentos, bem como a sua integração, devem, por isso, adequar-se às pessoas, aos processos e às necessidades das organizações. Contrariamente ao que sucede nas empresas de maiores dimensões, nas microempresas, face à sua reduzida dimensão, inexistem, muitas vezes, um departamento que assegure esses serviços e infraestruturas, sendo eles frequentemente realizados em regime de parceria ou em *outsourcing*, assumindo-se os riscos e os benefícios inerentes.

A estratégia organizacional, segundo Mintzberg (1987), consubstancia-se em planos ou linhas orientadoras, em que se verificam pelo menos duas características essenciais: i) elaboração com base em decisões antecipadas, relativamente às ações a que se vão aplicar; ii) implementação consciente e adequada aos objetivos organizacionais. Nas microempresas, a estratégia nem sempre é clarificada ou formalizada nem mesmo obedece a um acompanhamento regular, pelo facto de o número reduzido de colaboradores não proporcionar um meio profícuo para a sua implementação. Assim, e muito embora o proprietário da empresa saiba exatamente o que pretende, nem sempre se preocupa em desdobrar essas preocupações em aspetos fundamentais da estratégia, tais como: missão, valores e metas a atingir.

Não obstante, Edvardsson e Durst (2013) consideram a GC uma atividade sistemática e estreitamente vinculada à estratégia organizacional, associando-a a um conjunto de fatores considerados importantes para o sucesso das empresas, de um modo geral, e das microempresas, em particular, entre os quais citam, para além da já mencionada vantagem competitiva, os seguintes: crescimento sustentável e de vendas, aumento de lucros e de eficiência, flexibilidade, redução de custos e de perdas, ampliação da quota de mercado, desenvolvimento de funcionários, satisfação de clientes, desenvolvimento de novos produtos, inovação, criatividade, melhorias na tomada de decisões, no tempo de resposta em processos-chave e nas relações externas, sobretudo ao nível da comunicação.

Com efeito, a GC desenvolve-se num processo complexo de identificação, criação, armazenamento, transferência, utilização e reutilização do conhecimento (Durst & Edvardsson, 2012), sendo este um dos seus ativos mais valiosos. Considerada ela própria como um fator crítico na obtenção de vantagem competitiva, por parte das empresas, especialmente as de pequena dimensão (Wang & Yang, 2016), a adoção e a implementação de estratégias que visem à sua consecução tornam-se particularmente cruciais, devendo ser encaradas como um procedimento sistemático e contínuo, que passa pela criação de uma cultura de partilha e pela implementação de um conjunto de processos associados a sistemas de informação capazes de os suportar (Miklosik & Zak, 2015).

Também na aceção de Alberghini, Cricelli e Grimaldi (2010), os processos de GC são os que asseguram a sua própria execução consistente e eficiente, devendo aliar-se, por isso mesmo, às políticas, de modo a consolidar a cultura organizacional e alcançar benefícios para o negócio. No que respeita às empresas de pequena dimensão, eles exercem impacto na capacidade que elas possuem para fazer face aos seus próprios desafios. Na adoção das estratégias de GC que irão validar esses processos intervêm, segundo esses autores, outros dois fatores que, em termos do potencial de impacto, são considerados igualmente relevantes para o seu sucesso ou o seu fracasso, sendo eles: as pessoas e as tecnologias (figura 3). Conforme também concluem, no fator “pessoas” encontram-se implicados outros, tais como: gestão/liderança, envolvimento ativo dos gestores, cultura, experiências partilhadas, expectativas, regras e procedimentos informais, bem como hábitos sociais que influenciam a perceção e as comunicações

entre os funcionários. Por sua vez, no fator “tecnologias” encontram-se imbricadas as infraestruturas a adotar, devendo elas, no âmbito da GC, orientar-se para as pessoas, promovendo a intercolaboração.

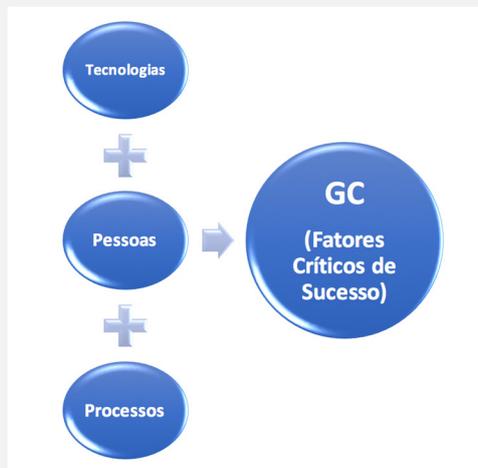


Figura 3. Fatores envolvidos na GC.

Fonte - Adaptado de: Alberghini, Cricelli e Grimaldi (2010).

De igual modo, e tendo como ponto de vista o facto de que as infraestruturas tecnológicas condicionam a implementação dos processos de GC, Wong e Aspinwall (2005) identificam o que designam como “Fatores Críticos de Sucesso (FCS)”, e que devem ser tidos em consideração nos projetos de desenvolvimento desses processos, aplicáveis às microempresas, estabelecendo uma ordem de prioridade entre eles, em razão do potencial de impacto, indo do mais importante (1) ao menos importante (11) (figura 4):

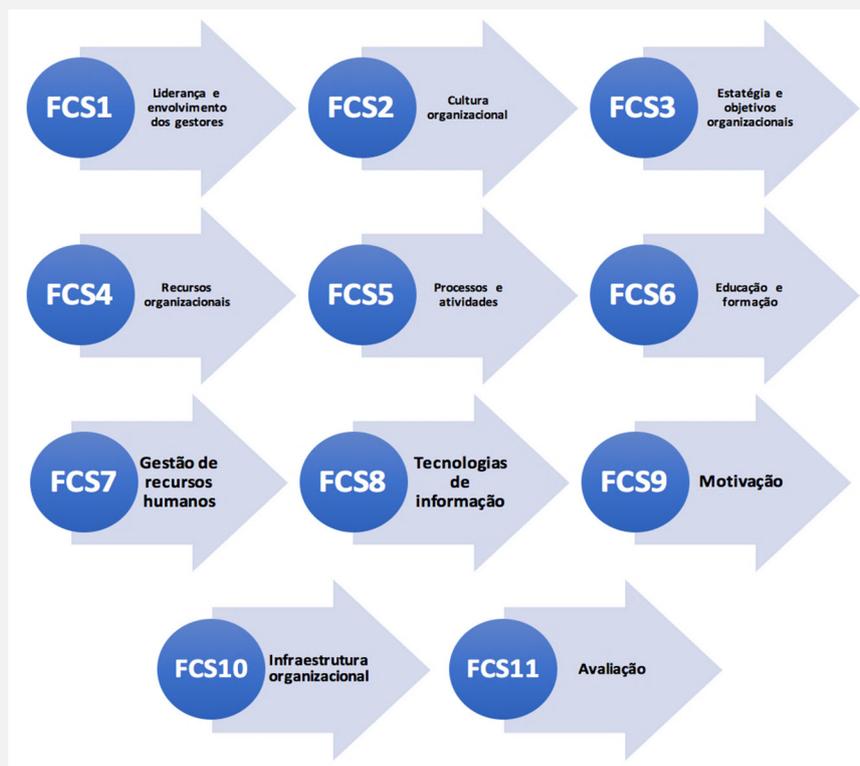


Figura 4. Fatores Críticos de Sucesso nos processos de GC, ordenados pela relevância.

Fonte – Adaptado de Wong e Aspinwall (2005).

Na perspetiva de Miklosik e Zak (2015), e de modo a evitar os previsíveis constrangimentos que se verificam na implementação desses mesmos processos, são sugeridas algumas medidas positivas que, pelo carácter genérico e preparatório do terreno, podem aplicar-se a diversas realidades, entre as quais as vivenciadas pelas microempresas, onde o seu impacto será certamente benéfico, e adotar-se sistematicamente (figura 5):

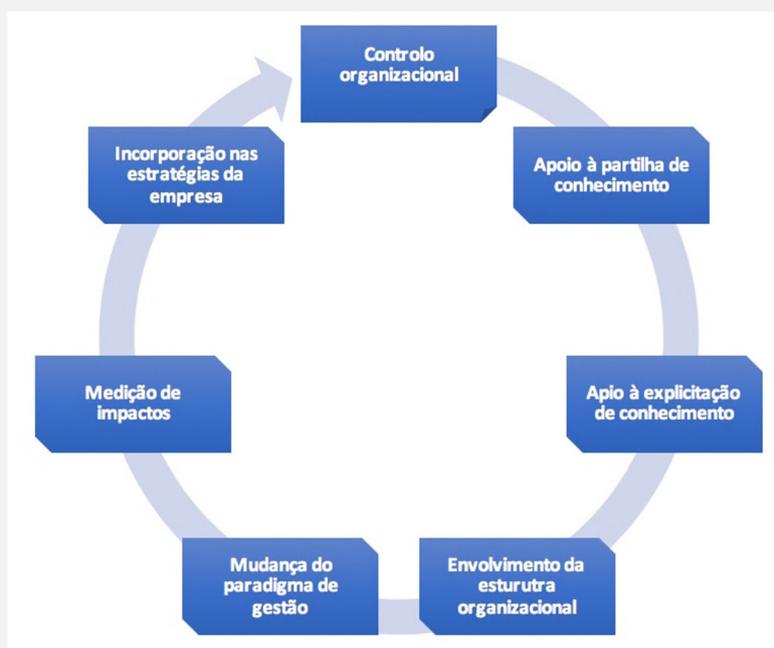


Figura 5 - Medidas que beneficiam a implementação de processos de GC.

Fonte – Adaptado de: Miklosik e Zak (2015).

Finalmente, e a modo de síntese, Miklosik e Zak (2015) encimam que a capacidade dos gestores para motivar os diversos intervenientes, internos e externos à organização, pode também contribuir, e de forma decisiva, para o sucesso ou para o fracasso destas iniciativas.

REQUISITOS E RECOMENDAÇÕES NA AQUISIÇÃO DE SGC PELAS MICROEMPRESAS

Os Sistemas de Gestão do Conhecimento (SGC) são plataformas utilizadas para gerir o conhecimento organizacional, baseadas em tecnologias desenvolvidas e aplicadas para promover e dar suporte aos processos de criação, armazenamento, transferência e utilização de conhecimento (Alavi & Leidner, 2001). Tratam-se, pois, de infraestruturas tecnológicas que, associadas às pessoas, condicionam a forma como se efetua a implementação dos processos de GC, tornando possível, desse entrecruzamento, a concretização de estratégias de GC.

Com a evolução tecnológica, têm surgido soluções adequadas às diversas categorias de empresas, capazes de potenciar os processos de socialização entre os diferentes intervenientes, interna e externamente, permitindo-lhes contornar os obstáculos financeiros e a escassez de recursos (Cerchione, Esposito & Spadaro, 2015).

Não obstante essas evidentes vantagens, que à partida e genericamente se encaixam no contexto particular das microempresas, as aquisições tecnológicas devem ser precedidas de um minucioso estudo,

sob pena de estar-se a investir recursos em ferramentas que, se usadas isoladamente, tornam difícil o acesso ao conhecimento e a interoperabilidade, e mesmo a sua concentração num único portal ou plataforma que facilite a sua integração. Para minimizar esse tipo de problema, é importante que na fase de aquisição do SGC sejam tidos em atenção fatores tais como as necessidades, os objetivos, as características e os recursos organizacionais, para além de atentar-se para o preenchimento de um conjunto de requisitos, por parte desses sistemas a adquirir, os quais envolvem, segundo Cagnazzo, Tiacci, Rossi, Industriale e Duranti (2014) diferentes aspetos a avaliar (figura 6):

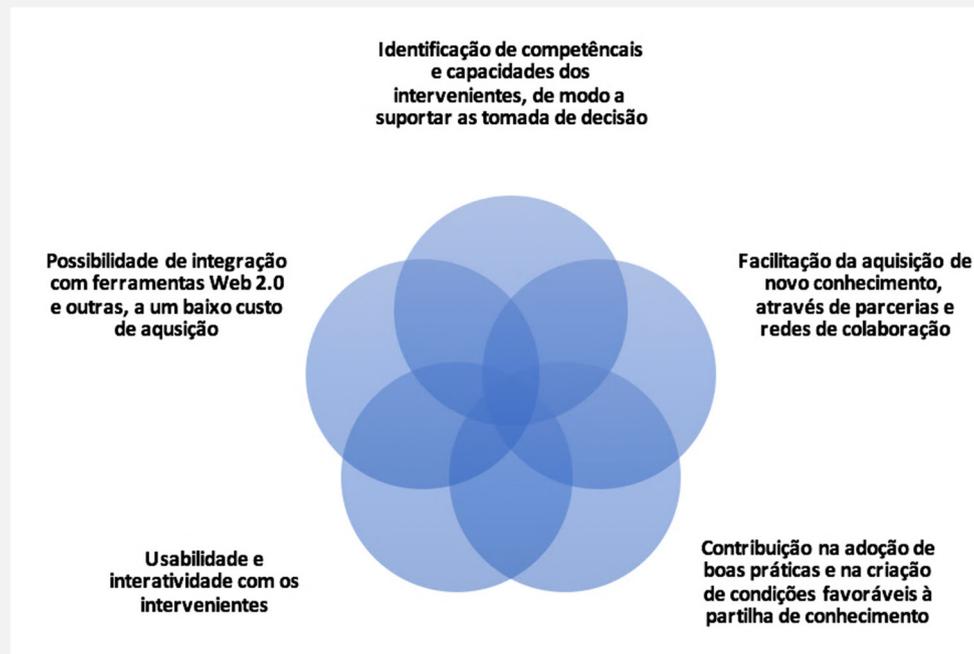


Figura 6. Requisitos essenciais a cumprir pelos SGC empresariais.

Fonte – Adaptado de: Cagnazzo, Tiacci, Rossi, Industriale e Duranti (2014).

No contexto atual, verifica-se que muitas das soluções comerciais de SGC disponíveis orientam-se para as empresas de média e de grande dimensão, representando custos elevados – de milhares ou dezenas de milhares de euros, – bem como a intervenção de recursos humanos especialmente dedicados à sua gestão, tornando-se por isso desajustadas às necessidades e aos meios disponíveis nas microempresas. Acresce que estas são soluções normalmente constituídas por vários módulos (*e.g.*, gestão de clientes, gestão de recursos humanos, gestão financeira, gestão comercial, logística, gestão de processos, gestão de produção, planeamento, gestão de existências), com pouca aplicabilidade na generalidade das microempresas, na medida em que são desenhadas para empresas em que existem recursos humanos suficientes para o que, vulgarmente, se designa por “alimentar o sistema”. Acresce, ainda, o facto de que várias destas soluções associam-se ao conceito de SaaS (*Software as a Service*), em que, para além do investimento inicial, existem custos de manutenção, de periodicidade variável, nem sempre suportáveis por empresas de escassos recursos. A heterogeneidade das PME é também um fator que dificulta a viabilização de um modelo único e que possa ser adotado como SGC adequado a todas as empresas integradas nesta classificação. Nesse caso, não se verificando a viabilidade na adoção de uma solução comercial, recomenda-se o desenvolvimento de um sistema sob medida e que permita uma integração de diversas ferramentas (*e.g.*, bases de dados, aplicações de gestão comercial, email, redes sociais, fóruns, blogs, vídeo conferência, ferramentas *Web 2.0*) numa única plataforma, evitando-se posições conflituantes entre pessoas e aplicações, relativamente às opções tecnológicas. Devem, pois,

ser integradas ferramentas que se complementem e que não sejam redundantes (Yuan, Zhao, Liao & Chi, 2013).

Em adição, também se recomenda que o processo de implementação desses sistemas nas microempresas seja gradual, dinâmico e desencadeado a médio e longo prazo, e que o seu acompanhamento seja continuamente ajustado aos objetivos e às necessidades, conforme o que foi dito anteriormente, permitindo assim manter e tirar partido de características (*e.g.*, escassez de recursos, reduzida dimensão, informalidade, flexibilidade) que são tão próprias dessa classe tão peculiar de empresas.

CONCLUSÃO

As pequenas e médias representam uma parte considerável do universo empresarial, dentro e fora de Portugal, contribuindo sobremaneira para o desenvolvimento. Os estudos analisados relacionam a melhoria ao nível da inovação, da vantagem competitiva e da sustentabilidade das empresas com as atividades de GC por elas implementadas. As microempresas não fogem a essa regra.

A GC deverá ser encarada pelas microempresas como parte integrante da sua estratégia, identificando-se objetivos e intervenientes, de acordo com as características, recursos e necessidades da organização. O compromisso e o envolvimento dos gestores são fatores indispensáveis na implementação e no acompanhamento desses projetos, de modo a criar uma cultura que promova a criação, a aquisição, a partilha e a utilização do conhecimento organizacional.

As estratégias de GC nas organizações envolvem o entrecruzamento de três importantes componentes: pessoas, processos e tecnologias. A correta identificação dos FCS, bem como dos requisitos essenciais dos SGC, facilitam a implementação e a utilização do conhecimento organizacional, tendo um impacto positivo em aspetos cruciais como a inovação, a competitividade e a sustentabilidade. A ênfase deve ser colocada nas pessoas, pois o seu envolvimento e a sua motivação para a criação, a partilha e a utilização do conhecimento são também encarados como requisitos essenciais dos SGC.

As microempresas caracterizam-se pela escassez de recursos, pela dimensão reduzida, pela informalidade e pela flexibilidade dos seus processos. A correta adequação destas características traduz-se em vantagem competitiva e não em limitações para o seu sucesso. Com isso, na escolha e implementação de um SGC, os requisitos e fatores críticos, bem como as medidas positivas que vão ao encontro dos processos e das estratégias de GC devem ser analisados e aplicados sob medida, não sendo aconselhável recorrer-se à replicação de modelos.

Ainda são relativamente escassos os estudos relacionados com o modo como as microempresas portuguesas usam e gerem o conhecimento organizacional e com os impactos que a implementação da GC e dos SGC causam na sua *performance* global, inovação e sustentabilidade. Consequentemente, na revisão bibliográfica realizada não foram obtidos resultados consistentes e que permitam identificar de que forma as microempresas portuguesas gerem, na prática, o seu conhecimento ou que modelos de SGC se encontram em curso nessas realidades, sendo essa uma linha a ser explorada futuramente.

REFERÊNCIAS BIBLIOGRÁFICAS

- Alavi, M., & Leidner, D. E. (1999). Knowledge management systems: issues, challenges, and benefits. *Communications of the AIS*, 1(7), 1–37. <http://doi.org/10.1002/jhrm.20064>.
- Alavi, M., & Leidner, D. E. (2001). Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly*, 25(1), 107–136. <http://doi.org/10.2307/3250961>
- Alberghini, E., Cricelli, L., & Grimaldi, M. (2010). Implementing Knowledge Management Through IT Opportunities: Definition of a Theoretical Model Based on Tools and Processes Classification. *Proceedings of the European Conference on Intellectual Capital*, 21–33. Recuperado de: <http://0-search.ebscohost.com.library.vu.edu.au/login.aspx?direct=true&db=bth&AN=49548997&site=bsi-live>
- Antunes, J. (2009). As pequenas e médias empresas. *Vida Económica*, 22-02-2009, 33. Recuperado de: <https://www.occ.pt/fotos/editor2/VidaEconomica20FevereiroII.pdf>.
- Cagnazzo, L., Tiacci, L., & Rossi, V. (2014). Knowledge management system in SMEs within stable enterprise networks. *WSEAS Transactions on Business and Economics*, 11(1), 155-174.
- Cerchione, R., Esposito, E., & Spadaro, M. R. (2015). The spread of knowledge management in SMEs: A scenario in evolution. *Sustainability*, 7(8), 10210–10232. <http://doi.org/10.3390/su70810210>.
- Comissão Europeia (2003). *Recomendação 2003/361/CE, de 6 de maio, relativa à definição de micro, pequenas e médias empresas*. Recuperado de: <https://www.iapmei.pt/getattachment/PRODUTOS-E-SERVICOS/Qualificacao-Certificacao/Certificacao-PME/Recomendacao-da-Comissao-2003-361-CE.pdf.aspx>.
- Decreto-Lei n.º 143/2009, de 16 de julho. Recuperado de: <https://www.iapmei.pt/getattachment/PRODUTOS-E-SERVICOS/Qualificacao-Certificacao/Certificacao-PME/Decreto-Lei-143-2009.pdf.aspx>.
- Decreto-Lei n.º 372/2007, de 6 de novembro. Recuperado de: <https://www.iapmei.pt/getattachment/PRODUTOS-E-SERVICOS/Qualificacao-Certificacao/Certificacao-PME/Decreto-Lei-372-2007.pdf.aspx>.
- Durst, S., & Edvardsson, I. R. (2012). Knowledge management in SMEs: a literature review. *Journal of Knowledge Management*, 16(6), 879–903. <http://doi.org/10.1108/13673271211276173>.
- Edvardsson, I. R., & Durst, S. (2013). Does Knowledge Management Deliver the Goods in SMEs? *Business and Management Research*, 2(2), 52–60. <http://doi.org/10.5430/bmr.v2n2p52>.
- I-Ching, L., Seidel, R., Shahbazzpour, M., & Howell, D. (2013). Knowledge Management in Small and Medium-sized Enterprises: A New Zealand Focus. *New Zealand Journal Of Applied Business Research*, 11(2), 1-21.
- Mateus, A., Mateus, J., Ferreira, N., Silva, C., Farinha, D., Rodrigues, H., ... Escária, V. (2015). *Três décadas de Portugal europeu: balanço e perspetivas*. Recuperado de: <https://infoeuropa.eurocid.pt/files/database/000066001-000067000/000066242.pdf>.

Miklosik, A., & Zak, S. (2015). Framework for Effective Removal of Knowledge Management Implementation Barriers. *Procedia Economics and Finance*, 30(15), 513–521. [http://doi.org/10.1016/S2212-5671\(15\)01263-0](http://doi.org/10.1016/S2212-5671(15)01263-0).

Mintzberg, H. (1987). The Strategy Concept I: Five Ps For Strategy. *California Management Review*, 30(1), 11–24. <http://doi.org/10.2307/41165263>.

Nonaka, I., Toyama, R., & Konno, N. (2000). SECI, Ba and Leadership: a Unified Model of Dynamic Knowledge Creation. *Long Range Planning*, 33(1), 5–34. [http://doi.org/10.1016/S0024-6301\(99\)00115-6](http://doi.org/10.1016/S0024-6301(99)00115-6).

OCDE (2017). *Enterprises by business size (indicator)*. Recuperado de: <https://data.oecd.org/entrepreneur/enterprises-by-business-size.htm>. doi: 10.1787/31d5eeaf-en.

PORDATA. (2016). *Pequenas e médias empresas em % do total de empresas: total e por dimensão: Portugal*. Recuperado de: <http://www.pordata.pt/Portugal/Pequenas+e+médias+empresas+em+percentagem+do+total+de+empresas+total+e+por+dimensão-2859>.

Wang, M. H., & Yang, T. Y. (2016). Investigating the success of knowledge management: An empirical study of small- and medium-sized enterprises. *Asia Pacific Management Review*, 21(2), 79–91. <http://doi.org/10.1016/j.apmr.2015.12.003>.

Wong, K. Y., & Aspinwall, E. (2005). An empirical study of the important factors for knowledge-management adoption in the SME sector. *Journal of Knowledge Management*, 9(3), 64–82. <http://doi.org/10.1108/13673270510602773>.

Yuan, Y. C., Zhao, X., Liao, Q., & Chi, C. (2013). The use of different information and communication technologies to support knowledge sharing in organizations: From e-mail to micro-blogging. *Journal of the American Society for Information Science and Technology*, 64(8), 1659–1670. <http://doi.org/10.1002/asi.22863>.



ESTUDO DE PROSPECÇÃO E VIGILÂNCIA TECNOLÓGICA PARA BUSCA E RECUPERAÇÃO DA INFORMAÇÃO

Maria do Carmo Oliveira Ribeiro¹, Fernanda Maria Melo Alves²

¹Doutoranda do Programa de Pós-graduação em Ciência da Informação da Universidade Federal da Bahia (UFBA). 0000-0003-0029-7636, mdocarmo@fieb.org.br

²Orientadora. Pós-doutoranda no Programa de Pós-graduação em Ciência de Informação da Universidade Federal da Bahia, Salvador, Brasil, com bolsa financiada pela CAPES. 0000-0002-8396-4053, fmelo2@hotmail.com

RESUMO O estudo apresenta conceitos e metodologias para a realização de estudos de prospecção, utilizando o documento de patente como fonte de informação tecnológica e artigos científicos para verificar o estado da arte de Robótica Autônoma Submarina (*Autonomous Underwater Vehicle*). O estudo visa demonstrar a importância da gestão da informação, como mecanismo de apoio ao monitoramento tecnológico para a tomada de decisão e o papel do profissional da informação no uso de estratégias de busca em bancos e bases de dados para localizar informação como matéria prima para a realização de estudos de prospecção. Os métodos de prospecção tecnológica iniciaram-se por buscas de anterioridade e similaridade em bancos de dados de patentes, nacionais e internacionais, através da ferramenta *Thomson Innovation*, base de dados de patentes e literatura científica, onde foram usados os termos de busca Robótica Autônoma Submarina e seus subtemas, com filtros por códigos da classificação internacional de patentes (IPC). Em seguida, selecionou-se um conjunto de patentes e artigos científicos, com os resultados apresentados em gráficos, que correspondem à fase da inteligência competitiva, com um conjunto de informações extraídas para geração de mapas de conhecimento, como subsídio para apoiar a tomada de decisão em relação aos investimentos no desenvolvimento do setor industrial. Com os resultados foi possível identificar a percentagem de documentos de patentes, oriundos da produção científica acadêmica, das empresas, das instituições de pesquisa e de pessoas físicas; e identificar suas rotas tecnológicas e as empresas detentoras dessas tecnologias no mundo.

PALAVRAS-CHAVE *Estudos de prospecção; Vigilância tecnológica; Busca e recuperação da informação.*

ABSTRACT The study presents concepts and methodologies for the accomplishment of search studies, using the patent document as source of technological information and scientific articles to verify the state of the art of Autonomous Underwater Vehicle. The study seeks to demonstrate the importance of the administration of the information, as a mechanism to support technological monitoring for decision-making and the role of the professional of information using search strategies in databases to locate information as a raw material for the accomplishment of search studies. The methods of technological search began for anteriority and similarity searches in databases of patents, national and international, through the Thomson Innovation tool, database of patents and scientific literature, where were used the terms of search Underwater Autonomous Robotics and their subthemes, with filters for codes of the international classification of patents (IPC). After, a group of patents and scientific articles it was selected, with the results presented in graphs, corresponding to the phase of the competitive intelligence, with a group of extracted information for the generation of knowledge maps, as subsidy to support the decision-making in relation to the investments in the development of the industrial section. With the results it was possible to identify the percentage of documents of patents, originating from of the scientific academic production, of the companies, of the

research institutions and of physic persons; and to identify their technological routes and the companies holders of these technologies in the world.

KEYWORDS *Prospecting studies; Surveillance technology; Search and retrieval of information.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

O desenvolvimento e especialização da ciência e da tecnologia, e do conhecimento em geral, aumentou exponencialmente o número de publicações científicas, fato que dificulta o processo de seleção e uso da informação das fontes e recursos de informação, que podem ser de natureza, classificação e suportes variados e contribuem para o desenvolvimento das atividades de seus usuários.

Entende-se por prospecção informacional o método ou técnica que visa a identificação inicial de dados, informação e conhecimento relevantes para a organização. Vigilância tecnológica é o método ou técnica de observação e acompanhamento constante de dados, informação e conhecimento relevantes ao negócio da organização.

A prospecção tecnológica, por meio de pesquisa em bases de dados de patentes e artigos técnico-científicos, tem se mostrado uma importante ferramenta de apoio à decisão, tendo em vista que seu conteúdo informacional permite identificar, fundamentalmente, informações sobre o desenvolvimento tecnológico e científico da área de interesse, que permitam a pesquisadores e tomadores de decisões, ter uma “visão” do estado da arte e sua possível evolução futura de determinada área, segmento ou tema.

Os mapas tecnológicos, são ferramentas que emergiram, nos últimos anos, para explorar a dinâmica das tecnologias emergentes nas indústrias, em um horizonte de longo prazo e, especialmente, desenvolver e executar mapas estratégicos, de modo a articular alinhar a estratégia da empresa às suas capacidades tecnológicas.

A aplicação dessa metodologia permite, também, obter informações mais específicas, como por exemplo, tecnologias relevantes, possíveis parceiros para pesquisas, nichos de mercado para atuação, inovações incrementais e movimentos de concorrência (CANONGIA, 2004).

A Ciência da Informação contribui expressivamente nesta temática, contribuindo para a construção e análise de indicadores científicos e tecnológicos de qualidade e confiabilidade. Os estudos métricos da informação são importantes para o reconhecimento dos temas que estão sendo pesquisados, e, principalmente, para demonstrar a influenciadessa produção no meio científico, aspetos promissores para a análise da dinâmica e das relações da Ciência, Tecnologia e Inovação (CT&I) em qualquer área do conhecimento. Também fornecem elementos que podem subsidiar tomadas de decisão em políticas, científicas e tecnológicas. A avaliação da produção intelectual pode contribuir para o delineamento de campos ou áreas específicas do conhecimento, detectando tendências temáticas, metodológicas, e até mesmo a evolução do conhecimento em frações temporais e espaciais.

O processo de busca, análise, interpretação e uso da informação, fato que cria conhecimento e permite uma sensível transformação e desenvolvimento de conceitos e sua aplicação sistemática. Essa visão demonstra o potencial do estudo de prospecção tecnológicas, principalmente no que se refere ao

tratamento automático da informação, como ferramenta de apoio ao processo decisório, possibilitando uma visão sistêmica sobre as dinâmicas da pesquisa, desenvolvimento e inovação de um setor e/ou tecnologias, em curto espaço de tempo (AMPARO et al, 2012).

Le Coadic (2004) explicita que o objeto social de estudo da Ciência da Informação é a “informação”. Ressalta que a dinâmica da informação passa por um processo de explosão quantitativa e implosão do tempo para sua comunicação e uso. A partir dessas premissas, consideram-se de interesse aspectos como o fluxo de informação e os processos de comunicação científica, destacando-se a colaboração científica, a comunicação formal e informal, a comunidade científica, as redes sociais, e outros.

A visão orientada para o futuro é o caminho para uma melhor sustentabilidade e fortalecimento da capacidade dos indivíduos, organizações e países, de modo a aproveitarem as oportunidades futuras numa economia globalizada. Na área tecnológica, essencial para qualquer projeto de desenvolvimento, os estudos prospectivos são essenciais. As mudanças tecnológicas, cada vez mais rápidas, indicam que ainda temos um longo caminho para percorrer, tornando-se necessário a utilização de estratégias alternativas para orientar o futuro, entre as quais os estudos de prospecção tecnológica.

O termo prospecção tecnológica designa atividades de prospecção centradas nas mudanças tecnológicas, em mudanças na capacidade funcional ou no tempo e significado de uma inovação. Visa incorporar informação ao processo de gestão tecnológica, tentando predizer possíveis estados futuros da tecnologia ou condições que afetam sua contribuição para as metas estabelecidas. Castells (1999) defende que o surgimento da economia da informação caracteriza-se pelo desenvolvimento de uma nova lógica organizacional baseada na convergência e na interação entre as TIC, os modelos de gestão e as articulações crescentes de redes de pessoas e empresas.

Contudo, entre a decisão estratégica de investir e a apropriação dos ganhos de competitividade resultantes, interpõe-se um grande número de etapas, cujo ponto crítico é a incerteza. Para Coelho (2005) a prospecção de tecnologia, por meio da gestão de informação tecnológica, é extremamente útil para apresentar o estado-da-arte de determinada área tecnológica, com o objetivo de gerar informações sobre a sua trajetória passada e sobre as tendências de mercado e percepção de sinais fracos.

A gestão da informação associada aos avanços das tecnologias vem sendo cada vez mais valorizada e seu uso sistemático é considerado como investimento face ao potencial de agregação de valor e geração de novos saberes (CANONGIA et al., 2004). A informação tecnológica necessária para os estudos de prospecção está disponível nos bancos de dados de artigos científicos e de patentes, um recurso valioso e confiável. Considera-se vantajoso o uso desta fonte de informação tecnológica, principalmente a facilidade de acesso às bases de dados disponibilizadas gratuitamente na internet.

Para Valentim (2003), a etapa que segue a prospecção exige conhecimento do negócio, conhecimento especializado da área de negócio e conhecimento informacional, de modo a filtrar apenas o que de fato é relevante ao negócio, para somente nesse momento iniciar a análise e tratamento da informação. Tal etapa deve ser realizada por uma equipe multidisciplinar integrada ao ambiente organizacional.

Os estudos de prospecção tecnológica, também chamados de estudos de futuro, ou *forecast (ing)*, *foresight (ing)* ou *future studies*, fornecem as principais tendências no contexto mundial sendo possível segmentar estas tecnologias por setor da economia. Estes estudos auxiliam a identificação de tecnologias promissoras, úteis para uma determinada organização, bem como apontam para possibilidades de negócios e parcerias (Borschiver, 2008).

O termo monitoramento ou vigilância tecnológica tem origem na corrente americana, similar a corrente francesa *veille* (vigília) e “designa uma técnica que tem como característica principal a observação e coleta (monitoramento) de fatores científicos, telemáticos, tecnológicos e outros que possam afetar a organização de forma intensa” (Tarapanoff, 1995).

A partir da década de 1990, robôs submarinos, que navegam de forma autônoma, começaram a surgir para auxiliar nas pesquisas oceanográficas e na exploração de petróleo e gás no fundo do mar. Com sensores para navegação submersa e GPS, quando estão na superfície, além de motores e equipamentos de comunicação por rádio, eles podem ser programados para ir e voltar de um local predeterminado. Receberam o nome de Veículos Autônomos Submersos (AUV, sigla para *Autonomous Underwater Vehicle*) e são produzidos por empresas de vários países, como Estados Unidos, Noruega, Japão e França.

No Brasil, as pesquisas e os desenvolvimentos nessa área são recentes e existem pelo menos três protótipos atualmente em fase de testes.

Todas as operações submarinas, efetuadas a uma profundidade limite para mergulho humano, são realizadas remotamente, por ter um custo muito alto enviar sondas ou submarinos tripulados para tais profundidades, colocá-los no lugar certo, conectar e desconectar cabos, ligar e desligar válvulas ou até mesmo estar lá apenas como observador, enquanto todos os comandos são enviados da superfície, a partir de uma plataforma de perfuração ou exploração.

A estes aspectos devem-se acrescentar outros riscos envolvidos no trabalho efetuado a tal profundidade, os equipamentos com suporte a vida são grandes, pesados e extremamente caros. Portanto, um dos objetivos da exploração submarina foi sempre descobrir estratégias de realização de tarefas de forma segura e à distância, seja da superfície imediatamente acima, seja, de uma confortável sala de controle em terra.

Tais fatos justificam a escolha do tema Veículos Autônomos Submersos como objeto de estudo do presente trabalho e utilizou como ferramenta de busca a ferramenta *Thomson Innovation*, que permite o acesso a um dos maiores e mais completo banco de dados de pesquisas científicas do mundo e está dividida em três bases de dados: patentes, literatura científica e negócios.

A base de patentes contém textos completos de pedidos de patentes, publicados nos EUA, Europa, WIPO (Organização Mundial de Propriedade Intelectual), Inglaterra, França, Alemanha, Japão e Coreia; históricos de arquivos de patentes, dados bibliográficos do INPADOC (Documentação Internacional de Patentes). Enquanto a base de literatura científica permite o acesso a Web of Science, a ISI Proceedings e a Current Contents Connect, conforme Quadro 1.

Quadro 1 - Scientific Literature Collections on Thomson Innovation

<i>Collection</i>	Description	Source	Dates Covered	File Size	Update Frequency
<i>Web of Science</i>	Offering retrospective searching of scientific literature, includes Science Citation Index, Social Sciences Citation Index and	Thomson Reuters	1898 - present	48 million records	Daily per business week

	Arts and Humanities Citation Index				
Current Contents Connect	Offering complete tables of contents, abstracts and bibliographic information from leading scholarly journals and books.	Thomson Reuters	1998 - present	11,000 journals 2,000 books	Daily per business week
Conference Proceedings	Helps researchers access the published literature from the most significant conferences, symposia, seminars, colloquia, workshops and conventions worldwide.	Thomson Reuters	1990 - present	6,5 million papers from 150,000 conferences	Weekly
Inspec	Relevant journal articles, conference papers and other documents for current awareness, new product information, technological forecasting, competitive intelligence or patent-related searching.	Institution of Engineering and Technology (IET)	1898 – present	13 million records	Weekly

Fonte: Elaboração das autoras a partir de informações extraídas da Thomson Innovation

O quadro apresenta as coleções de literatura científica da ferramenta *Thomson Innovation*, que inclui uma breve descrição, as fontes de informação, a data inicial da indexação dos dados, o período de atualização das bases e a quantidade de informações armazenadas em cada base.

O objetivo deste artigo é apresentar métodos de estudos de prospecção e vigilância tecnológica, através de buscas em bancos de dados de patentes e de artigos científicos, usando a ferramenta Thomson Innovation. Procura-se obter informação e apresentar o estado-da-arte das tecnologias sobre Robótica Autônoma Submarina e seus subtemas.

METODOLOGIA

A pesquisa é exploratória e descritiva com uma abordagem qualitativa e quantitativa. Apresenta um estudo de caso, tendo usado os métodos de prospecção tecnológica como procedimento de coleta de dados. Iniciou-se a pesquisa por buscas de anterioridade e similaridade em bancos de dados de patentes e artigos científicos, nacionais e internacionais no Portal de Periódicos da CAPES e no Google Acadêmico e na ferramenta *Thomson Innovation*, utilizando-se o documento de patente com fonte de informação tecnológica e os artigos científicos, para verificar o estado da arte da área de Robótica Autônoma Submarina.

Este trabalho desenvolveu-se em duas etapas. A primeira, a “fase pré-prospectiva”, de pesquisa preliminar, é uma busca menos direcionada, procurando informações acerca do objeto de estudo e que pretende a constituição de uma base de informações, que suportará a próxima fase. A segunda etapa, a “fase de prospecção tecnológica”, é uma busca mais direcionada, baseada em uma metodologia definida

com palavras-chave específicas, buscas em documentos técnicos (artigos científicos e patentes) e é acompanhada por uma análise mais detalhada, segundo vários critérios pré-definidos, tais como ano de publicação, país de origem, tipo de autor e foco sobre o objeto de estudo.

A pesquisa na ferramenta *Thomson Innovation*, realizou-se nas bases de dados “patentes” e “literatura” separadamente. A busca utilizou o descritor principal “*Autonomous Underwater Vehicle*” e seus subtemas “*Underwater Inspection / Underwater Navigation / Autonomos Underwater Docking*”, utilizando a estratégia de busca com descritores em campos específicos comou: título e abstract *Derwent World Patents Index* (DWPI), termos técnicos normalizados conforme a Classificação Internacional de Patentes (IPC) e período da busca para identificar documentos de patentes, suas rotas tecnológicas e empresas detentoras das referidas tecnologias no mundo.

Nesta etapa, efetuou-se uma análise das tecnologias priorizadas para cada uma das áreas em estudo. Os critérios de análise das tecnologias identificadas foram os seguintes: tema central analisado, número de publicações relacionadas; subtemas analisados, número de publicações relacionadas; tendências das publicações – evolução de patentes e artigos, conforme dados apresentados no Quadro 2. Os assuntos/temas foram definidos por especialistas na área de robótica submarina e o período da busca corresponde ao espaço temporal compreendido entre 2000 e 2016.

Os resultados são apresentados em vários quadros e figuras, os quais correspondem à fase da inteligência competitiva, isto é, um conjunto de informações extraídas para geração de mapas de conhecimento, como subsídio para apoiar a tomada de decisão em relação aos investimentos no desenvolvimento do setor industrial. Com os resultados alcançados foi possível identificar o percentual de documentos de patentes, oriundos da produção científica acadêmica, das empresas, das instituições de pesquisa e de pessoas físicas e identificar suas rotas tecnológicas e as empresas detentoras das referidas tecnologias no mundo.

A análise e o tratamento das informações filtradas pela prospecção tecnológica exigem conhecimentos e competências investigadoras e informacionais. Por essa razão, esta etapa é realizada geralmente por equipes multidisciplinares, que possibilitam um trabalho integrador, visando agregar valor aos dados, informação e conhecimento de forma mais eficaz e eficiente. Essa etapa envolve análise, reflexão, síntese, contexto e mediação, sendo, portanto, uma das etapas mais demoradas do estudo de prospecção tecnológica, cujas atividades são de natureza complexa e necessitam de recursos tecnológicos. O monitoramento, etapa posterior à análise e tratamento de dados, requer uma atividade diária do acompanhamento de dados, informação e conhecimento (Valentim & Molina, 2004).

Através do presente mapeamento de prospecção tecnológica sobre Robótica Autônoma Submarina e seus subtemas, utilizando a ferramenta *Thomson Innovation*, alcançaram-se resultados avançados. Além disso, foi possível avaliar e apresentar o estado-da-arte dos estudos e aplicações da Robótica Autônoma Submarina no mundo, identificando os pontos relevantes e as novas tecnologias que podem diminuir o tempo de processo e direcionar os investimentos em projetos de pesquisa e de desenvolvimento.

RESULTADOS

As pesquisas foram realizadas nas bases de dados de patentes e de literatura científica da ferramenta *Thomson Innovation*. Para definir a estratégia de busca, utilizaram-se descritores previamente selecionados, campos representativos, período definido, filtros disponíveis como recursos da ferramenta utilizada. Além desses recursos, utilizaram-se os números da Classificação Internacional de Patentes (IPC) para as buscas nas bases de dados de documentos de patentes, que facilitam a identificação dos documentos com maior similaridade com o objeto da pesquisa.

O Quadro 2 apresenta a matriz de estratégia de busca utilizada para a pesquisa bibliográfica nas bases de dados de artigos científicos e de documentos de patentes da ferramenta *Thomson Innovation*.

Quadro 2 - Matriz de estratégia de busca

Descritores	Ferramenta	Base de dados	Campo	Período	IPC (patentes)	Filtro
Autonomous Underwater Vehicle	Thomson Innovation	Patentes Artigos	Title/ Abstract DWPI	2000 a 2016	B63G 8/00 G01V 1/38	Empresas Instituições Países Depositantes
AUV	Thomson Innovation	Patentes Artigos	Title/ Abstract DWPI	2000 a 2016	B63G 8/00 G01V 1/38	Empresas Instituições Países Depositantes
Underwater Inspection / navigation	Thomson Innovation	Patentes Artigos	Title/ Abstract DWPI	2000 a 2016	B63C 11/42	Empresas Instituições Países Depositantes
Autonomous underwater docking	Thomson Innovation	Patentes Artigos	Title/ Abstract DWPI	2000 a 2016	B63B 27/36	Empresas Instituições Países Depositantes

Fonte: Elaboração das autoras

RESULTADOS DA PESQUISA NAS BASES DE DADOS DE DOCUMENTOS DE PATENTES

A pesquisa de anterioridade da tecnologia “*Autonomous Underwater Vehicle*” foi realizada na base de dados de patentes da ferramenta já indicada. A estratégia de busca identificou 488 documentos de patentes com o tema principal “*Autonomous Underwater Vehicle*” no campo “*Abstract DWPI*”. Em seguida, utilizando-se os termos “*AUV*”, *Underwater Inspection / Navigation / Docking* no campo

“Title DWPI”, alcançaram-se 281 documentos de patentes, publicados no mundo no período de 2000 a 2016, conforme dados abaixo.

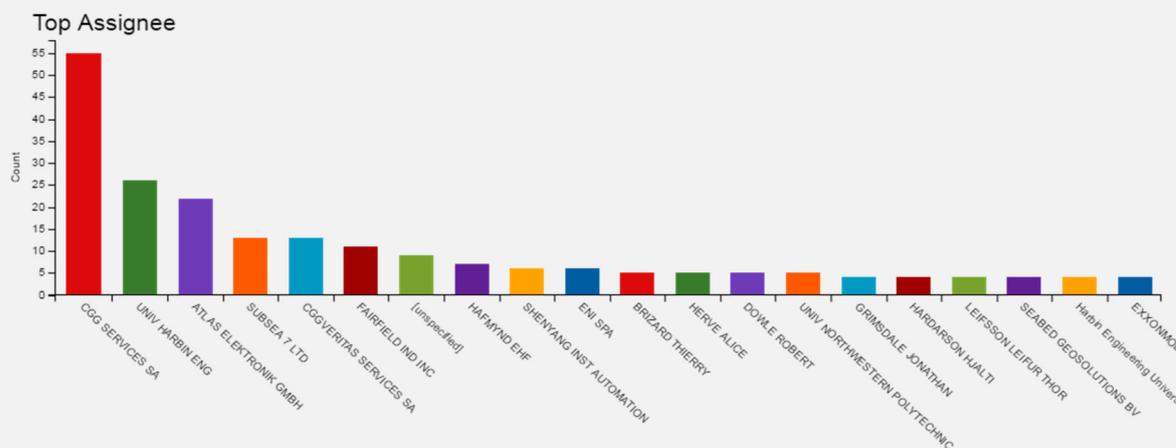


Figura 1. Principais depositantes de patentes (2010-2016)

De acordo com a Figura 1, os principais depositantes de patentes de Veículos Autônomos Subaquáticos são as empresas CGG Services, Univ. Harbin Eng, Atlas Elextronic e outras, cujos pedidos, registrados em vários países, se apresentam na Figura 2 abaixo.

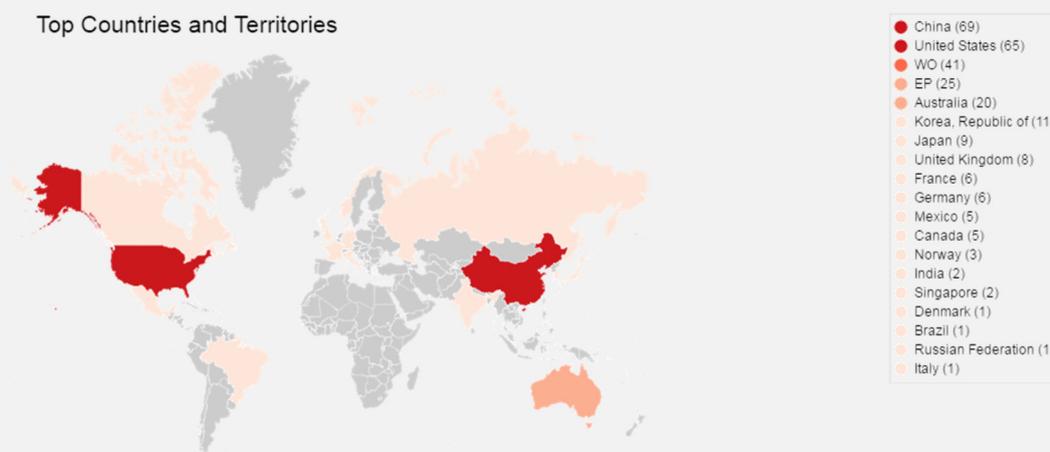


Figura 2. Distribuição por países de patentes sobre Veículos Autônomos Subaquáticos (2010-2016)

Conforme dados apresentados na Figura 2, a China domina a tecnologia de construção de Veículos Autônomos Subaquáticos com 69 registros de patentes, os Estados Unidos estão no segundo lugar com 65 registros de patentes, seguidos da Europa com 25 registros. Os 41 registros de patentes com a sigla WO são de famílias de patentes registradas na Organização Internacional de Propriedade Intelectual (OMPI/WIPO), através do *Patent Cooperation Treaty* (PCT), que permite o depósito de patentes em aproximadamente 145 países durante 30 meses, conhecida como fase internacional. Após os 30 meses, os depositantes devem indicar os países para entrada no pedido de patentes na fase nacional.

Em relação a totalidade dos pedidos de patentes realizados por ano, durante o período da pesquisa de anterioridade, que foi de 2000 a 2016, foram apresentados os seguintes resultados. Houve uma evolução inicial no período de 2001 a 2003 e em seguida um declínio entre 2004 e 2005. Uma nova evolução foi

apresentada no período de 2008 a 2010. Mas, o maior número de pedidos de patentes foram realizados no período de 2014 e 2015.

Em relação a Classificação Internacional de Patentes (CIP) o maior número de registros de patentes foram indexados na classificação B63G 8/00 (Barcos submersos / submarinos) com 73 pedidos. Em seguida 61 registros de patentes na classificação G01V 1/38 e 51 registros de patentes na classificação B63C 11/42 (Equipamento para permanência ou trabalho embaixo d'água com propulsão independente ou controle direcional).

RESULTADOS DAS PESQUISAS NAS BASES DE DADOS DE LITERATURA CIENTÍFICA

A pesquisa realizada nas bases de literatura científica da ferramenta *Thomson Innovation* apresentou um resultado de 5.703 documentos publicados no período de 2000 a 2016. Para a estratégia de busca utilizou-se o descritor “*Autonomous Underwater Vehicle*” em “*All Text Fields*”, que recuperou os seguintes tipos de documentos: artigos, bibliografias, livros, material editorial, relatórios, seminários, etc. Em seguida, utilizando a sigla “AUV” no mesmo campo, alcançaram-se 3.173 documentos publicados no mesmo período.

A mesma estratégia de busca utilizando como um diferencial no filtro por “tipo de documento” identificou 896 registros de artigos científicos, para o período em análise. Visando aperfeiçoar o resultado da pesquisa, utilizou-se mais um filtro com um descritor importante, o termo “*inspection*”, através do qual se alcançaram 133 registros de artigos científicos publicados, cujas análises estão representadas nas figuras abaixo.

No período de 2000 a 2016, houve uma evolução no período de 2000 a 2002. Em seguida houve uma certa estabilidade no número de artigos publicados entre 2005 a 2010 com uma média de 10 artigos por ano. Mas, 2015 foi o ano com o maior número de artigos científicos publicados sobre “*Autonomous Underwater Vehicle*”.

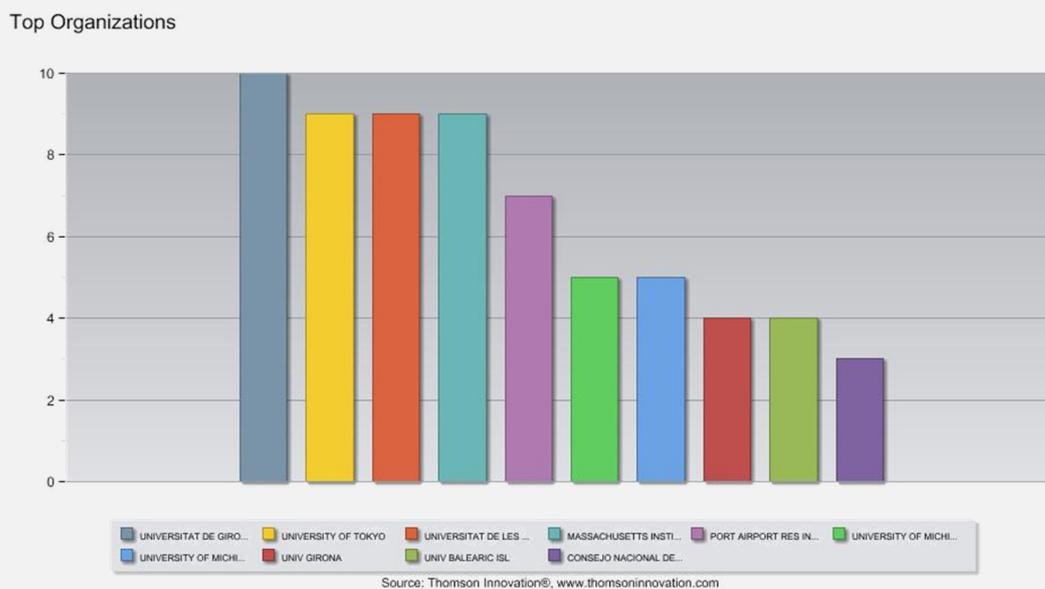


Figura 3. Principais organizações com maior número de artigos publicados

De acordo com a Figura 3, as principais instituições com maior número de artigos científicos publicados são: Universidad de Girona (Espanha) com 10 artigos; a University of Tokyo, a Universidad de Islas Baleares (Espanha) e o Massachusetts Institute of Technology (Estados Unidos), três instituições com 9 artigos publicados.

A evolução do número de publicações de artigos científicos por organizações / ano, apresenta destaque para a Universidade de Girona que publicou em 2004, 2005, 2007, 2011 e 2015; a University of Tokyo em 2002, 2004, 2005, 2009 e 2012; a Universidad de las Islas Baleares em 2000, 2002, 2005, 2008, 2009, 2011 e 2014 e o MIT em 2002, 2004, 2005, 2009 e 2012.

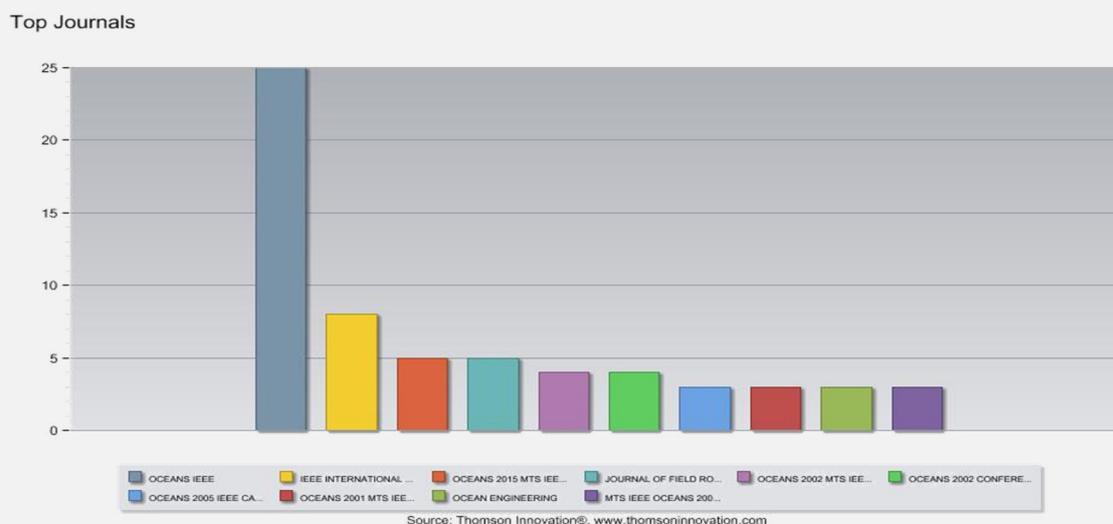


Figura 4. Principais revistas científicas com maior número de artigos

Conforme a Figura 4 acima, a revista científica com maior número de artigos publicados é a Oceans IEEE, com 25 artigos, seguida da IEEE International Conference on Robotics and Automation (ICRA) com 8 artigos e a Oceans '04 MTS/IEEE Techno-Ocean e a Journal of Field Robotics, ambas com 5 artigos.

Em 2015 foi registrado o maior número de artigos científicos publicados, entre os quais se destacam os seguintes periódicos científicos: Journal of Field Robotics com publicação de artigos científicos em 2008, 2010, 2015 e 2016; Oceans '04 MTS/IEEE Techno-Ocean em 2015 e Oceans IEEE com publicação de artigos científicos no período de 2005 a 2015.

Em relação ao destaque para os autores com maior número de artigos publicados, foram apresentados os seguintes resultados: URA, T. (University of Tokyo) tem 14 artigos científicos publicados; RIDAL, P. e ORTIZ, A. com 8 artigos publicados cada um; TANAKA, T. Com 7 artigos publicados; RIBAS, D.; CARRERAS, M. e ANTICH, J. com 6 artigos publicados cada. O menor número de artigos publicados durante o período da pesquisa foram os seguintes autores, PALOMERAS, N. (University of Michigan System), MALLIOS, A. (Woods Hole Oceanog Inst) e HOVER, F. (Massachusetts Institute of Technology-MIT) com 5 artigos cada.

CONCLUSÕES

A prospecção tecnológica é um método fundamental para o desenvolvimento de projetos, que deve ser utilizado sistematicamente, pois influencia todo o processo de pesquisa e de desenvolvimento, e pode ser efetuado por várias ferramentas disponíveis no mercado.

Através do presente mapeamento de prospecção tecnológica sobre Robótica Autônoma Submarina e seus subtemas, utilizando a ferramenta *Thomson Innovation*, alcançaram-se resultados avançados, que permitem verificar a importância de se utilizar a prospecção, focada na busca por anterioridade.

Além disso, foi possível avaliar e apresentar o estado-da-arte dos estudos e aplicações da Robótica Autônoma Submarina no mundo, identificando os pontos relevantes e as novas tecnologias que podem diminuir o tempo de processo e direcionar os investimentos em projetos de pesquisa e de desenvolvimento.

Esse tipo de estudo enriquece a Ciência da Informação e favorece o estabelecimento de novas linhas de investigação, por esclarecer a importância dos estudos de prospecção tecnológica com um serviço de informação com alto valor agregado.

Por outro lado, o presente estudo demonstra a possibilidade de ampliação da atuação do profissional da informação, para as quais necessita novos conhecimentos, competências, habilidades e atitudes, e confirma como o uso efetivo do conhecimento, registrado em artigos, patentes, relatórios e outros documentos, podem nortear a pesquisa, o desenvolvimento e a inovação tecnológica, para o bem-estar dos povos e dos países.

REFERÊNCIAS BIBLIOGRÁFICAS

Amparo, K.S. et al. (2012). Estudo de caso utilizando mapeamento de prospecção tecnológica como principal ferramenta de busca científica. *Perspectivas em Ciência da Informação*, v.17, n.4, pp.195-209, out./dez. Disponível em <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/1533>.

Borschiver, S. et al. (2014). Roadmap tecnológico Sisal. In: *Proceeding of ISTI/SIMTEC*, v. 2, n.1, pp.111-121. D.O.I.: 10.7198/S2318-3403201400020014.

CANONGIA, C. *Gestão do Conhecimento e a Competitividade. Reflexão*. Centro de Gestão e Estudos Estratégicos (CGEE).Brasília-DF: CGEE, 2002.

CASTELLS, M. *A sociedade em rede*. São Paulo: Editora Paz e Terra, 8. ed., 2005.

COELHO, G. M.; SANTOS, D. M.; SANTOS, M. M.; FELLOWS FILHO, L. Caminhos para o desenvolvimento em prospecção tecnológica: *Technology Roadmapping - um olhar sobre formatos e processos*. *Parcerias Estratégicas*, Brasília, v. 21, p. 199-234, 2005.

LE COADIC, Yves-François. *A ciência da informação*. Briquet de Lemos: Brasília, 2004.

Valentim, M. L. P. (2003). *Prospecção e monitoramento informacional no processo de inteligência competitiva*. Londrina: Infohome. Disponível em <https://periodicos.ufsc.br/index.php/eb/article/viewFile/1518-2924.2004v9nesp1p59/5282>.

Valentim, M L. P. & Molina, L. G. (2004). Prospecção e monitoramento informacional no processo de inteligência competitiva. *Encontros Bibli: R. Eletr. Bibliotecon. Ci. Inf.*, Florianópolis, n. esp., 1º sem. 2004. Disponível em <https://periodicos.ufsc.br/index.php/eb/article/viewFile/1518-2924.2004v9nesp1p59/5282>.

Tarapanoff, K. (Org.). (2001). *Inteligência organizacional e competitiva*. Brasília: UnB.

Thomson Innovation, www.thomsoninnovation.com.



REVISITANDO WILLIAM SHAKESPEARE EM SISTEMA DE ORGANIZAÇÃO DO CONHECIMENTO NA ÁREA DE LITERATURA INGLESA

Márcia Barcelos Gomes¹, Vânia Lisboa da Silveira Guedes², Maria José Veloso da Costa Santos³

¹Universidade Federal do Rio de Janeiro, 0000-0002-2507-5186, marcinhaaaa007@yahoo.com.br

²Universidade Federal do Rio de Janeiro, 0000-0001-5854-5677, vanialisboa@facc.ufrj.br

³Universidade Federal do Rio de Janeiro, 0000-0003-0473-5680, msantos1402@gmail.com

RESUMO William Shakespeare soube como poucos pensar e refletir sobre as questões primordiais da própria existência humana e, por essa razão, suas obras literárias são lidas, relidas e adaptadas em diferentes gêneros. Este estudo, motivado pela importância deste dramaturgo, tem como objetivo desenvolver um protótipo de tesauro sobre Shakespeare e sua obra. Dessa forma, para embasamento, utilizaram-se abordagens teóricas e metodológicas interdisciplinares da análise documentária, terminologia, teoria do conceito e dos sistemas de organização do conhecimento (SOC), sob a perspectiva da Linguística Documentária. Nesse sentido, destaca-se a importância da interdisciplinaridade entre a Linguística e a Biblioteconomia que permite a aproximação de diferentes áreas do saber com vistas a permitir o desenvolvimento de SOCs especializados. Para a composição da amostra, selecionaram-se artigos de periódicos científicos, em Português do Brasil, publicados no período de 2006 a 2015, que tratam de temas relacionados à biografia de Shakespeare e sua obra. Como resultado, é apresentado um protótipo de tesauro sobre o tema em questão e, finalmente, sugere-se dar continuidade a esse empreendimento em pesquisas posteriores.

PALAVRAS-CHAVE Organização do Conhecimento, William Shakespeare, Tesauro, Linguagem Documentária.

ABSTRACT This study develops a system of organization of knowledge about Shakespeare and his work in English literature area. William Shakespeare knew how few to think and reflect on the key issues of human existence and, therefore, his literary works are read, reread and adapted in different genres. Thus, this study aims to contribute to the development of a controlled vocabulary prototype, namely, a thesaurus on the biography of the author and his work. Therefore, for basement, were used: interdisciplinary theoretical and methodological approaches of documentary analysis, classification, terminology, concept of theory and knowledge organization systems, from the perspective of the documentary language. From this perspective, it highlights the importance of interdisciplinarity between linguistics and librarianship that allows the approach of different areas of knowledge in order to develop knowledge organization systems. For sample composition, were selected articles from scientific journals, in Portuguese of Brazil, published in the period 2006-2015, which deal with issues related to Shakespeare's biography and his work. Finally, a controlled vocabulary prototype is presented on the topic at hand and it is suggested to continue this project in subsequent research.

KEYWORDS Knowledge Organization. William Shakespeare. Micro thesaurus. Documentary Linguistics.

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

As formas de tratamento documentário podem ser observadas em diferentes momentos da história da humanidade, tais ações permitem que um indivíduo recupere o documento de seu interesse, evitando o manuseio excessivo de informações irrelevantes e facilitando o acesso à informação relevante disponível. Segundo Mey e Silveira (2009), pode-se observar essa atividade, na Antiguidade, no trabalho de Calimachus, sábio grego e bibliotecário da Biblioteca de Alexandria, que criou o Pinakes, catálogo que continha inscrições documentárias sobre o acervo da biblioteca. Com o advento da imprensa, no século XV, Dias e Naves (2013) ressaltam que a produção e a circulação de livros foram otimizadas, fato que propiciou a disseminação e o acesso ao conhecimento não só entre pessoas das classes dominantes, mas também da população que, em sua maioria, estava à margem dos discursos veiculados nos meios de comunicação disponíveis nessa época. Price (1976, p. 126) chama a atenção para o fato de que, já no início do século XIX, “as resenhas e artigos científicos eram tão numerosos que indivíduo algum poderia lê-los ou pretender assimilá-los completamente”.

Mais recentemente, as tecnologias da informação e comunicação (TIC) possibilitaram que maior volume de informação pudesse ser produzido e disponibilizado para diferentes usuários, em lugares distintos e em pouco tempo. Nesse ambiente, para o acesso à informação, é necessária a representação de documentos, seguindo parâmetros internacionalmente utilizados na área de Organização do Conhecimento (OC), particularmente, no processo de construção de informação documentária, como meio de representação e de organização do conhecimento em sistemas de recuperação da informação.

Neste estudo, propõe-se a elaboração de um protótipo de tesouro, a partir da análise documentária de seis artigos de periódicos, publicados no período de 2006-2015, em língua portuguesa, sobre Shakespeare e sua obra. O estudo se insere na interface entre a Linguística e a CI, especificamente na subárea de pesquisa de OC denominada Linguística Documentária (LTD).

FUNDAMENTAÇÃO TEÓRICA

A OC é a área da Ciência da Informação (CI) que, segundo Dahlberg (2006), estuda a estruturação e o arranjo sistemático dos conceitos de uma determinada área do conhecimento, levando em conta as características semânticas que são inerentes aos mesmos e que possibilitam a criação de Sistemas de Organização do Conhecimento (doravante SOC). Os SOC são instrumentos usados na representação e na recuperação da informação em um sistema de informação. Hjørland (2007) considera que os SOC são ferramentas semânticas e, nesse sentido, o tesouro, sendo uma delas, consiste em uma seleção de conceitos suplementada com informações sobre suas relações semânticas.

Conforme Dalhberg (1978), o conhecimento é construído e divulgado por meio da linguagem, na medida em que novos conhecimentos surgem é preciso que novos termos sejam criados para dar conta deste processo que perdurará enquanto o homem existir. Dessa forma, a linguagem é a habilidade que o homem possui de se comunicar e designar os objetos e fatos que são observados. A autora define ainda dois tipos de linguagens: as linguagens naturais, utilizadas pelas pessoas na realização de suas atividades diárias, e as linguagens artificiais ou linguagens formalizadas, criadas pelo homem, como por exemplo, a linguagem da matemática ou a linguagem química.

Lancaster (2004) apresenta dois tipos de linguagem: a linguagem natural (LN) que é utilizada pelos indivíduos para a comunicação no cotidiano e por autores na produção de textos ou documentos, utilizando os termos adequados para que as ideias possam ser veiculadas em um grupo específico; e a linguagem documentária (LD) que se refere às linguagens artificiais que necessitam de regras explícitas para o seu uso, na indexação e recuperação da informação.

A Linguística Documentária (LTD), um subdomínio da Ciência da Informação, se propõe a estudar “[...] as características da linguagem dos ambientes informacionais que combinam as referências da produção informacional, os objetivos institucionais e os elementos cognitivos e comunicacionais de grupos de usuários.” (Tálamo & Lara, 2006, p.204). As autoras esclarecem ainda que a LTD se caracteriza como uma sucessão de tomadas de decisão, uma vez que “Desde a seleção do que compõe ou não o conjunto a ser analisado, a pertinentização, a construção de ‘conteúdos’, até a difusão das formas de sua disseminação está em jogo um universo de opções.” (Tálamo & Lara, 2006, p.2).

Nesse contexto, Tálamo e Lara (2009) acrescentam que a terminologia utilizada por uma comunidade científica, o uso de jargões e termos técnicos, pode criar espaço para relações de poder, ou seja, grande parte da população estaria excluída dos novos conhecimentos veiculados por um determinado grupo. Para as autoras a (LTD), pode estabelecer uma ponte entre os discursos vinculados por pesquisadores e o público em geral, por meio do uso de linguagem de tratamento da informação, permitindo uma maior democratização da informação na sociedade.

Lancaster (1987) afirma que o tesouro é uma variedade de vocabulário controlado, sendo seu uso recorrente para representar os conteúdos de publicações, por meio da padronização de descritores. Gomes (1990) acrescenta que o tesouro é uma linguagem documentária, uma vez que não advém de um processo evolutivo, necessita de regras explícitas e não permite exceções. Dessa forma, no tesouro, o controle de termos se faz necessário para evitar que não se atribua mais do que um conceito ao termo e vice-versa. Outras características são apontadas pela autora, tais como: é uma ferramenta dinâmica; permite a inserção novos termos; é específico e apresenta termos conectados semântica e logicamente.

O conceito é a representação mental que um indivíduo faz sobre determinado objeto, uma vez que Dahlberg afirma que “[...] todo o enunciado sobre objetos contém um elemento do respectivo conceito.” (1978, p.102). Acrescenta-se ainda a definição de Gomes “[...] um constructo mental que representa um objeto material ou imaterial.” (1990, p.18). A apropriação do conceito na Linguística Documentária, mais especificamente para construção de tesouros, é fundamental para a elaboração de tesouros, segundo Gomes (1990), pois cada significado deve ser representado por conceitos e não palavras.

Para Gomes (1990, p.18), a reunião dos conceitos de uma área do conhecimento forma “[...] um sistema de conceitos ou parte de um sistema, o que significa dizer que os conceitos se relacionam entre si. Portanto, devem ser ordenados sistematicamente”. Ela esclarece que “No âmbito dos tesouros os conceitos são designados por termos.” (Gomes, 1990, p.18). Acrescenta ainda que o termo é a palavra utilizada para representar, em um documento, um significado desejado.

METODOLOGIA

O estudo foi elaborado por meio da leitura de teóricos nas áreas da Literatura, Ciência da Informação e Linguística. Tal tipo de pesquisa é caracterizada por Gerhardt e Silveira (2009) como bibliográfica. Além disso, em diferentes etapas, utilizou-se tanto a abordagem qualitativa quanto a e quantitativa, uma

vez que, a abordagem quantitativa, “tem suas raízes no pensamento positivista lógico, tende a enfatizar o raciocínio dedutivo, as regras da lógica e os atributos mensuráveis da experiência humana” enquanto que, a qualitativa “tende a salientar os aspectos dinâmicos, holísticos e individuais da experiência humana, para apreender a totalidade no contexto daqueles que estão vivenciando o fenômeno”. (Polit, Becker e Hungler, 2004 apud Gerhardt & Silveira, 2009, p. 201).

A abordagem qualitativa se fez necessária para a avaliação do contexto, em diversas fases do estudo, como a definição de conceitos na estrutura do protótipo de tesouro, por exemplo. Para ilustrar a quantitativa, pode-se citar o instrumento de análise e a seleção dos artigos, que compõem a amostra, de acordo com a categorização da base WebQualis, da Coordenação de Aperfeiçoamento de Pessoal de Nível (CAPES), composta por um “conjunto de procedimentos utilizados para estratificação da qualidade da produção intelectual dos programas de pós-graduação”. A base apresenta “uma lista com a classificação dos veículos utilizados pelos programas de pós-graduação para a divulgação da sua produção” (CAPES, 2016).

Para a seleção da amostra na WebQualis, optou-se pelos seguintes critérios: evento de classificação Qualis 2014; área de avaliação Letras/Linguística; e classificação do veículo no estrato A1, sendo esses, títulos de periódicos editados em língua portuguesa. Selecionados os títulos com essas características, buscou-se artigos sobre William Shakespeare e sua obra que apresentassem uma média de 3.000 a 7.000 palavras. Os termos e conceitos selecionados para a construção do protótipo de tesouro foram compilados a partir de leitura e análise desses artigos. O marco cronológico abrangeu o período de 2006 a 2015.

Os periódicos selecionados foram: Cadernos de Tradução do Curso de Pós-Graduação em Estudos de Tradução da Universidade Federal de Santa Catarina (UFSC); Aletria: Revista de Estudos de Literatura do Programa de Pós-Graduação em Estudos Literários da Faculdade de Letras da Universidade Federal de Minas Gerais (UFMG). Na tabela 1, a seguir, constam os títulos de periódicos selecionados e os respectivos artigos para a construção do protótipo de tesouro.

Tabela 1- Periódicos e Artigos Selecionados (2006-2015)

Periódicos	Artigos
Cadernos de Tradução	1)Diniz, Thaís Flores Nogueira (2015) Como os pintores “traduziram” <i>Hamlet</i> . Cadernos de Tradução, Florianópolis, v.35, n.1, p. 86-99.
Aletria	2) Galery, Maria Clara Versiani (2006) Romance de Romeu e Julieta: tradição, memória e cultura popular. Aletria: Revista de Estudos de Literatura, Minas Gerais, v.13, n. Especial, p. 155-164.
Cadernos de Tradução	3) Lohmer, José Eduardo dos Santos; Freitas, Renata Cazarini de (2014) Reconhecer e traduzir traços de Sêneca em Shakespeare. Cadernos de Tradução, Florianópolis, v.1, n.33, p. 97-118.
Cadernos de Tradução	4) Martins, Márcia do Amaral Peixoto. Reescritas de peças de Shakespeare para o público jovem: a série Mangá Shakespeare (2014) Cadernos de Tradução, Florianópolis, v.2, n.34, p. 61-84.
Aletria	5) Ramalho, Erick. Shakespeare e o drama satírico (2009) Aletria: Revista de Estudos de Literatura, Minas Gerais, v.19, n. Especial, p. 109-123.
Cadernos de Tradução	6) Salem, Robert (2014) Os Simpsons em Hamlet: um petisco de Shakespeare. Cadernos de Tradução, Florianópolis, v.1, n.33, p. 261-285.

Referência: as autoras (2016)

Para a elaboração do protótipo de tesauro, realizou-se um levantamento dos termos que abordam temas relativos às peças e aos poemas de Shakespeare. Gomes (1990, p.30) aponta que esse procedimento “Oferece [...] garantia literária para a formação das classes, facilitando, pois, a estruturação dos termos em etapa posterior.” A estruturação dos conceitos em tesauros permitiu mapear as relações semânticas entre os termos, incluindo relações hierárquicas e não hierárquicas, bem como relações associativas e de equivalência. Nesse sentido, as Diretrizes da UNESCO (1993) estabelecem três classes de relações para a elaboração de tesauros monolíngues. São elas:

Relações de equivalência: USE (precede o termo preferido) e UP (precede o termo não preferido); b) Relação hierárquica: TGM (precede o maior termo genérico de uma hierarquia); TG (precede o termo superordenado); TGP (precede o termo superordenado na relação todo - parte); TE (precede o termo subordinado) e TEP (precede o termo subordinado na relação todo- parte); c) Relação associativa: TR (precede o termo relacionado na relação associativa).

Segundo as Diretrizes da UNESCO (1993, p. 50) a relação associativa “cobre as relações entre pares de termos que não são membros de um conjunto de equivalência nem podem ser organizados em uma hierarquia onde um termo se subordina a outro”, no entanto, são “termos mentalmente associados de tal maneira que a conexão entre eles deve ser feita explicitamente no tesauro [...]”.

De acordo com o tutorial para a Elaboração de Tesauro Documentário, a relação entre conceitos pode ser de dois tipos: lógica, reconhecida pelas abreviaturas TG/TE; e ontológica reconhecida pelas abreviaturas TGP/TEP. A relação lógica se caracteriza por ser hierárquica, incluindo termos específicos (TE) e termos genéricos (TG), sendo que a relação hierárquica se subdivide em dois tipos: subordinação (formação de cadeia) e coordenação (formação de renque). Gomes (2004) classifica cadeia como uma série vertical de conceitos, onde o conceito subordinado é chamado de termo específico e o superordenado é chamado de termo genérico. A autora descreve o renque como conceitos coordenados, que formam uma série horizontal de conceitos. A relação ontológica é necessária para a análise de um conceito, inclui as relações partitivas (todo/parte) e associativas (ex. produto/processo). (Gomes, 2004).

Quanto à estruturação o tesauro, podem-se observar diversas formas de se demonstrar o relacionamento entre os termos, mas Gomes (1990) estabelece a ordem alfabética como a mais usual, incluindo também nessa ordem as remissivas. Como exemplo cita-se uma remissiva tirada do protótipo apresentado.

CISNE DE AVON (sinônimo)

USE WILLIAM SHAKESPEARE (termo)

Para a definição de conceitos, controle de fenômenos lexicográficos e mapeamento de relações lógicas e ontológicas, utilizaram-se nesse estudo, símbolos adotados no Tesauro sobre Literatura do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) editado em 1985, citados a seguir: **NE:** Nota explicativa. Definição do termo; **UP:** Usado Para. Precede o termo não preferido; **USE:** Use. Precede o termo preferido; **TG:** Termo Genérico. Precede o termo hierarquicamente superior; **TE:** Termo Específico. Precede o termo hierarquicamente inferior; **TGP:** Termo Genérico Partitivo. Precede o termo partitivo superior; **TEP:** Termo Específico Partitivo. Precede o termo partitivo inferior; **TR:** Termo Relacionado. Precede o termo relacionado ou associado (TA).

Além dessas relações adotadas no Tesauro sobre Literatura do IBICT, acrescentou-se a representação de relações lógicas baseada nas Diretrizes da UNESCO (1993), que indicam o uso do símbolo **TGM**, Termo Genérico Maior, para a classe mais ampla à qual pertence o conceito, em uma cadeia hierárquica.

Optou-se pela ordem alfabética para a organização do presente protótipo de tesouro, uma vez que é a forma mais tradicional de apresentação de tesouros, sendo a praticidade uma de suas maiores qualidades.

RESULTADOS

Na construção do protótipo de tesouro sobre Shakespeare e sua obra, foi estabelecido na metodologia que seriam retirados de cada artigo selecionado, entre 05 e 10 termos. A tabela 2, a seguir mostra os termos compilados a partir dos artigos 1 a 6 selecionados.

Tabela 2 – Termos selecionados

1	Peça. Shakespeare. Pintura. Cena. Hamlet. Drama. Tragédia. Literatura Inglesa. (8)
2	Cânone. Memória Coletiva. Romeu e Julieta. Cultura. Cultura popular. (9) Shakespeare. Cordel. Romance. Folheto
3	Shakespeare. Adaptação. Mangá. História em quadrinhos. Literatura juvenil. Literatura infantil. Hamlet. Tragédia. Drama. (9)
4	Teatro latino. Tragédia. Sêneca. William Shakespeare. Comédia. Teatro. Drama. Literatura. (8)
5	Shakespeare. Adaptação. Mangá. História em quadrinhos. Literatura juvenil. Literatura infantil. Hamlet. Tragédia. Drama. (9)
6	Shakespeare. Teatro. Drama. Tragédia. Sátira. Comédia. Peça. Tragicomédia. Gênero textual. Paródia. (10)

Referência: As autoras (2016)

Verificou-se que foram compilados dos seis artigos analisados um total de 53 termos que compuseram o protótipo de tesouro, do qual apresenta-se duas entradas, com algumas relações:

CISNE DE AVON USE WILLIAM SHAKESPEARE

WILLIAM SHAKESPEARE

NE: Apelidado de Bardo de Avon ou Cisne de Avon, batizado em 26 de Abril de 1564, em Stratford-upon-Avon, condado de Warwickshire, Inglaterra e morreu em 23 de Abril de 1616, em Stratford-upon-Avon. Poeta inglês, dramaturgo e ator, muitas vezes chamado de o poeta nacional Inglês, sendo considerado por muitos como o maior dramaturgo de todos os tempos. (ENCICLOPAEDIA BRITANNICA, 2016c).

UP BARDO DE AVON

UP CISNE DE AVON

TGM LITERATURA

TG LITARATURA INGLESA

TG DRAMA ELISABETANO

TG TRAGÉDIA

TE HAMLET (PEÇA)

TE ROMEU E JULITA (PEÇA)

TR TEATRO

TR TEATRO LATINO

TR TRAGICOMÉDIA

TR SÊNECA

CONCLUSÕES

O objetivo geral do presente estudo foi elaborar um protótipo de tesauro, tendo como objeto de análise artigos, escritos em Português do Brasil, que abordassem os seguintes assuntos: Shakespeare e sua obra. Foram incluídos também artigos que tratam de adaptações das obras do dramaturgo para diversos gêneros textuais, desde pinturas até mangás.

Para auxiliar no desenvolvimento da pesquisa, utilizaram-se como base fundamental as discussões sobre as contribuições da Linguística, por meio da Linguística Documentária, para a formação de Sistemas de Organização do Conhecimento, uma vez que a linguagem permeia diferentes campos do saber.

Os estudos de Lancaster (1987; 2004) foram utilizados para o estabelecimento de critérios, desde a forma como o profissional bibliotecário deve trabalhar com os documentos, qual tipo de leitura técnica deve ser adotado, bem como para a definição da quantidade de conceitos/termos analisados em cada artigo de periódico selecionado.

Além de Lancaster, outros autores que produzem frequentemente importantes trabalhos sobre vocabulários controlados e tesouros, sob a perspectiva da Organização do Conhecimento na CI, foram utilizados para nortear o presente estudo, com destaque para a Professoras Hagar Espanha Gomes e Maria Luiza de Almeida Campos. Seus manuais, livros e, até mesmo, tutoriais na web são ferramentas riquíssimas para auxiliar o bibliotecário- indexador a entender como organizar os conceitos e termos do campo semântico de uma área do conhecimento e, assim, representar os documentos de forma a disponibilizar os mesmos com qualidade, sempre pensando em estratégias para criar sistemas de recuperação da informação que tenham uma interface mais clara e próxima das necessidades do usuário. Para Gomes (1990) o uso de um vocabulário controlado, como o tesauro, é interessante, uma vez que sua estrutura permite a atualização e a inserção de novos termos de acordo com a dinâmica do conhecimento.

Trabalhar com vocabulários controlados na área da Literatura é um grande desafio, uma vez que grande parte dos dicionários, glossários e textos de referência não são atualizados, com frequência. Durante a elaboração do presente estudo sentiu-se dificuldade em encontrar definições para alguns termos, tais como: desenho animado, história em quadrinhos e mangá. Tal situação pode estar relacionada a certo preconceito linguístico com estes tipos de gêneros, mas “[...] não concebamos os gêneros como modelos estanques, nem como estruturas rígidas, mas como formas culturais e cognitivas de ação social corporificadas de modo particular na linguagem, [...] ver os gêneros como entidades dinâmicas.” (Marcuschi, 2008. p. 156). Ou seja, a linguagem é dinâmica e deve ser entendida como a ferramenta sistêmica por meio da qual o homem se expressa, se comunica e entende o mundo ao redor e o representa.

Pensando neste dinamismo e vivacidade do conhecimento e da linguagem, o bibliotecário que trabalha com a análise e indexação de documentos deve estar atento às demandas da comunidade em que atua, bem como procurar estabelecer critérios para desenvolver vocabulários controlados e descritores de áreas especializadas do conhecimento que ainda não foram muito exploradas, como a da Literatura Inglesa, por exemplo.

Outro grande desafio, que a presente pesquisa pretende instigar é a construção de um vocabulário controlado mais completo de conceitos e termos que pertencem ao campo semântico da obra de Shakespeare em sua totalidade, em Língua Portuguesa, apesar de difícil execução, uma vez que a linguagem do Bardo é riquíssima e cheia de detalhes. Acredita-se que este tipo de ferramenta seria

interessante para estudantes, pesquisadores e curiosos no assunto, uma vez que permitiria a recuperação mais eficiente e ágil de uma obra de interesse para esses usuários e de grande relevância para a área de Literatura.

Sendo assim, um objetivo futuro para esta pesquisa seria estabelecer uma parceria com profissionais (bibliotecários, docentes e discentes) que atuam neste campo tão rico e surpreendente, que é a Literatura, uma vez que o trabalho holístico, com diferentes visões e experiências sobre o uso e o conhecimento da linguagem e da Literatura, permitirá a construção de um vocabulário controlado com maior qualidade e eficiência.

REFERÊNCIAS BIBLIOGRÁFICAS

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (2016) Portal de Periódicos. Recuperado em 6 junho, 2016, de <http://www.periodicos.capes.gov.br/>

Dahlberg, I. (1978) Teoria do conceito. *Ciência da Informação*, v.7, n.2, p. 101-107,1978. Recuperado em 4 janeiro, 2016, de <http://revista.ibict.br/index.php/ciinf/article/view/1680/1286>

Dias, E. W. & Naves, M. M. L. (2013) *Análise de assunto: teoria e prática*. Brasília, DF: Briquet de Lemos.

Enciclopaedia Britannica. English literature. Recuperado em 4 junho,2016, de <http://global.britannica.com/art/English-literature> >. Acesso 04 jun. 2016.

Gomes, H. E. (1990) *Manual de elaboração de tesauros monolíngues*. Brasília: CNPq/PNBU.

Hjørland, B. (2007) *Knowledge Organization Systems*. Recuperado em 27 junho, 2017, de http://www.isko.org/cyclo/knowledge_organization

Instituto Brasileiro de Informação em Ciência e Tecnologia (1985) *Tesauros sobre literatura*. Brasília, D.F.

Lancaster, F. W. (1987) *Construção e uso de tesauro: curso condensado*. Brasília: IBICT.

Lancaster, F. W. (2004) *Indexação e resumos: teoria e prática*. Brasília, D.F.: Briquet de Lemos.

Marcuschi, L.A. (2008) *Produção de gêneros e compreensão*. São Paulo: Parábola.

Mey, E. S.A.& Silveira, N. C. (2009) *Catálogo no plural*. Brasília, D.F.: Briquet de Lemos.

Price, D. S. (1976) *A ciência desde a Babilônia*. Belo Horizonte: Itatiaia; São Paulo: EDUSP.

Silveira, D. T. & Córdova, F. P. (2009) A pesquisa científica: tipos de pesquisa quanto à abordagem. In: GERHARDT, T. E.; SILVEIRA, D. T. (Org.). *Métodos de pesquisa*. Porto Alegre: Editora UFRGS, p. 31-33.

Tálamo, M. F. G. M. & Lara, M. L. G. O campo da Linguística Documentária. *TRANSINFORMAÇÃO*, v. 18, n. 3, p. 203-211, set./dez., 2006. Recuperado em 12 novembro, 2015, de <http://periodicos.puc-campinas.edu.br/seer/index.php/transinfo/article/view/668>

Unesco (1993) Diretrizes para o estabelecimento e desenvolvimento de tesouros monolíngües. [2ª ed. rev. por Derek Austin e Peter Dale]; Trad. de Bianca Amaro de Melo; rev. de Lígia Maria Café de Miranda. Brasília: IBICT; SENAI.



CATEGORIZAÇÕES DE SOC: UMA INVESTIGAÇÃO LITERÁRIA

Márcio Bezerra Da Silva¹, Zeny Duarte de Miranda²

¹Universidade de Brasília (UnB) – Brasil, 0000-0002-0052-7174, marciobdsilva@unb.br

²Universidade Federal da Bahia (UFBA) – Brasil, 0000-0003-0365-6905, zenydu@gmail.com

RESUMO Estudo que objetiva identificar propostas de categorização dos sistemas de organização do conhecimento na literatura. Resulta, a partir de uma pesquisa bibliográfica, de cunho qualitativo, que os sistemas de organização do conhecimento evoluíram de esquemas pré-coordenados para pós-coordenados, especialmente a partir da Internet, além da diversificação de rótulos de categorização dos sistemas de organização do conhecimento, sendo uns com enfoque nos modelos tradicionais, outros nos arranjos digitais, enquanto outros mais naqueles que nasceram bibliográficos, mas que apresentam um arcabouço teórico que os permitem à aplicação nos ambientes digitais, como os sistemas facetados, além das abordagens tradicionalmente não consideradas pela maioria dos autores, atuando como complementos à construção e funcionamento dos sistemas de organização do conhecimento. Conclui-se que as estruturas pré-coordenadas representam os sistemas de organização do conhecimento mais tradicionais, indo ao encontro de uma organização bibliográfica, enquanto que os estudiosos convergem que a ontologia é o tipo mais contemporâneo entre os sistemas de organização do conhecimento elencados, bem como a classificação facetada representa a ideia de transição de esquemas antigos para modelos mais modernos, especialmente por criarem estratégias de busca variadas sobre a mesma informação e navegação multidimensional, características comuns em *websites* recentes.

PALAVRAS-CHAVE *organização do conhecimento, sistemas de organização do conhecimento, pré-coordenados, pós-coordenados, sistemas facetados.*

ABSTRACT Study that aims to identify proposals for categorization of knowledge organization systems in the literature. Based on a qualitative bibliographical research, the study results that the systems of knowledge organization evolved from pre-coordinated to post-coordinated schemes, especially from the Internet, as well as the diversification of categorization labels of knowledge organization systems, Some with a focus on traditional models, others with digital arrays, and others with bibliographic origins, but with a theoretical framework that allows them to be applied in digital environments, such as faceted systems, in addition to approaches traditionally considered by most Of the authors, acting as complements to the construction and functioning of knowledge organization systems. Concludes that the pre-coordinated structures represent the most traditional systems of knowledge organization, meeting a bibliographic organization, while scholars converge that ontology is the most contemporary type among the systems of knowledge organization listed, as well As the faceted classification represents the idea of transition from old schemes to more modern models, especially by creating varied search strategies on the same information and multidimensional navigation, characteristics common in recent websites.

KEYWORDS *knowledge organization, knowledge organization systems, pre-coordinated, post-coordinated, faceted systems.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Em 1988, na cidade de Pittsburgh (Pensilvânia – EUA), foi cunhada a expressão *Knowledge Organization System* (KOS), conhecida em português pela tradução sistema de organização do conhecimento (SOC) e conceituada da seguinte forma:

[...] representações de domínios do conhecimento que delimitam o significado de termos no contexto desses domínios, estabelecem relações conceituais que auxiliam a posicionar um conceito no sistema conceitual e são utilizadas como instrumentos de organização e recuperação da informação. (Bräscher e Carlan 2010: 149)

Como os objetivos propostos pelos SOC “[...] visam à organização da produção intelectual humana” (Vignoli, Souto e Cervantes, 2013: 61), permitindo “[...] apoiar a recuperação de itens relevantes de uma coleção da biblioteca” (Hodge, 2000: 1), Lambe (2011) os enxerga como importantes instrumentos no desenvolvimento e progresso da ciência, pois normalizam a linguagem, o que possibilita a coordenação e a construção de um sistema em torno da linguagem compartilhada e das entidades descritas por essa linguagem; identificam conexões ou relações entre diferentes áreas do conhecimento; e sobrepõem estruturas salientes e úteis a um domínio de conhecimento difuso.

Na verdade, SOC refere-se a uma denominação recente para as linguagens documentárias (LD), assim como defende Carlan (2010: 28-29), pois “[...] agregam elementos incorporados nas inovações tecnológicas da era digital”. Com base em Fujita, Rubi e Boccato (2009), as LD objetivam à organização e à disseminação de informação, funcionando como um controle terminológico que afetará no desempenho da recuperação e filtragem de informações.

A expressão SOC, neste caso KOS, foi apresentada na primeira reunião da equipe de trabalho intitulada *Networked Knowledge Organization Systems* (NKOS). De acordo com o NKOS (2016), trata-se de um grupo dedicado a discussões sobre modelagem funcional e de dados para permitir que o conhecimento seja organizado de forma sistêmica, os chamados sistemas de organização do conhecimento (SOC), atuando como serviços de informação interativos em rede para apoiar a descrição e recuperação de diversos recursos de informação, através da Internet.

A *International Society for Knowledge Organization* (ISKO) é outro espaço de discussão sobre organização do conhecimento (OC) e seus sistemas, físicos e digitais, promovendo debates acadêmicos nos âmbitos internacional, Espanha-Portugal e Brasil. Neste sítio de discussão, a OC pode ser vista como uma “[...] ciência que estrutura e organiza sistematicamente unidades do conhecimento (conceitos) segundo seus elementos de conhecimento (características) inerentes [...]” (Dahlberg, 1993: 211) com vistas “[...] à construção de modelos de mundo que se constituem em abstrações da realidade (Bräscher e Café, 2008: 6).

A ISKO, devido à forte penetração das recentes TIC, abarca profissionais além de cientistas da informação e bibliotecários, tais como linguísticos, filósofos e cientistas da computação, inclusive programadores, analistas de sistemas e informáticos com enfoque na saúde, objetivando contribuir no avanço do trabalho conceitual na OC em todos os tipos de formas (ISKO, 2016).

Entre as temáticas discutidas na ISKO estão as perspectivas epistemológicas em OC; formação e atuação profissional em OC; perspectivas de investigação em representação e OC: atualidade e tendências; e desafios à organização e acesso ao património cultural. Os citados subtemas, por exemplo,

configuram o III congresso ISKO Espanha-Portugal e XIII congresso ISKO Espanha, a ser realizado no mês de novembro, na cidade de Coimbra (Portugal), do corrente ano (ISKO España-Portugal, 2017).

Direcionando a presente pesquisa entre os subtemas dos referidos eventos, o vislumbre em identificar propostas de categorização dos SOC, na literatura da ciência da informação (CI), vai ao encontro do painel, nos supracitados eventos, intitulado “perspectivas de investigação em representação e OC: atualidade e tendências”, pois os esquemas de organização transitam de tradicionais aos digitais, a partir de sua época de cunho e perspectivas informáticas vigentes. Também deve-se mencionar àqueles SOC que nasceram bibliográficos, isto é, com fins de organizar acervos físicos, mas que apresentam um arcabouço teórico, e inclusive visionário, de aplicação nos ambientes digitais, comumente encontrado em *websites*, como é o caso dos sistemas facetados (esquemas analítico-sintéticos/analítico-facetados, como um modelo categorial proposto pelo indiano, matemático e bibliotecário Shiyali Ramamrita Ranganathan (1897-1972) e que representa o tema de investigação de uma pesquisa doutoral a partir da teoria da classificação facetada (TCF) em bibliotecas digitais, estimulada pela perspectiva da citada teoria oferecer caminhos de navegação e busca variados sobre o mesmo assunto/objeto, cabendo ao usuário escolher o trajeto a ser percorrido. Trata-se de um ambiente multidimensional (poli hierárquico), onde a cognição do usuário será o real guia no uso de determinado ambiente, seja ele físico ou digital.

Com base na diversificação tipológica dos SOC, diante do acima panorama apresentado, vale destacar que:

Embora alguns SOC sejam mais flexíveis e de fácil adaptação, sempre vai existir algum tipo de visão parcial ou tendenciosa. Esta visão parcial na estrutura dos SOC é natural, pois reflete os interesses e as considerações de uma coleção, dos usuários e de suas prioridades. (Carlan, 2010: 31)

A partir de um levantamento teórico, de cunho qualitativo, percebe-se que os SOC evoluíram de esquemas pré-coordenados para pós-coordenados, especialmente a partir do advento da Internet, bem como os autores estudados possuem seus próprios rótulos de categorização dos SOC, uns focando em modelos tradicionais, outros nos arranjos de nascedouro digital, enquanto outros mais intercedem entre os mais conhecidos e os menos tradicionais, além de citar abordagens tradicionalmente não consideradas pela maioria dos autores, atuando como complementos à construção e funcionamento dos SOC.

METODOLOGIA

Na realização de uma pesquisa científica, a metodologia desempenha papel fundamental, pois ela apresenta os caminhos a serem usados em uma investigação. A partir do problema definido, ao qual “[...] tem por objetivo estabelecer uma série de compreensões no sentido de descobrir respostas para as indagações e questões que existem em todos os ramos do conhecimento humano [...]” (Oliveira, 2001: 117), esta investigação formaliza-se pela pesquisa bibliográfica e por uma abordagem de coleta de dados qualitativa, realizada no período de 10 de janeiro até 30 de março do corrente ano, segundo uma juntada literária constituída por livros, artigos científicos e *websites* especializados.

Partindo da compreensão de que a pesquisa bibliográfica “[...] tem por finalidade conhecer as diferentes formas de contribuição científica que se realizam sobre determinado assunto ou fenômeno” (Oliveira, 2001: 119), inclusive de que se trata de uma técnica que “[...] acaba se transformando em rotina para os pesquisadores e profissionais que necessitam de constante atualização” (Oliveira, 2001: 119), como

objetiva a presente pesquisa, realizou-se a identificação das propostas de classificação dos SOC segundo os mais variados autores (nacionais e internacionais), mostrando suas visões peculiares sobre os esquemas de organização na rotulagem de cada categorização aqui elencada. Ao adotar o apresentado percurso metodológico, acredita-se na possibilidade de identificação dos diferentes nomes de categorias, inclusive para o mesmo tipo de SOC, assim como distanciamento e aproximação entre os sistemas.

RESULTADOS

Cada SOC possui peculiaridades, o que permite diferentes categorizações quanto as técnicas de construção, bem como diversas nomenclaturas conforme a perspectiva do estudioso. Souza (2012), por exemplo, numa ótica bibliográfica (documentária), apresenta a seguinte proposta de categorização: **enumerativos**, formados por categorias que dividem o universo do conhecimento e são apresentadas em números, como são os casos da classificação decimal de Dewey (CDD) e da classificação decimal universal (CDU); e **facetados**, também chamados de classificação analítico-sintética, em que os assuntos são decompostos (síntese) conforme facetas que organizam categorias diversas que apresentam características comuns.

Ainda na visão bibliográfica, Ranganathan rotula os sistemas de classificação, de sua época, como **esquemas descritivos** ou **arranjos dinâmicos** (Ranganathan, 1967). Enquanto que a proposta **descritiva** representa “[...] o estado atual do conhecimento e não tem mecanismos que permitam atender às mudanças advindas das diversas áreas do conhecimento” (Campos, 2001: 31), isto é, refletem o conhecimento naquele específico momento tempo-histórico, a versão **dinâmica**, em uma perspectiva filosófica, associa-se a um “[...] Universo de Conhecimento como uma espiral que está em movimento contínuo agregando novos conceitos, trazendo para o Universo de Trabalho da Classificação uma perspectiva dinâmica” (Campos, 2001: 33), ou seja, permitindo que: sejam acomodados novos assuntos, analisados em facetas e sem prejudicar a sequência útil de funcionamento do sistema; e a hospitalidade de assuntos. Ainda sobre os **arranjos dinâmicos**, vale salientar que a *colon classification* (CC), idealizada pelo próprio Ranganathan, foi o primeiro sistema desenvolvido na perspectiva dinâmica, deflagrando a facetagem como um novo tipo de classificação bibliográfica. De acordo com Ranganathan (1967), os SOC **descritivos** podem ser **enumerativo**, ou seja, que existe apenas uma tabela que enumera os assuntos básicos a partir de uma notação monolítica, como a *Library of Congress classification* (LCC); **quase enumerativo**, isto é, sistema que amplia a tabela do esquema anterior, não sendo mais única e permitindo a enumeração dos assuntos básicos e dos compostos, além de tabelas de isolados, como a CDD; e **quase facetado**, em outras palavras, um esquema menos monolítico e que contém todos os elementos do arranjo quase enumerativo, além de tabelas de isolados especiais e de orientações para a elaboração da notação, como a CDU. No que se refere aos SOC **dinâmicos**, para Ranganathan (1967), são classificados como **rigidamente facetado**, ou seja, as facetas e suas sequências são permanentes para todos os assuntos, neste caso, em cada classe básica; ou **livremente facetado**, isto é, não há influência das tabelas de classificação existentes, pois qualquer faceta que ocorra em um assunto composto é descoberta pela análise da faceta daquele assunto, ao qual cada assunto composto determina suas próprias facetas, seu número de classes e a própria sequência de facetas.

Em um sentido mais ampliado da ótica bibliográfica indicada por Souza (2012) e mais próximo ao modelo ranganathiano, Souza (2007) sugere uma proposta de categorização de SOC não totalmente enumerativo ou completamente facetado: **enumerativos** (método dedutivo), como a CDD;

semienumerativos, como a CDU; e **facetados** (método indutivo), como a CC e a classificação facetada do *classification research group* (CRG).

Na visão de Broughton, Hansson, Hjørland e López-Huertas (2005), a tipologia dos esquemas de organização é mais abrangente, trazendo abordagens normalmente não citadas nos estudos em SOC, a saber: **tradicionais sistemas de classificação**, usados em bibliotecas e bases de dados na organização de documentos como os esquemas da CDD e da CDU; **abordagem analítico-facetada** do Ranganathan e posteriormente desenvolvida pelo CRG e pelos editores da segunda edição da classificação bibliográfica de Bliss (CBB); a **tradição da recuperação de informação**, encarado como um campo fundado nos anos de 1950 e que vem se desenvolvendo, especialmente a partir de pesquisas na Internet junto aos buscadores (motores de busca); **visualizações (organização) orientadas (de forma amigável) aos usuários**, ou seja, uma linha de estudo mais focada na recuperação de informação do que na OC propriamente dita, atuando como uma espécie de estudo de usuários orientada a participação dos mesmos, como ocorre na folksonomia, neologismo criado por Thomas Vander Wal, a partir da junção das palavras *folk* (gente/povo/pessoas) com *taxonomy* (taxonomia/classificação), e que se constitui no “[...] modo de indexação cooperativa que adota palavras-chave para representar a informação de maneira espontânea e livre (Moreiro González, 2011: 44-45); **abordagens bibliométricas**, aos quais representam uma proposta que investiga os documentos e padrões de citação entre documentos e que se apresentam como tentativas de combinar a bibliometria, enquanto “[...] o estudo dos aspectos quantitativos da produção, disseminação e uso da informação registrada” (Macias-Chapula, 1998: 134), com OC e recuperação de informação no objetivo de apresentar o conhecimento que foi levantado/analísado/apurado em uma área de domínio; **abordagem analítica de domínio** enquanto uma linha em que o conhecimento humano é substituído por reivindicações do conhecimento (presentes nas obras/materiais), pois a construção, avaliação e utilização de um SOC deve basear-se numa reflexão sobre os pontos de vista identificados na análise do domínio organizado; e **outras abordagens**, entre as quais estão a semiótica, a crítico-hermenêutica, analítico-discursiva, gênero, linguagens de marcação, representação de documentos, tipologia documental e descrição, arquitetura de documentos, sistemas de arquivo (baseados no princípio da proveniência) etc.

Com base em Hodge (2000), Zeng (2008) apresenta uma proposta de evolução dos SOC (figura 1), iniciando pelos sistemas considerados mais simples, como as listas de termos, entre os quais cita-se os glossários, até os esquemas que se desenham em listas de relacionamento como as ontologias.

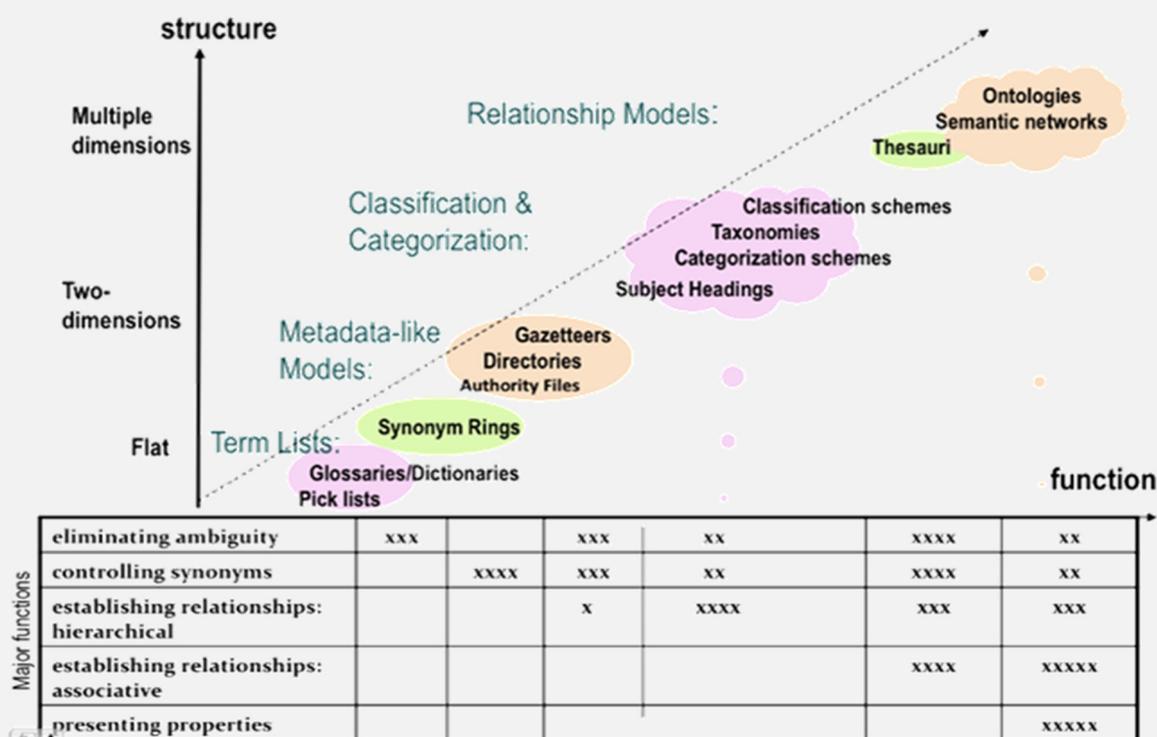


Figura 1. Evolução tipológica dos SOC.

Referência: NKOS, 2000.

Destrinchando a proposta da figura 1, Hodge (2000, p. 5) categoriza os SOC com um vislumbre de organização da informação (OI) em bibliotecas digitais, “[...] baseados em características tais como a estrutura e complexidade, as relações entre os termos e função histórica”. A proposta formaliza-se em três categorias básicas: listas de termos; classificações e categorias; e listas de relacionamento. A **lista de termos** é representada pelos arquivos de autoridade, entre nomes, entidades ou um domínio particular; dicionários são listas organizadas alfabeticamente e que contêm definições, além de apresentarem informações sobre a origem de uma palavra, variações e múltiplos significados; e os *gazetteers*, que são listas confeccionadas como uma espécie de dicionário geográfico (nomes de lugares), apresentado como livros ou como índices para atlas. As **classificações e categorias** enfatizam a criação de conjuntos de assuntos a partir de cabeçalhos de assunto, esquemas de classificação, como as taxonomias, e esquemas de categorização, que por sua vez possuem agrupamentos além do modelo hierárquico da taxonomia. As **listas de relacionamento** enfatizam as ligações entre conceitos, como os tesouros (relações hierárquicas, de equivalência ou associativas); como as redes semânticas, ou seja, estruturas não hierárquicas, mas como uma rede ou uma teia (pensada como nós), constituídas pela linguagem natural e que podem ter relações do tipo todo-parte, causa-efeito ou pai-filho; e como a ontologia enquanto um modelo que apresenta o rótulo mais recente entre os SOC mencionados, formalizado pela criação de relacionamentos e adoção de regras e axiomas em uma área específica do conhecimento.

Vickery (2008) expõe uma indicação contemporânea das categorizações supracitadas, defendendo que os SOC podem ser classificados como estrutura pré-coordenada (estática), estrutura pós-coordenada, a Era da Internet e a Era da *web* semântica. No caso **pré-coordenação**, propõe-se que as entradas de assunto sejam termos únicos ou combinados e os arranjos sejam ordenados de forma alfabética, enumerativa (hierarquia) ou facetada, via categorias como entidade, parte, atributo, operação, lugar e

tempo. Quanto à **estrutura pós-coordenada**, os elementos do SOC passaram a ser manipuláveis a partir de combinações dinâmicas, inclusive fazendo uso dos operadores AND (e) e OR (ou), como os tesouros, por exemplo. Na **Era da Internet**, as buscas *online* ocorrem a partir de uma classificação hierárquica, inclusive facetada, permitindo a navegação pelos resultados, conforme busca por referências de citação (*links de uniform resource locator – URL*) de um item da *web* para outro ou a partir de pesquisa por um índice criado de um buscador ao qual forma um conjunto de palavras extraídas de textos; e a **Era da web semântica**, proposta em que os SOC, conhecidos como ontologias, passam a ser usados por agentes de *software* "inteligentes" na busca por informações ao levar em consideração a relação entre dois termos criado de forma explícita e segundo inferência, pelos quais formalizará uma dedução ao fato analisado.

Miranda (2005) sintetiza as categorias anteriormente mencionadas em quatro tipos, perpassando pelos arranjos mais simples, como os usados para organizar livros, até os esquemas mais complexos, exemplificados pelas ontologias. Os **sistemas bibliográficos** surgiram com o objetivo de organizar os livros nas estantes, a partir do uso de notações, como a CDD e a CC. No caso dos **sistemas alfabéticos**, são arranjos que apresentam uma ordenação sequencial assim como os dicionários, os glossários e as listas de cabeçalho de assunto. Os **sistemas alfabéticos sistemáticos** baseiam-se em conceitos que apresentam relacionamentos entre seus termos, tais como hierárquico, equivalente ou associativo, presentes em tesouros, por exemplo. Por fim, os **sistemas inferenciais** são esquemas abalizados em regras lógicas de raciocínio, aos quais permitem a organização automática de domínios do conhecimento por meio de relacionamentos complexos entre objetos, que por sua vez incluem regras e axiomas para configurar redes semânticas.

Conforme as sugestões de categorização dos SOC elencadas, construídas em modelos que seguem as necessidades dos usuários e propostas de disseminação de informação do seu tempo, inclusive influenciadas pelas recentes TIC, o quadro 1 tenta ilustrar uma relação de aproximação e distanciamento entre os autores frente a diversidade de categorizações, sintetizadas em pré-coordenadas, pós-coordenadas e outras abordagens.

Quadro 1. Propostas de categorização dos SOC

TIPOS DE SOC (abordagens)	AUTORES						
	Souza (2012)	Ranganathan (1967)	Souza (2007)	Broughton <i>et al.</i> (2005) e Hjørland (2008)	Hodge (2000)	Vickery (2008)	Miranda (2005)
Pré-coordenados	X	X	X	X	X	X	X
Listas de termos					X		X
Classificações e categorias	X		X	X	X	X	X
Enumerativos	X	X	X	X	X	X	X
Semienumerativos		X	X	X	X	X	X
Facetados	X	X	X	X	X	X	X
Pós-coordenados				X	X	X	X
Era da Internet				X	X	X	X
Listas de relacionamento				X	X	X	X
Era da <i>web</i> semântica				X	X	X	X
Outras abordagens				X	X	X	
Análise de domínio				X			
Bibliometria				X			
Recuperação da Informação				X			
Orientação ao usuário				X			

Referência: Dos autores, 2017.

Em suma, independente da proposta de categorização a ser seguida, Hodge (2000) defende que os SOC são mecanismos de OI que atuam como o coração de toda biblioteca, museu e arquivo, já que permitem “[...] organizar, padronizar e possibilitar a recuperação da informação e do conhecimento [...]” (Vignoli, Souto e Cervantes, 2013: 62). A assertiva apresentada pode ser vista tanto no prisma do ambiente físico como também transcender ao mundo digital, fazendo com que alguns sistemas, como o modelo pré-coordenado facetado, originalmente criado no uso da biblioteca física, apresente subsídios na OI em *websites*, bibliotecas digitais etc.

CONCLUSÕES

Os ambientes pré-coordenados retratam os SOC revelados pelas listas de termos, como os dicionários, e pelos sistemas de classificação/categorização bibliográfica, exemplificados pelas ordenações enumerativas, como a CDD, pelos esquemas semienumerativos, como a CDU, e pelos arranjos facetados, como a CC.

No caso da expressão pós-coordenada, os SOC são representados, inicialmente, por estruturas taxonômicas, que aos poucos foram evoluindo conforme as relações entre os termos presentes neste esquema, aos quais são montadas (combinadas) segundo as expressões dos usuários digitadas em buscadores *online*, por exemplo. Os tesouros abarcam uma outra exemplificação pós-coordenada ao expandir as relações entre os termos se comparadas as taxonomias, além de serem adotadas como controle terminológico (vocabulário). Ambos os esquemas de classificação supracitados serviram de subsídio ao surgimento de propostas inferenciais, neste caso, deflagradas na Era da *web* semântica, ou seja, a partir de esquemas como as redes semânticas e as ontologias, aos quais buscam explicitar um determinado domínio do conhecimento humano.

A terceira faixa do quadro 1 remete a outras abordagens citadas exclusivamente por Broughton *et al.* (2005) e Hjørland (2008) e que servem de complemento à construção e funcionamento dos SOC, bem como uma intercessão entre os esquemas mais simples e os arranjos mais complexos.

Conclui-se que as estruturas pré-coordenadas representam os SOC mais tradicionais pela maioria dos autores, relacionados a uma organização bibliográfica, enquanto que os estudiosos convergem que a ontologia é o tipo mais contemporâneo entre os SOC, bem como a classificação facetada representa a ideia de transição de esquemas antigos para modelos mais modernos, especialmente por criarem estratégias de busca variadas sobre a mesma informação e navegação multidimensional, características comuns em *websites* recentes.

REFERÊNCIAS BIBLIOGRÁFICAS

Bräscher, M. e Café, L. (2008). Organização da informação ou organização do conhecimento? Em Lara, M. L.G.; Smit, J. (Organização). Em *Temas de pesquisa em Ciência da Informação no Brasil*. São Paulo: Escola de Comunicação e Artes (USP). Recuperado em 12 de fevereiro, 2017, de <http://skat.ihmc.us/rid=1KR7TM7S9-S3HDKP-5STP/BRASCHERCAF%C3%89>.

Fujita, M. S. L., Rubi, M. P. e Boccato, V. R. C. (2009). As diferentes perspectivas teóricas e metodológicas sobre indexação e catalogação de assuntos. Em Fujita, M. S. L. (organizadora). *A indexação de livros: a percepção de catalogadores e usuários de bibliotecas universitárias. Um estudo de observação do contexto sociocognitivo com protocolos verbais*. São Paulo: Cultura Acadêmica. Recuperado em 14 de março, 2017, de <http://static.scielo.org/scielobooks/wcvbc/pdf/boccato-9788579830150.pdf>.

Broughton, V., Hansson, J., Hjørland, B., López-Huertas, M. J. (2005). Knowledge Organization. Em *European Curriculum Reflections on Library and Information Science Education*. Recuperado em 03 de fevereiro, 2017, de <https://www.asis.org/Bulletin/Dec-06/EuropeanLIS.pdf>.

Campo, M. L. (2001). A. *Linguagem Documentária: teorias que fundamentam sua elaboração*. Niterói: EdUFF. Recuperado em 20 de abril, 2017, de <http://www.uff.br/ppgci/editais/linguagem.pdf>.

Carlan, E. (2010) *Sistemas de Organização do Conhecimento: uma reflexão no contexto da Ciência da Informação*. Dissertação de Mestrado, Universidade de Brasília, Brasília, Brasil. Recuperado em 07 de abril, 2017, de <http://eprints.rclis.org/14519/1/Carlan-Eliana-Dissertacao.pdf>.

Carlan, E. e Bräscher, M. (2011). Sistemas de Organização do Conhecimento na visão da Ciência da Informação. Em *Revista Ibero-Americana de Ciência da Informação*, 4(2), 53-73. Recuperado em 12 de fevereiro, 2017, de <http://periodicos.unb.br/index.php/RICI/article/view/6209/5102>

Dahlberg, I. (1993). Knowledge organization: its scope and possibilities. Em *Knowledge Organization*, 20(4), 211-222.

Hjørland, B. (2008). What is Knowledge Organization (KO)? Em *Knowledge Organization* 35(2/3), 86-101. Recuperado em 02 de fevereiro, 2017, de http://www.ergon-verlag.de/downloads/35_2008_2-3.pdf.

Hodge, G. (2000). *Systems of Knowledge Organization for Digital Libraries: beyond traditional authorities files*. Washington (DC), the Council on Library and Information Resources. Recuperado em 11 de março, 2017, de <https://www.clir.org/pubs/reports/pub91/pub91.pdf>

International Society for Knowledge Organization. (2016). *About ISKO*. Recuperado em 02 de abril, 2017, de <http://www.isko.org/about.html>.

International Society for Knowledge Organization Espanha-Portugal. (2017). *III Congresso ISKO Espanha-Portugal / XIII Congresso ISKO Espanha*. Recuperado em 02 de abril, 2017, de <http://sci.uc.pt/eventos/index.php/ISKO/ISKO2017>.

Lambe, P. (2011). Knowledge organization systems as enablers to the conduct of science. Em *ISKO UK*. Recuperado em 15 de março, 2017, de http://www.iskouk.org/sites/default/files/ISKOUK-2011_PatrickLambe.pdf.

Macias-Chapula, A. M. (1998). O papel da informetria e da cienciometria e sua perspectiva nacional e internacional. Em *Ciência da Informação*, 27(2), 134-140. Recuperado em 11 de março, 2017, de <http://www.scielo.br/pdf/ci/v27n2/macias.pdf>.

Miranda, M. L. C. (2005) *Organização e representação do conhecimento: fundamentos teórico-metodológico na busca e recuperação da informação em ambientes virtuais*. Tese de Doutorado, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.

Moreiro González, J. A. (2011). *Linguagens documentárias e vocabulários semânticos para a web: elementos conceituais*. Salvador: EDUFBA.

Networked Knowledge Organization Systems. (2000). *Taxonomy of Knowledge Organization Sources/Systems*. Recuperado em 02 de abril, 2017, de http://nkos.slis.kent.edu/KOS_taxonomy.htm.

Networked Knowledge Organization Systems. (2016). NKOS. Recuperado em 02 de abril, 2017, de <http://nkos.slis.kent.edu/>.

Oliveira, S. L. (2001). *Tratado de Metodologia Científica: Projetos de Pesquisas, TGI, TCC, Monografias, Dissertações e Teses*. São Paulo: Pioneira.

Ranganathan, S. R. (1967). *Prolegomena to library classification*. Bombay: Asia Publ. House.

Vickery, B. (2008). *On 'knowledge organisation'*. Recuperado em 10 de março, 2017, de <http://web.archive.org/web/20080404103206/www.lucis.me.uk/knowlorg.htm>.

Vignoli, R. G., Souto, D. V. B. e Cervantes, B. M. N. (2013). Sistemas de organização do conhecimento com foco em ontologias e taxonomias. Em *Informação & Sociedade: estudos*, 23(2), 59-72. Recuperado em 20 de março, 2017, de <http://www.ies.ufpb.br/ojs2/index.php/ies/article/view/15160/9685>

Zeng, M. L. (2008). Knowledge Organization Systems (KOS). Em *Knowledge Organization*, 35(2/3). Recuperado em 02 de abril, 2017, de http://nkos.slis.kent.edu/KOS_taxonomy.htm.



ANÁLISE DA PAISAGEM RURAL SOB A ÓTICA DA TEORIA DA ESPACIALIDADE: SUA APLICAÇÃO AOS CENSOS AGROPECUÁRIOS DE 1970 E 2006 E O CAMPO TRAÇADO NO ATLAS DO ESPAÇO RURAL BRASILEIRO (2011)

Vera Lucia Punzi Barcelos Capone¹, Rosa Inês de Novais Cordeiro²

¹*Bibliotecária do Instituto Brasileiro de Geografia e Estatística (IBGE), Doutoranda do Programa de Pós-graduação em Ciência da Informação – Universidade Federal Fluminense (Niterói, Brasil), caponevera@gmail.com*

²*Professora associada do Departamento de Ciência da Informação da Universidade Federal Fluminense (Niterói, Brasil), rosacordeiro@vm.uff.br*

RESUMO Comunicação que aborda o estudo da paisagem rural sob a ótica da teoria da espacialidade no campo da geografia, a partir do método da análise de domínio para a organização do conhecimento, sob o ponto de vista da ciência da informação. Para tal, está sendo desenvolvido um modelo conceitual resultante dos estudos de espacialidades detectados na literatura e os quais serão contextualizados nos Censos agropecuários brasileiros de 1970 e 2006. Os conceitos modelados serão comparados com as articulações do pensamento geográfico traçados no Atlas do espaço rural brasileiro (2011). A aplicação deste modelo conceitual visa à organização do conhecimento e da informação, ou seja, os sistemas de organização do conhecimento (KOS) da ciência da informação geográfica.

PALAVRAS-CHAVE *Análise da paisagem rural, Teoria da espacialidade, Organização e representação do conhecimento, Análise de domínio, Censos agropecuários.*

ABSTRACT Communication that approaches the study of rural landscape from the standpoint of the theory of spatiality in the field of geography, using the method of domain analysis for knowledge organization from the point of view of information science. For such, a conceptual model resulting from spatiality studies, detected in literature and contextualized in Brazilian agricultural censuses of 1970 and 2006, is being developed. The modeled concepts will be compared with the articulation of the geographic thought outlined in the Brazilian rural space atlas (2011). The application of this conceptual model aims at the organization of knowledge and information, that is, knowledge organization systems (KOS) of geographic information science.

KEYWORDS *Rural landscape analysis, Theory of spatiality, Organization and representation of knowledge, Domain analysis, Agricultural censuses.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Esta comunicação é resultante de projeto de pesquisa em desenvolvimento junto ao Programa de Pós-Graduação em Ciência da Informação (PPGCI) da Universidade Federal Fluminense (UFF) no campo da ciência da informação e no âmbito da análise de domínio, direcionado à organização e representação do conhecimento geográfico registrado de diversas naturezas. O objetivo da pesquisa visa cartografar na área da geografia as estruturas lógicas e conceituais relacionadas à teoria da espacialidade, associadas à paisagem rural brasileira (regiões Sul e Centro-oeste). Soma-se a isto, a aplicação dessa rede lógica conceitual no contexto dos censos agropecuários de 1970 e 2006 sob a responsabilidade do Instituto Brasileiro de Geografia e Estatística (IBGE). Ter-se-á como consequência o desenvolvimento de um modelo conceitual que será comparado e intensificado com as articulações do pensamento geográfico traçado no *Atlas do espaço rural brasileiro* do IBGE (2011). Aliando-os aos pressupostos teóricos da ciência da informação e direcionados aos estudos de análise de domínio, o modelo proposto poderá ser aplicado na organização do conhecimento e da informação geográfica, contribuindo para a modelagem de serviços de informação; princípios de linguagens documentárias verbais e notacionais, como por exemplo: tesouros, classificações, ontologias, entre outros. Ou seja, seu propósito engloba os sistemas de organização do conhecimento (*Knowledge Organization Systems - KOS*) conforme elencados por Golub (2014, p. 9), a fim de responder ao pressuposto inicial da pesquisa que vem a ser compreender quais categorias primárias e secundárias governam e configuram a rede de conceitos da área de geografia com foco na teoria da espacialidade, fornecendo elementos que caracterizem a paisagem rural, e auxiliem em sua análise.

A propósito do termo ciência da informação geográfica (*Geographic Information Science - GISc*), sua denominação deve-se a Michel Goodchild, no início dos anos 1990, quando a GISc era observada como contexto teórico para o desenvolvimento dos sistemas de informação geográfica (*Geographic Information System - GIS*), na ocasião em que seu estudo sistemático adquire *status* de uma ciência própria, cujos temas fundamentais decorrem da informação geográfica, compreendendo uma classe do conhecimento bem definida (Longley et al., 2013 & Raper, 2009).

Explorar as particularidades da informação geográfica implica em examinar seus fundamentos teóricos, identificando suas entidades, relacionamentos e processos de acordo com os princípios científicos que norteiam as propriedades da informação geográfica. Seu domínio abrange as dimensões de lugar, espaço e tempo, que aplicadas na leitura da paisagem e consoante aos dados estatísticos contribuem para os processos de análise, organização e representação dos estoques informacionais destas áreas.

Os processos de análise, organização e representação da informação geográfica aplicados na leitura da paisagem de determinada região ressaltam, além dos elementos naturais, práticas de organização da sociedade e suas estruturas econômicas. Quanto à investigação estatística na coleta de informações detalhadas de uma região, a expressão numérica nos revela um conjunto de indicadores que refletem o grupo social que produziu uma determinada paisagem, atestando seu caráter polissêmico correspondente a uma realidade complexa resultante das múltiplas ações humanas.

A pluralidade de paisagens regionais retrata diferentes estilos de vida, espelhando os modos de organização de seus atores sociais. Sob este aspecto, podemos incluir a paisagem rural modelada por sucessivas gerações, políticas de Estado, ordenamento do território, entre outros.

O censo agropecuário reflete o Brasil agrário, cabendo ao IBGE à responsabilidade em planejar, executar e coordenar a atividade censitária no país de acordo com a legislação vigente, investigando os estabelecimentos agropecuários e as atividades aí desenvolvidas. O *Atlas do espaço rural brasileiro* (2011) apresenta-se como a interpretação geográfica das informações agropecuárias obtidas pelos censos agropecuários, destacando-se aquele realizado em 2006, mas também incorpora os resultados levantados desde o ano de 1940, permitindo uma visão histórica da evolução do espaço rural brasileiro.

A escolha dos censos agropecuários como objeto de análise é motivada pelo seu potencial informativo como fonte de pesquisa em torno da vida rural e a possibilidade de se extrair deles informações acerca da paisagem rural e também se verificando como o conhecimento estabelecido na literatura da área de geografia é sistematizado nos censos. Ainda, os censos agropecuários estruturam os conceitos direcionados à espacialidade, comprovando as mudanças ocorridas no espaço rural através das diferentes formas de apropriação e uso do território. Outrossim, os censos agropecuários representam a conjunção de idéias instituídas de áreas do conhecimento, geografia e estatística, que constituem um domínio.

A partir da década de 1970, o setor agropecuário apresenta uma expansão e modernização tecnológica através da ampliação de investimentos públicos na infraestrutura, no armazenamento e na modernização da produção agropecuária em grande escala.

O recenseamento geral de 1970 foi a oitava operação censitária englobando o conjunto do país, constituindo-se dos censos demográfico, predial, agropecuário, industrial, comercial e dos serviços, além de inquéritos especiais sobre as instituições de crédito e seguradoras. O censo agropecuário de 1970 foi o quinto a ser realizado no país, tendo abrangido as seguintes atividades: agricultura, pecuária, avicultura, apicultura, cunicultura, sericultura, horticultura, floricultura, silvicultura e extração de produtos vegetais.

O censo agropecuário de 2006 teve por escopo retratar o Brasil agrário considerando a inter-relação com atores, cenários, modos e instrumentos de ação e para tal implementou um processo de refinamento metodológico com respeito ao conteúdo da pesquisa relativo às novidades que se apresentaram ao universo agrícola nacional. Entre as variáveis investigadas, detalhou os estabelecimentos agropecuários, as atividades agropecuárias desenvolvidas abrangendo informações sobre as características do produtor, dos estabelecimentos, economia e emprego no meio rural, pecuária, lavoura e agroindústria.

Também foram coletadas informações sobre as diferentes práticas agrícolas e formas de ocupação da área; manejo e conservação do solo; utilização de agrotóxicos; agricultura orgânica; atividades agrícolas; e disponibilidade de água dentro dos estabelecimentos. O questionário eletrônico aplicado à pesquisa censitária foi a novidade tecnológica com o emprego do *Personal Digital Assistant* - PDA.

O censo agropecuário não apenas constitui uma representação do real em termos de economia rural, mas implica em diferentes sentidos que podem ser atribuídos ao espaço geográfico quando estudado sob a ótica da configuração espacial decorrente do arranjo entre elementos naturais e artificiais de uso social, entre os quais as plantações e técnicas adotadas, sobre o território (Santos, 2014).

Esses arranjos entre elementos naturais e artificiais implicam em como o espaço se organiza. Condensando estes elementos numa segmentação da realidade envolvendo região, lugar e tempo, juntos refletem características que são recortes de conteúdo, passíveis no contexto informacional, de organização e representação para transmissão e uso efetivo resultante do processo da análise de domínio.

A comparação das paisagens depreendidas no censo agropecuário em relação ao *Atlas do espaço rural brasileiro* (2011) evidencia as informações levantadas pelos censos agropecuários entre os anos de 1940 e 2006, exercendo ao mesmo tempo a função de materialização da paisagem através dos mapas. O *Atlas* traça um panorama no processo de ocupação do espaço rural brasileiro, examinando as transformações ocorridas no setor agropecuário no intervalo censitário de 1940 a 2006, considerando variáveis passíveis de comparação, tais como: evolução dos estabelecimentos, número e área; utilização da terra; as principais atividades produtivas; total de pessoal ocupado; número de tratores; efetivos da pecuária, entre outros, levando em conta as Unidades da Federação. Este é um estudo publicado em 2011, não tendo sido reeditado até a presente data.

A aplicação da teoria da espacialidade à análise da paisagem rural vai ao encontro de um sistema de realidades estruturado por um *corpus* teórico construído em torno de um conjunto de conceitos referentes à natureza do espaço, que nos permitem propor um modelo conceitual relacionados à paisagem e seus constituintes, considerando-se a vida que a preenche:

O espaço deve ser considerado como um conjunto indissociável, de que participam, de um lado, certo arranjo de objetos geográficos, objetos naturais e objetos sociais, e, de outro, a vida que os preenche e os anima, ou seja, a sociedade em movimento. O conteúdo (da sociedade) não é independente da forma (os objetos geográficos), e cada forma encerra uma fração do conteúdo (Santos, 2014, p. 31).

A questão da espacialidade se reveste de significância no domínio da geografia, pois se relaciona à disposição física dos elementos da natureza na superfície terrestre e direcionam nosso olhar. Esses elementos combinados, isto é, o lugar de cada elemento, configura uma paisagem. Gomes (2013, p. 17) trata essa questão com muita propriedade: “Espacialidade é o conjunto formado pela disposição física sobre esse plano de tudo que ele contém. Corresponde, assim, ao resultado de um jogo de posições relativas de coisas/e ou fenômenos que se situam, ao mesmo tempo, sobre esse mesmo espaço.”

A estrutura teórica direcionada à compreensão da paisagem importa ao campo da geografia, sendo a realidade perceptível o resultado da ação do homem na produção do espaço por meio de objetos naturais e artificiais que resultam numa configuração territorial, sendo estas as principais categorias defendidas por Santos na análise da paisagem:

No começo da história do homem, a configuração territorial é simplesmente o conjunto dos complexos naturais. À medida que a história vai fazendo-se, a configuração territorial é dada pelas obras dos homens: estradas, plantações, casa, depósitos, portos, fábricas, cidades etc; verdadeiras próteses. Cria-se uma configuração territorial que é cada vez mais o resultado de uma produção histórica [...] (Santos, 2009, p. 62).

Cada tipo de paisagem reflete as diferentes forças produtivas estando os diversos níveis de realidade associados à porção visível do espaço geográfico, refletindo sua fração aparente. Vista desse modo, as paisagens podem ser divididas em três categorias, quais sejam naturais, modificadas e organizadas. Cruz (2002), com base nos estudos de Dollfus (1991), ressalta que a paisagem natural é aquela que não sofreu a ação do homem. Em contrapartida, a paisagem modificada apresenta transformações provocadas pelo homem, tais como as queimadas ou aquelas resultantes de atividades pastoris. As paisagens organizadas derivam de ações meditadas, combinadas sobre o meio ambiente.

Saquet (2015) examina uma vasta literatura que aborda o conceito de paisagem em estudos pioneiros de geógrafos italianos, cuja ênfase recai sobre as relações entre forma e estrutura da paisagem associada

a fatores sociais. Com base nas pesquisas de Quaini (1973, 1973a), Saquet (2015, p. 140) vê a paisagem agrária “[...] como um produto histórico, com descontinuidade e continuidade, ou seja, com mudanças e permanências que resultam de uma combinação de fatores ambientais e históricos, especialmente, de técnicas, sistemas de cultivo e de circuitos comerciais que estão presentes na organização territorial.”

A incorporação de sistemas técnicos ao espaço rural brasileiro forjou a implementação de inúmeras redes, redesenhando a paisagem rural das regiões que se pretende estudar, imprimindo um novo conteúdo aos recortes regionais do território. Esse novo arranjo que se dá pela presença dos objetos técnicos é impregnado de informação: as atividades mais modernas, na cidade e no campo, passam a exigir adaptações do território, com a adição ao solo de acréscimos cada vez mais baseados nas formulações da ciência e na ajuda da técnica (Santos, 1994, p. 73).

Compreender o espaço significa adentrar numa teia de relações, formas, funções, organizações, estruturas, entre outros aspectos, que expõem as distintas maneiras de viver, num mesmo país, conjuntamente a seus meios de produção, porém em diferentes regiões. Nesta conjuntura, o estudo regional desempenha importante papel na apreensão dos diferenciados modos de vida nas diferentes regiões exercendo a função de categoria de análise, de acordo com afirmação de Santos (2014, p. 53): “Hoje a região, o regional, a regionalização têm de ser assim entendidos”.

No contexto da ciência da informação, a questão que se impõe à informação geográfica é como este campo teórico deve ser compreendido. Sobre a natureza e o conceito de espaço:

Uma das fontes mais frequentes de dúvida entre os estudiosos do tema parece ser o próprio conceito de espaço [...]. Como, na teoria e na prática, levar em conta os ingredientes sociais e “naturais” que compõem o espaço para descrevê-lo, defini-lo, interpretá-lo e, afinal, encontrar o espacial? O que caracteriza a análise do espaço? (Santos, 2009, p. 11).

Em artigo de revisão, Raper (2009) nos indica que o ponto de partida para a definição de informação geográfica e seus subcampos, entre eles a ciência da informação geográfica (*Geography Information Science – GISc*) depende de sua representação, ou seja, como se desenvolve sua construção teórica. Longley et al. (2013) argumentam que seu estudo sistemático constitui uma ciência própria e como tal seus temas fundamentais necessitam decorrer da informação geográfica enquanto uma classe bem definida da informação em geral.

Compreender e representar um domínio implica em conhecer e organizar seu conjunto de termos e conceitos adotados empregando-se um propósito lógico, cujo recorte reflita o objeto de estudo de uma comunidade científica, e qual o seu papel na sociedade, visto que a informação adquire o *status* de conhecimento quando compartilhada, descrita e disseminada. Uma comunidade de discurso é representada por seus pesquisadores, refletindo a construção social do conhecimento à medida que são identificados os pontos comuns que reunidos compartilham correntes teóricas semelhantes, porém autônomas, delimitadas por meio de fronteiras epistemológicas bem definidas e individualizam terminologias, taxonomias e organização padronizada do conhecimento, por exemplo (López-Huertas, 2006, p. 211, tradução nossa).

Segundo Hjørland e Albrechtsen (1995, p. 400, tradução nossa): “O paradigma da análise de domínio na ciência da informação estabelece que a melhor maneira de compreender uma comunidade de discurso é estudar os padrões e estruturas que compõem e refletem seus objetos”. Para esses dois autores, o paradigma na análise de domínio possui três dimensões: primeiramente, o caráter social tendo em vista a perspectiva social da ciência da informação. Segundo, a sua abordagem funcionalista dirigida ao

entendimento do mecanismo implícito e explícito da informação. Por último, o caráter filosófico-realista que busca encontrar as bases para a ciência da informação, cujos fatores externos às percepções subjetivas individuais se opõem ao paradigma cognitivo e comportamental (Hjørland & Albrechtsen, 1995, p. 400, tradução nossa).

A terceira dimensão abordada pelos autores, a partir de Wilson (1993), é que uma especialidade deveria ser vista em sua totalidade e não individualmente: “A questão não é com um aspecto individual pode afetar um grupo [uma comunidade discursiva], mas sim sua totalidade, visto que a informação possui um traço coletivo” (Hjørland & Albrechtsen, 1995, pp. 401, tradução nossa).

Análise de domínio em ciência da informação pode ser definida como uma metateoria, tendo em conta sua capacidade em explorar e analisar as definições, o escopo e o alcance, bem como o nível de especialização de um assunto ou uma comunidade de interesse, segundo Tennis (2012, p. 3-8), com base em Hjørland e Hartel (2003).

A tarefa de organização do conhecimento pressupõe uma análise dos elementos que constituem um domínio, uma área do conhecimento, em como sistematizá-lo, procurando entender logicamente a estrutura das categorias e conceitos e como estes se relacionam semanticamente visando um objetivo específico, que vem a ser o atendimento às necessidades de informação das diversas comunidades de usuários, através do diálogo que necessita ser estabelecido entre estas comunidades e as unidades informacionais (Hjørland, 2010).

Em relação à análise de domínio, Smiraglia (2012), segundo estudo de Guimarães (2014), para que bibliotecas e centros de informação atendam a objetivos específicos, as fontes de informação devem ser identificadas, descritas, organizadas e disseminadas. Para isso, elenca onze diferentes abordagens que podem caracterizar uma análise de domínio visando o atendimento às necessidades informacionais das comunidades de usuários, a saber: 1. Produção de guias de literatura; 2. Construção de classificações e tesouros especializados; 3. Indexação e recuperação da informação especializada; 4. Estudos empíricos de usuários; 5. Estudos bibliométricos; 6. Estudos históricos; 7. Estudos de documentos e gêneros; 8. Estudos epistemológicos e críticos; 9. Estudos terminológicos, linguagens para propósitos determinados, semântica de bases de dados e estudos dos discursos; 10. Estrutura e instituições da comunicação científica; e 11. Cognição científica, conhecimento perito e inteligência artificial.

López-Huertas (2015) ratifica a linha de pensamento de Hjørland e Albrechten ao dedicar estudos acerca da análise de domínio como sendo um caminho para se encontrar a estrutura do conhecimento, a dinâmica, a linguagem e o padrão de comunicação de uma comunidade de discurso, que representa o esforço de síntese das várias contribuições, relacionando-se diretamente com a questão da interdisciplinaridade interessando aos estudos de representação do conhecimento e recuperação de informações.

Percebe-se, desse modo, que a análise de domínio é uma abordagem metodológica sistematizada por meio da qual se promove a organização do conhecimento com vistas à indexação automática e recuperação da informação (Albrechtsen, 2015, pp. 560-561, tradução nossa).

METODOLOGIA

A natureza da presente pesquisa qualifica-se como uma abordagem qualitativa, tendo em vista representar de forma substancial a dinâmica das estruturas lógicas e conceituais que moldam os estudos de espacialidade compreendidos na literatura e produtos/dispositivos (censos agropecuários e *Atlas do espaço rural brasileiro*) da área de geografia. No contexto da ciência da informação geográfica volta-se para a organização e representação do conhecimento registrado no processo de transferência da informação, no âmbito da análise de domínio com vistas à inclusão nos KOS para posterior acesso e uso destes documentos.

Diante dos objetivos de se compreender o *corpus* teórico que estruturam a rede conceitual de espacialidade relacionados à paisagem rural, esta investigação tem como metodologia a proposta de análise de domínio estruturada na ciência da informação por Birger Hjørland (2002), partindo das abordagens sugeridas pelo autor, adotando-se a segunda abordagem, ou seja, a classificação especializada dirigida para o modo como os conceitos e suas relações semânticas se estruturam, de modo a orientar na compreensão dos modelos teóricos de espacialidade relativos à paisagem rural, objetivando sua recuperação pela perspectiva do usuário.

Os censos agropecuários de 1970 e 2006 juntamente com o *Atlas do espaço rural brasileiro* (2011) são as fontes documentais empregadas no desenvolvimento desta investigação, além da determinação de um núcleo significativo de autores relevantes já identificados na literatura da área da geografia, a saber: Gomes (2013), Quaini (1973), Santos (1994, 2009, 2014), Saquet (2015), entre outros, que tornam possível reconhecer e validar o que importa à análise da paisagem rural e viabiliza a garantia literária oriunda da rede conceitos da comunidade discursiva, configurando um domínio. Nas palavras de Smiraglia (2012, p. 113, tradução nossa), esse núcleo que produz conhecimento, proporciona outros novos: “A atividade social de uma comunidade de estudiosos é orientada por objetivos e limites a fim de produzir conhecimento e que este seja retido funcionando como uma fortaleza intelectual que proteja seus membros da influência externa”.

Esse domínio do conhecimento será detectado mediante duas dimensões, quais sejam: a dimensão da garantia literária, com base na literatura da área de geografia e pela dimensão empírica de dispositivos que registram o “consenso”, ou melhor, uma certa concordância, de uma comunidade discursiva representada nos censos agropecuários de 1970 e 2006 e no *Atlas do espaço rural brasileiro* (2011). O modelo teórico visa evidenciar as estruturas lógicas e conceituais desse domínio, para fins de aplicação nos sistemas de organização do conhecimento (KOS) na área de geografia, teoria da espacialidade e a paisagem rural fundamentado por essas duas dimensões selecionadas.

Em suma, nos procedimentos metodológicos conjuga-se ao estudo do pensamento geográfico acerca dos fundamentos teóricos que moldam o espaço agrário, as categorias conceituais que se aplicam à definição de espaço que se relacionam à paisagem rural por meio das regiões selecionadas, a saber: as regiões Sul e Centro-oeste, em vista dos diferentes aspectos espaciais, regionais, socioculturais e temporais que particularizam a paisagem das regiões indicadas nesta comunicação.

RESULTADOS PRETENDIDOS

Por meio da análise de domínio na área de geografia, o presente estudo permitirá a elaboração de um modelo conceitual de espacialidade relacionado à paisagem rural, que poderá ser aplicado nos sistemas de organização do conhecimento (KOS). Consequentemente, a aplicação desses protótipos teóricos detectados na literatura das áreas de geografia possibilitará comparar as categorias identificadas nos censos agropecuários de 1970 e 2006 daquelas representadas no *Atlas do espaço rural brasileiro* (2011).

Pretende-se também com esta pesquisa contribuir para a ampliação de temas e discussões no âmbito da informação geográfica mediante o uso da metodologia da análise de domínio.

CONSIDERAÇÕES FINAIS

Sob o viés da ciência da informação, percebe-se que o conceito de informação geográfica não se limita apenas aos mapas, fato que enfraqueceria sua conceituação. Em sentido lato, a representação da informação geográfica se constitui das dimensões de lugar, espaço e tempo, integrando-se ao subconjunto da ciência da informação, conforme o entendimento de Raper (2009), que chama atenção para o fato da ciência da informação se debruçar pouco sobre a informação geográfica, embora ofereça um vasto campo para discussão, visto seu aspecto multidisciplinar. Reconhece o desafio à ciência da informação na compreensão da representação do espaço geograficamente distribuído, pois como dado a ser recuperado possui seu estatuto próprio com relações e características variáveis que também requer tratamento, padronização e princípios de análise criteriosos, que assim oferece um considerável espectro de informações que tendem a se multiplicar produzindo novos conhecimentos.

REFERÊNCIAS BIBLIOGRÁFICAS

Albrechtsen, H. (2015). This is not domain analysis. *Knowledge Organization*, 42(8), pp. 557-561.

Atlas do espaço rural brasileiro. (2011). Rio de Janeiro: IBGE. Recuperado de <http://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=263372>.

Censo agropecuário: 1970. (1970). Rio de Janeiro: IBGE. 3 v. Recuperado de <http://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?acervo=periodico&campo=titulo&opeqry=&texto=censo%20agropecuario&digital=false&fraseexata=>.

Censo agropecuário: 2006: Brasil, grandes regiões e unidades da federação: segunda apuração. (2012). Rio de Janeiro: IBGE. Recuperado de <http://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?acervo=periodico&campo=titulo&opeqry=&texto=censo%20agropecuario&digital=false&fraseexata=>.

Cruz, R. de C. A. da. (2002). As paisagens artificiais criadas pelo turismo. In Yázigi, E. (Org.), *Turismo e paisagem*. São Paulo: Contexto.

Dollfus, O. (1991). *O espaço geográfico*. 5 ed. Rio de Janeiro: Bertrand Brasil.

Golub, K. (2014). *Subject access to information: an interdisciplinary approach*. Santa Barbara, CA: ABC-CLIO, pp. 1-40.

Gomes, P. C. da C. (2013). *O lugar do olhar: elementos para uma geografia da visibilidade*. Rio de Janeiro: Bertrand Brasil.

Guimarães, J. A. C. (2014). Análise de domínio como perspectiva metodológica em organização da informação. *Ciência da Informação*, Brasília, DF, 41(1), pp. 13-21.

Hjørland, B. Domain analysis in information science: eleven approaches – traditional as well as innovative. (2002). *Journal of Documentation*, pp. 422-462, 2002.

_____. Domain analysis in information science. (2010). In Bates, M. J., & Maack, M. N. (Ed.), *Encyclopedia of Library and Information Sciences*. 3rd. ed. New York: Taylor and Francis. v. 7, pp. 1648-1654.

Hjørland, B. & Albrechtsen, H. (1995). Toward a new horizon in information science: domain-analysis. *Journal of the American Society for Information Science*, Hoboken, 46(6), pp. 400-425.

Hjørland, B. & Hartel, J. (2003). Afterward: ontological, epistemological and social dimensions of domains. *Knowledge Organization*, 30(3/4), pp. 239-245).

Longley, P. A. et al. (Org.). (2013). *Sistemas e ciência da informação geográfica*. 3. ed. Porto Alegre: Bookman.

López-Huertas, M. J. (2006). Análisis del dominio interdisciplinar para la representación y organización del conocimiento. In González de Gómez, M. N. & Orrico, E. G. D. (Org.). *Políticas de memória e informação: reflexos na organização do conhecimento*. Natal: EdUFRN, pp. 209-235.

_____. (2015). Domain analysis for interdisciplinary knowledge domains. *Knowledge Organization Journal*, 42(8).

Quaini, M. (1973). *Per la storia del paesaggio agrario in Liguria: note di geografia storica sulle strutture agrarie della Liguria medievale e moderna*. Savona: C.C.I.A.A.

_____. (1973a). Geografia storica o storia sociale del popolamento rurale? *Quaderni Storici*, n. 24, pp. 691-744.

Raper, J. (2009). Geographic information science. *Annual Review of Information Science and Technology*, 43(1), pp.1-117.

Santos, M. (1994). *Técnica, espaço, tempo: globalização e meio técnico-científico informacional*. São Paulo: Hucitec.

_____. (2009). *A natureza do espaço: técnica e tempo. Razão e emoção*. 4. ed. 5. reimpr. São Paulo: EdUSP.

_____. (2014). *Metamorfoses do espaço habitado: fundamentos teóricos e metodológicos da geografia*. 6. ed. 2. reimp. São Paulo: Edusp.

Saquet, M. A. (2015). Território e paisagem: da construção material à representação?. In _____. *Abordagens e concepções de território*. 4. ed. São Paulo: Outras Expressões.

Smiraglia, R. P. (2012). Epistemology of domain analysis. In: Lee, H. L. & Smiraglia, R. P. (Ed.). *Cultural frames of knowledge*. Wurzburg: Ergon, pp. 111-124.

Tennis, J. T. (2012). What does a domain analysis look like in form, function, and genre?. *Braslian Journal of Information Science: research trends*, Marília, 6(1), pp. 3-15. Recuperado de <http://www2.marilia.unesp.br/revistas/index.php/bjis/index>

Wilson, P. (1993). Communication efficiency in research and development. *Journal of the American Society for Information Science*, 44, pp. 376-382.

DESAFIOS À ORGANIZAÇÃO E ACESSO AO PATRIMÓNIO CULTURAL



TRATAMENTO E ANÁLISE DOS OBJETOS PRESENTEADOS AOS REITORES DE UNIVERSIDADES: QUESTÕES SIMBÓLICAS E DIPLOMÁTICAS

Samanta Prado¹, Luzia S. F. Costa², Luciana S. Gracioso³

¹Universidade Federal de São Carlos (UFSCar). Dept. Ciência da Informação, 0000-0002-7386-3419, samanta_prado@hotmail.com

²Universidade Federal de São Carlos (UFSCar). Dept. Ciência da Informação, 0000-0001-8530-4000, luziasigoli@gmail.com

³Universidade Federal de São Carlos (UFSCar). Dept. Ciência da Informação, 0000-0002-6320-4946, luciana@ufscar.br

RESUMO A Universidade Federal de São Carlos (UFSCar), ao longo de sua existência, acumulou um significativo número de objetos, recebidos pelos reitores. Esses objetos representam, à luz das ciências sociais, as relações diplomáticas construídas entre os atores sociais, bem como a circulação de saberes. O objetivo do projeto foi reunir, tratar tecnicamente e analisar a coleção dos mimos - "objetos presenteados"- recebidos pelos reitores da UFSCar ao longo dos 45 anos da universidade. Metodologicamente a pesquisa realizou-se por meio de 4 etapas: reunião do acervo e definição das coleções compostas pela séries de objetos, tais como troféus, placas, moedas comemorativas, e outros; inventariação dos objetos à maneira museológicas, com descrição física, conteúdos informacionais e registros fotográficos; composição de um catálogo digital, contendo as informações textuais e fotográficas, em acordo com a Associação Brasileira de Normas Técnicas (ABNT). O objeto de estudo desta pesquisa correspondeu a uma coleção de 115 objetos. No decorrer da pesquisa, alguns objetos se destacaram pelos significados e relações que representam para a história da UFSCar. Alguns desses objetos contêm informações escritas sobre premiações, reconhecimentos, alianças e agradecimentos, referentes tanto à figura do reitor quanto à universidade como instituição. O estudo realizado pode ser visto como uma contribuição para a preservação da memória e para a construção da identidade institucional da UFSCar, que se revela marcada pela integração e pela cooperação com associações e outras instituições.

PALAVRAS-CHAVE *Memória Institucional, Coleções Históricas; Patrimônio Cultural*

ABSTRACT The Federal University of São Carlos (UFSCar), over your existence, has accumulated a significant number of objects received by the rectors. These objects represent, in the light of the social sciences, diplomatic relations built between the social actors, as well as the circulation of knowledge. The objective of the project was to gather, treat technically and analyze collection of Pats-"present objects"- received by the rectors of UFSCar over the 45 years of the University. Metodologicamente the survey was held through 4 steps: meeting of collection and definition of compound collections the series of objects, such as trophies, plaques, commemorative coins, and others; inventory of the objects the Museum way, with physical description, informational content and records taking photographs; composition of a digital catalogue, containing textual and photographic information, in accordance with the Brazilian Association of technical standards (ABNT). The object of study of this research corresponded to a collection of 115 objects. In the course of the research, some outstanding objects by meanings and relationships that represent for the history of UFSCar. Some of these objects contain written information about awards, recognitions, alliances and thanks, referring to both the Dean's figure as the University as an institution. The study can be seen as a contribution for the

preservation of memory and the construction of the institutional identity of UFSCar, which is marked by integration and cooperation with associations and other institutions

KEYWORDS *Institutional Memory, Historical collections; Cultural Heritage.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A Universidade Federal de São Carlos (UFSCar), localizada no interior do Estado de São Paulo/Brasil, completou 45 anos em 2015. Ao longo de sua existência, a instituição acumulou um significativo número de objetos, recebidos pelos reitores que por ela passaram. Esses objetos representam as relações diplomáticas construídas ao longo da história da universidade, bem como a circulação de saberes e significados que fizeram parte dessa história. A troca de pequenos mimos e agrados, à luz das ciências sociais, pode ajudar a entender as relações políticas estabelecidas e a importância estratégica que existe na troca de presentes entre personalidades e representantes públicos.

A coleção de objetos e agrados, guardada na Reitoria da UFSCar, é formada por presentes diplomáticos. São peças de valor estético e, às vezes, incomuns, que materializam as relações estabelecidas com lideranças, personalidades públicas e outros atores sociais. Esses objetos carregam significados relacionados ao contexto das trocas ou recebimento. Por essa razão, os contextos agregam valor aos objetos, portanto são relevantes para o processo de pesquisa que aqui relatamos. Nesse sentido, a descrição física dos objetos nos revela apenas uma dimensão dos mesmos, a dimensão material, que deve ser relacionada à dimensão simbólica. Só a partir dessa relação é possível compreender o papel da coleção na construção da história e da memória da UFSCar.

Para o desenvolvimento da pesquisa (UFSCAR, 2017), foi possível propor algumas questões de estudo. Essas questões foram as seguintes: O que representam os presentes aos reitores da universidade, acumulados nas últimas quatro décadas, em termos de prática social? Que vivências eles simbolizam através de sua materialidade? E que leitura nos possibilitam da história da universidade? Qual face dessa história essa coleção torna visível? Quais conceitos da Ciência da Informação e da Ciência Sociais foram úteis para pensarmos nosso objeto de estudo?. A pesquisa visou e possibilitou, em certo grau, ainda que de forma limitada, responder a essas perguntas.

Em relação à área da Ciência da Informação, tratamos aqui os mimos/presentes recebidos pelos reitores como objetos enquanto documentos informacionais. Para Paul Otlet (1937), um dos principais estudiosos da área de Documentação, o documento tem a capacidade de informar, organizando e disseminando esta informação em qualquer tipo de suporte até então conhecido. Neste caso, até mesmo os objetos tridimensionais, aqueles relacionados ao museu (artefatos, objetos, esculturas) também, na concepção de Otlet, são considerados documentos. Logo, nessa perspectiva, entende-se documento como algo além daqueles que são somente escritos. De forma geral, o documento agrega três ordens de elementos: materiais (sua substância, forma e acabamento), gráficos (textos, imagens) e elementos intelectuais. Os documentos (Crippa, Mostafa, 2011) são carregados de características sociais, de processos de comunicação, tanto de acumulação quanto de transmissão de conhecimento, de instrumento de pesquisa e de cultura.

A documentação (1937) está presente em todas as fases do tratamento do documento. Ela é um processo de operações que envolvem elementos humanos (autor, copista) e elementos materiais (produção, distribuição, conservação e utilização) e imateriais (mundo das ideias). De forma geral, como apontado por Litton (1976), a necessidade de documentar a pesquisa vai além de registrar os dados para as populações futuras. A documentação tem o papel de registrar novas informações, assim como classificar os dados. No contexto desta pesquisa, a documentação visa informar e educar o usuário da informação por meio da descrição física, ou seja, da representação descritiva e da explicitação das relações que foram estabelecidas ao longo dos anos por certos atores sociais, da universidade ou advindos de outros lugares, reconstituindo, assim, parte da memória da instituição.

No que tange a pesquisa em questão, os mimos como documentos refletem a informação contida nas relações estabelecidas entre agentes, instituições, alianças, etc. Como patrimônio, estes objetos se relacionam com uma memória, considerada um ponto crucial para sua existência.

Podemos afirmar que o patrimônio é o resultado de um trabalho de memória que, no correr do tempo e dos critérios extremamente variáveis, seleciona determinados elementos herdados do passado para coloca-los na categoria dos objetos patrimoniais. Definitivamente, o patrimônio funciona como aparelho ideológico da memória e esta se coloca como estímulo para a composição do patrimônio.” (Crippa, Mostafa, p.60)

Esses objetos cunhados como patrimônios (Soares, 2011), uma vez que caracterizam um conjunto de bens culturais, institucionais, estão relacionados com a ideia da construção de uma identidade social, científica. Podemos relacionar o patrimônio à ideia de um documento com alto valor agregado, já que carrega em si aspectos históricos, ideais, simbólicos.

Patrimônio pode ser definido como um conjunto de bens materiais e imateriais herdados dos antepassados. Ele está associado à lembrança e é uma fonte para a construção da memória de uma determinada sociedade, comunidade. Ele representa a identidade e as diferenças étnicas em relação a outros grupos, ou seja, as particularidades e especificidades de um local, região da sociedade (SOARES; KLAMT, 2004) que pode ser representado por objetos, documentos, edifícios, conjuntos urbanos, sítios de valor histórico, criações científicas, artísticas e tecnológicas, bem como todo objeto ou manifestação compartilhada que diga respeito a identidade e a memória de um grupo.

O patrimônio pode ser reconhecido em uma crença, um costume, um objeto, dentre tantas outras possibilidades. A noção de conservação e preservação do patrimônio cultural é fundamental para as futuras gerações como elemento da construção da própria história e da memória institucional.

Pode-se dizer que a construção da identidade (UFSCAR, 2017) é dada por um conjunto de crenças e costumes que foram herdados por um determinado grupo. Sua preservação forma a memória e esta é refletida por meio do patrimônio cultural. O patrimônio pode ser inventado ou reinventado de várias formas, de acordo com o que o grupo determina ou preserve como um patrimônio que os represente enquanto uma comunidade. O patrimônio histórico institucional informa as características sociais, a sua história, os costumes predominantes, e as ideias que circulam em determinados períodos. No caso desta pesquisa, os objetos informam os acordos, as alianças, as amizades que foram construídas ao longo dos anos de existência da instituição. Segundo Soares (2007), o objeto não é criado para ser um patrimônio, ao longo do tempo pela sua identificação com determinado grupo, suas características, enquadra-se em um determinado contexto histórico, diferenciando dos demais grupos pelas suas particularidades, sua cultura.

Em relação a questão cultural que está inserida no mimo/objeto, devemos caracterizar o que a palavra “cultura” significa para nós. Nesse sentido, a concepção de cultura de Raymond Williams pode nos trazer uma visão interessante.

Williams (2007) foi um dos intelectuais fundadores da linha dos estudos culturais, que nasce na Inglaterra dos anos 1950 e 1960. Em suas origens, esse movimento teórico representou um processo de ressignificação da cultura, sendo Williams um de seus principais debatedores.

A diferença fundamental que a contribuição de Williams traz ao debate é a percepção materialista de cultura: os bens culturais são resultados de meios também eles materiais de produção (indo desde a linguagem como consciência prática aos meios eletrônicos de comunicação), que concretizam relações sociais complexas envolvendo instituições, convenções e formas. Definir cultura é pronunciar-se sobre o significado de um modo de vida. Esse o vasto campo de estudo e intervenção aberto aos estudos culturais no momento de sua formação (Cevasco, 2012, p. 23).

Podemos perceber, no destaque de Cevasco (2012), alguns aspectos importantes do pensamento do autor que servem aos propósitos deste estudo. Um primeiro ponto é justamente a visão materialista de cultura. Para Williams, a cultura não se resume a um conceito abstrato, ligado apenas às produções intelectuais e morais da humanidade, como em “cultivo do espírito”. Ela corresponde também, e talvez principalmente, à experiência humana em sua dimensão material e social. Os bens culturais, para esse autor, são aqueles que concretizam as relações sociais, podendo estas estarem associadas aos mais diversos contextos. Desse modo, podemos pensar em “cultura universitária” ou “cultura institucional”, sendo estas definidas pelas convenções e formas reconhecidas em seu contexto próprio.

Assim, podemos pensar a coleção de mimos dos reitores como resultado de certas formas de comportamento e convenções sociais que da alta hierarquia universitária. Esses objetos arquivados na UFSCar podem ser vistos como produtos culturais que materializam uma prática social, e estabelecem um elo entre o material e o simbólico.

Outra referência que cabe neste trabalho são os estudos de Marcel Mauss (2003), que buscam o entendimento da constituição da vida social por meio de relações de trocas e reciprocidades. Ele, em sua obra, define a “dáviva” de forma ampla, abrangendo presentes, visitas, comunhões e heranças. Nessa roga um entendimento da constituição da vida social por processos que se materializam no cotidiano das dádivas, que nada mais são do que os atos de “dar” e de “retribuir”. As trocas podem ser concebidas de formas distintas, variando de tempo/época e lugares.

Os objetos, nesse caso, passam a representar as relações de sociabilidade, reciprocidade, fidelidade e gratidão, e estão associados a um determinado momento, de um encontro ou estada do Reitor em algum lugar, ou representam uma passagem de alguma outra personalidade pela UFSCar. Nesse sentido, as contribuições de Mauss (2003) sobre os conceitos de troca e reciprocidade assumem importância teórica para a interpretação das situações de recebimento dos objetos. Também são pertinentes os conceitos de sociabilidade, reciprocidade, fidelidade e gratidão, propostos por Simmel (1983 & 2004).

Simmel (2004) discute o papel que a fidelidade adquire nas relações sociais. Para este autor, este papel é duplo: por um lado, a fidelidade complementa outros sentimentos; de outro lado ela também sustenta relações e condições sociais.

a existência de um estado psíquico e sociológico específico, que assegura a continuação de uma relação para além das forças que primeiro lhe deram origem; um estado que sobrevive a essas forças com o mesmo efeito sintetizador que elas próprias tinham originalmente (Simmel, 2004, p. 32).

A etnografia demonstra a incorporação de formas de sociabilidade e de reciprocidade, assim como de relações de fidelidade e gratidão no conteúdo das trocas. Em conteúdo, produzido pelos envolvidos, está presente nas manifestações culturais da instituição, mas, ao mesmo tempo, mescla-se com as histórias de vida, na medida em que proporciona novos rumos, cria “pontes”, abre novas “portas” e dá reforço às condições de cidadania.

Uma terceira referência adequada para este trabalho é pensarmos a importância da pesquisa para a construção da memória institucional da UFSCar. Contextualizar os objetos e os momentos, situando-os dentro de um determinado tempo e espaço, além de tornar mais claras as relações diplomáticas e seus objetivos para a universidade, contribui para a construção dessa memória. Conhecer a história da UFSCar por meio do estudo dos objetos reunidos, conhecer os convênios que foram firmados entre a instituição e outros órgãos, os tipos de alianças que foram formadas, as vivências que os mimos proporcionaram para estes atores, como a criação de laços voltados para o desenvolvimento da educação, de incentivo a programas educacionais, entre outros, é um movimento importante para a construção da memória da universidade.

Neste projeto, utilizamos o termo “memória institucional” com a justificativa de que a análise de um conjunto de objetos pode contribuir para a construção da história da UFSCar. Pensando a universidade como uma instituição de conhecimento, de saberes, que é legitimada e sustentada pela sociedade, sua memória se torna um elemento importante para a comunidade e para o processo de construção da própria instituição.

Nos estudos de Costa (1997), a memória é um elemento fundamental no funcionamento das instituições. Por meio da memória as instituições se reproduzem na sociedade, retendo apenas as informações que interessam para seu funcionamento. A Memória Institucional (MI), para a autora, está em constante evolução, pois é elaborada com o passar do tempo, sendo o reflexo de uma trajetória social e histórica. Uma característica importante do conceito de MI, segundo Costa, é que este inclui as relações estabelecidas entre as instituições em uma sociedade, e não apenas as informações internas de uma única instituição. Por essa razão, esse conceito é adequado para o estudo aqui realizado, pois, por meio deste, foi possível destacar relações entre a UFSCar e outras instituições.

Em síntese os principais conceitos utilizados para construir o olhar teórico dessa pesquisa foram: o conceito de documento de Paul Otlet, o conceito de patrimônio de André Soares, o conceito de cultural de Raymond Williams, a dádiva de Marcel Mauss, os conceitos de sociabilidade, reciprocidade, fidelidade e gratidão de Simmel e o conceito de Memória Institucional de Icléia Costa.

O objeto de estudo desta pesquisa correspondeu a uma coleção de 115 objetos. Todos esses objetos foram presentes recebidos pelos reitores da UFSCar, ao longo dos 45 anos da história da instituição, por ocasião de uma variedade de eventos, estadias, visitas, parcerias, associações e outros tipos de relações sociais de que a universidade participou.

Essa coleção é composta de 20 moedas comemorativas, 14 placas, 12 certificados, 12 troféus, 8 pratos, 7 bottons, 5 brasões, 4 quadros, 3 chaveiros, 3 medalhas, 2 abridores de carta, 2 bandejas comemorativas, 2 broches, 2 copos, 2 fotos, um azulejo, uma bandeira, um envelope, uma escultura, uma flâmula, um suporte para prato, uma tapeçaria, um vaso decorativo, e mais 9 objetos decorativos variados.

METODOLOGIA

A pesquisa (UFSCAR, 2017) foi realizada de acordo com as seguintes etapas:

- a) Reunião do acervo e definição das coleções compostas pelas séries de objetos, tais como placas, troféus, flâmulas, copos e outros;
- b) Inventariação dos objetos à maneira dos museológicos, com descrição física, conteúdos informacionais e registros fotográficos;
- c) Composição de um catálogo digital, contendo as informações textuais e fotográficas, em acordo com a Associação Brasileira de Normas Técnicas (ABNT);
- d) Seleção de um subconjunto de mimos ligados a momentos marcantes da história da universidade, seguida de análise dos significados contidos nesses mimos.

Esta pesquisa gerou resultados de diversas formas. Entre elas, podemos destacar que foi feita uma reunião de conceitos teóricos para a construção de um olhar para o objeto. Além disso, a pesquisa resultou na construção de fichas de descrição, bem como na análise dos significados dos mimos.

O processo de descrição visou a obtenção e o registro de um conjunto de informações a respeito dos objetos, entre elas o tipo de objeto, título, autoria, descrição intrínseca e extrínseca, quando couber, dimensões, tipo do material, eventuais observações e data, quando informada (Tabela 1.). Em síntese, esse processo de descrição foi baseado no seguinte modelo:

Tabela 1. Modelo da Ficha de Descrição

ITEM	DESCRIÇÃO
Objeto:	
Título:	
Autor/Autoridade:	
Descrição Intrínseca:	
Descrição Extrínseca:	
Dimensões:	
Material:	
Observação:	
Data:	

Referência: Prado (31/08/2016)

É importante ressaltar que, dependendo do tipo de objeto analisado e de suas condições, foi possível obter maior ou menor quantidade e qualidade de informações para descrição. Como exemplo de ficha obtida, apresentamos a ficha (Tabela 2.) a seguir:

Tabela 2. Exemplo da Ficha de Descrição

ITEM	DESCRIÇÃO
Objeto:	Placa (PL)
Título:	Associação Pró Casa do Pinhal
Autor/Autoridade:	Casa do Pinhal/São Carlos-SP
Descrição Intrínseca:	
Descrição Extrínseca:	A placa está dentro de um suporte de veludo. Dimensões da caixa: 15,7 cm

	comprimento x 10,8 cm largura e 1,4 cm espessura.
Dimensões:	(Placa) 14 cm comprimento x 9 cm de largura
Material:	Metal
Observação:	Homenagem da Fazenda Conde do Pinhal localizada no município de São Carlos/SP à UFSCar.
Data:	07/06/2008

Referência: Prado (31/08/2016)

RESULTADOS

No decorrer da pesquisa (UFSCAR, 2017), alguns objetos se destacaram pelos significados e relações que representam para a história da UFSCar. Alguns desses objetos contêm informações escritas sobre premiações, reconhecimentos, alianças e agradecimentos, referentes tanto à figura do reitor quanto à universidade como instituição. Nesses casos, geralmente verificados na forma de certificados e placas, foi possível estabelecer os significados de cada mimo com clareza a partir dos objetos em si.

Em outros casos, (UFSCAR, 2017) os mimos trazem informações mais breves, identificando apenas os nomes das instituições com as quais a UFSCar estabeleceu relações, e eventuais datas comemorativas. Esse grupo de objetos, do qual fazem parte diversas moedas e medalhas comemorativas, e objetos afins, por um lado é mais restrito no que se refere às informações apresentadas nos objetos, por outro possibilita a demarcação de significados de uma forma indireta.

Alguns objetos (UFSCAR, 2017) demarcam a presença de relações internacionais estabelecidas ao longo da história da UFSCar (Figura 1.). São exemplos que se destacam nesse grupo: o prato da Universidade Nacional de Entre Rios (Argentina), o prato da Universidade de Aveiro (Portugal), a moeda da Universidade do Minho (Portugal), o abridor de cartas da Università degli Studi di Padova (Itália), a faixa do Timor Leste e a placa da Associação de Universidades do Grupo Montevideo (AUGM). Esses mimos são evidências da inserção da UFSCar no cenário internacional, a partir de eventuais alianças, cooperações, intercâmbios e parcerias com universidades, grupos, associações e estudantes de outros países. Isso nos mostra que as relações de troca superam os limites das fronteiras nacionais.



Figura 1. Mimos Internacionais

As moedas da Pontifícia Universidade Católica de São Paulo (PUC-SP), referente à comemoração dos cinquenta anos dessa instituição (1946-1996), e da Universidade Federal do Rio Grande do Sul (UFRGS) (Figura 2.), comemorativa do centenário dos cursos de Farmácia e Engenharia, são exemplos de relações entre a UFSCar e outras universidades brasileiras. Estes são mimos que visam o cultivo da memória das histórias dessas instituições.



Figura 2. Moeda PUC-SP (à esquerda e ao centro) e Moeda UFRGS (à direita)

CONCLUSÕES

Neste trabalho previamente realizamos um estudo teórico. De forma interdisciplinar trabalhamos com vários enfoques e perspectivas teóricas a fim de olharmos esta coleção de várias maneiras. As contribuições das Ciências Sociais, da História, dos Estudos Culturais, da Ciência da Informação e da Documentação podem ser consideradas de grande importância.

O modelo de ficha de descrição adotado foi adequado para a descrição da coleção, podendo dessa forma obter o maior número de informação que objeto trás, suas características, suas peculiaridades, além de suas condições de conservação. Por meio da descrição dos objetos percebemos que há uma importância de que preservar memórias por meio de coleções de objetos é fundamental, além disso, é preciso realizar um registro cuidadoso da história de cada item, assim como de sua conservação ao longo do tempo.

Por fim, foi possível identificar alguns significados que se destacaram entre os objetos da coleção da reitoria da UFSCar. Alguns objetos representam relações nacionais, ou seja, vínculos entre a UFSCar e outras universidades brasileiras. Esses vínculos entre universidades, expressos por meio dos objetos, podem ser interpretados como relações diplomáticas travadas entre instituições da mesma natureza, ou seja, atuantes em um mesmo setor da sociedade, instituições voltada para o ensino, pesquisa e extensão. Por essa razão, é natural que essas instituições estabeleçam parcerias e cooperações, e celebrem essas parcerias por meio de presentes que simbolizam momentos de sua história, alianças ou datas comemorativas.

Do mesmo modo, (UFSCAR, 2017) alguns itens da coleção representam relações internacionais, da UFSCar com universidades de outros países, ou mesmo com associações de pesquisa internacionais. De forma semelhante, essas instituições compartilham o setor social no qual atuam, ou seja, compartilham metas, objetivos, formas de atuação e eventualmente recursos humanos. Associações, cooperações, intercâmbios e articulações de pesquisa são exemplos dos vínculos possíveis entre elas. Nesse sentido, a troca de objetos entre os agentes institucionais pode ser vista como um ritual social que reforça os vínculos institucionais estabelecidos.

De modo geral, todos esses mimos podem ser lidos como elementos culturais, dotados de materialidade e, ao mesmo tempo, de significados ou simbologias. A perspectiva de cultura de Raymond Williams, segundo a qual o material e o simbólico são dimensões complementares dos objetos, é adequada para essa pesquisa.

Por fim, podemos dizer que esse processo de descrição dos objetos e identificação dos significados nos fornece alguns elementos interessantes que fazem parte da Memória Institucional da UFSCar

REFERÊNCIAS BIBLIOGRÁFICAS

Bachelard, Gaston. (1993). *A poética do espaço*. São Paulo, Martins Fontes.

Cevasco, Maria Elisa. (2012). *Dez lições sobre estudos culturais*. São Paulo: Boitempo Editorial.

Costa, Icléia Thiesen Magalhães. (1997) *Memória Institucional: a construção conceitual numa abordagem teórico-metodológica*. Tese apresentada ao Programa de Pós-Graduação em Ciência da Informação da UFRJ. Rio de Janeiro.

Crippa, Giullia & Mostafa, Solange Puntel (org.) (2011). *Ciência da informação e documentação*. Campinas : Editora Alínea.

Litton, Gaston. (1976). *A documentação*. São Paulo: McGraw-Hill do Brasil.

Mauss, Marcel. (2003) Ensaio sobre a dádiva. In: _____. *Sociologia e Antropologia*. São Paulo: Cosac Naify.

Molina, Letícia Gorri, & Valentin, Marta Lúcia Pomim. (2011). Memória organizacional, memória corporativa e memória institucional: discussões conceituais e terminológicas. *Revista EDICIC*, v.1, n.1, p.262-276, Jan./Mar.

OTLET, Paul. Documentos e documentação. In: CONGRESSO MUNDIAL DA DOCUMENTAÇÃO UNIVERSAL, 1937, Paris. Disponível em: <<http://www.conexaorio.com/bit/otlet/>>. Acesso em: 16. mai. 2007.

Simmel, Georg. (1983). *Sociologia: Grandes Cientistas Sociais*. São Paulo: Ática.

Simmel, Georg. (2004). *Fidelidade e Gratidão e Outros Textos*. Lisboa : Relógio D'água Editores.

Soares, A. L. R (org,) (2007). *Educação patrimonial: teoria e prática*. Santa Maria: Ed. UFSM.

Universidade Federal de São Carlos. (2017). *Congresso de Iniciação Científica e Tecnológica*. Recuperado em 21 fevereiro, 2017 de <http://eventweb.com.br/cict2016/home-event/schedule.php?area=1182>

Williams, Raymond. (2007) *Palavras-Chave: um vocabulário de cultura e sociedade*. São Paulo: Boitempo Editorial.



ANÁLISE DO TERMO COMPOSTO NA INDEXAÇÃO DE UMA COLEÇÃO DE POSTAIS ILUSTRADOS (1940-1960) À LUZ DA NF Z 47-200 (1985)

Maria da Graça de Melo Simões¹, María Begoña López-Ávila², Blanca Rodríguez-Bravo³,
Maria de Fátima Moura Carvalho⁴, Giovana Deliberati Maimone⁵

¹*Departamento de Filosofia, Comunicação e Informação, Universidade de Coimbra, Portugal,
0000-0001-5575-0261, gsimoies@fl.uc.pt*

²*Departamento de Información y Comunicación, Universidad de Granada, España,
0000-0002-5740-1599, begolopez@ugr.es*

³*Departamento de Patrimonio Artístico y Documental, Universidad de León, España,
0000-0002-9476-7602, blanca.rodriguez@unileon.es*

⁴*Biblioteca Geral da Universidade de Coimbra, Portugal, 0000-0002-4610-1822, fmoura@bg.uc.pt*

⁵*Departamento de Biblioteconomia e Documentação da Universidade de São Paulo, Brasil,
0000-0002-4150-8084, gdmaimone@usp.br*

RESUMO As particularidades da indexação da imagem assentam na dicotomia: real e denotativo versus imaginário e conotativo, características que contribuem para a complexidade deste processo. À subjetividade que caracteriza a indexação acrescem os pormenores explícitos e implícitos da imagem. Assim, a aplicação de uma norma e o recurso a uma terminologia específica são imprescindíveis. Entre os termos para descrever uma imagem, salienta-se o composto para representar a especificidade. Propõe-se contribuir para o estudo do termo composto na indexação de imagens, através da análise do seu uso na indexação de uma coleção de postais ilustrados (1940-1960), comparando essa prática com as formas consideradas na NF Z 47-200 (1985). Partiu-se de uma abordagem qualitativa exploratória, baseada na revisão da literatura e num estudo empírico. Para a análise dos dados, recorreu-se às técnicas de análise de conteúdo e da análise comparativa, dado a relevância que assumem na interpretação dos dados, em contexto qualitativo. Os resultados mostram que o uso do termo composto coincide, de um modo geral, com os princípios teórico-metodológicos recomendados pela NF Z 47-200 (1985), revelam ainda a incoerência e inconsistência na aplicação do termo composto quanto à sua estrutura. Infere-se que o termo composto deve ser utilizado na indexação de imagens, particularmente nos postais ilustrados, dado permitir representar o específico. Concluiu-se haver necessidade da aplicação de normas, para harmonizarem a representação das especificidades deste recurso informativo.

PALAVRAS-CHAVE *Termo composto, NF Z 47-200 (1985), postais ilustrados, indexação.*

ABSTRACT The particularities of the indexation of images are based on the dichotomy: real and denotative versus imaginary and connotative, which are the characteristics that contribute to the complexity of this process. The subjectivity that characterizes the indexation adds the explicit and implicit details of the image. Thus, the application of a standard and the use of specific terminology are essential. Amongst the terms to describe an image, the multi-word term stands out to represent the specificity. It is proposed to contribute to the study of the multi-word term in the indexation of images, through the analysis of its use in indexing a collection of illustrated postcards (1940-1960), comparing this practice with the forms considered in NF Z 47-200 (1985). It was based on an exploratory qualitative method, based on literature review and an empirical research. For the analysis of the data, we used the techniques of content analysis and comparative analysis, due to the relevance that assume in the interpretation of the data, in qualitative context. The results

show that the use of the multi-word term, in general, coincides with the theoretical and methodological principles recommended by NF Z 47-200 (1985), also reveal the incoherence and inconsistency in the application of the multi-word term in its structure. It is inferred that the multi-word term must be used in the indexation of images, particularly in illustrated postcards, since it allows to represent the specific. It was concluded that it is necessary to apply the norms to harmonize the representation of the specificities of this information resource.

KEYWORDS *Multi-word term, NF Z 47-200 (1985), illustrated postcards, indexation.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A imagem foi desde sempre um dos principais veículos de comunicação na cultura da humanidade, embora, em alguns períodos, a comunicação escrita se tenha sobreposto (Rodrigues, 2007). O acesso ao conteúdo dos documentos visuais e o processo de descrever imagens é complexo e merece especial consideração Maillat (1991). A complexidade inicia-se na análise das imagens estendendo-se depois à sua representação. Na imagem encontra-se presente uma dicotomia que se traduz no real e denotativo versus imaginário e conotativo. O denotativo é o referente (a imagem em si, o concreto) enquanto o conotativo é a leitura que se faz da imagem, o abstrato (Valle, 1999). Enquanto a imagem, considerada no aspeto denotativo se mantém autónoma e inalterável, porém quando é considerada sob o aspeto conotativo são-lhe atribuídas numerosas apropriações semânticas, que concorrem para diversos e distintos significados específicos, próprios do contexto em que se integra. A indexação de imagens assume-se como uma moeda, com duas faces. Uma relativa ao referente, outra correspondente à interpretação. Para esta última dimensão, contribui o contexto, a própria técnica usada na sua produção e os elementos implícitos da imagem que não são visíveis por parte de quem a observa (Lopes, 2006). Na fotografia, Manini (2002), designa a este processo dimensão expressiva. Sendo estas duas dimensões duas faces da mesma moeda, elas sobrepõem-se, podendo confluír no momento da recuperação da informação para um conflito difícil de conciliar, entre a linguagem daqueles que a indexam e a dos que recuperam a informação. A subjetividade de uns pode entrar em colisão com a objetividade dos outros (Sánchez, 2006). É nesta simbiose polissémica que radica a complexidade da indexação da imagem, independentemente da sua manifestação física (fotografia, iconografia, postal, etc.). As imagens raramente apresentam informação escrita, excetuando-se, eventualmente o registo de uma legenda que, em determinadas situações, não é atribuída pelo autor. De acordo com (Manini, Lima-Marques & Miranda, 2007) a legenda representa uma breve informação que deve ser considerada na sua indexação. A informação facultada pelas legendas é geralmente ampla de sentido, e, em alguns casos, não descreve com rigor a imagem, o que concorre para uma indexação geral e não específica, como seria o desejável. Neste sentido, Lopes (2006) ao referir-se à indexação de fotografia menciona que esta deve considerar os objetivos institucionais e as necessidades de informação de quem a procura, em particular, no que se refere aos níveis de exaustividade e especificidade. Dadas as características diferenciadas da imagem Smit (1996), refere que não se pode adaptar ao tratamento da imagem, os mesmos procedimentos que se aplicam à indexação do texto impresso. A imagem faculta um conjunto de pormenores que são menos evidentes e, menos perceptíveis quando comparados aos dos textos impressos, situação que contribui para um nível considerável de subjetividade, em detrimento da dimensão denotativa. No mundo da imagem, destaca-se o postal ilustrado, incrementado por Emanuel Herrmann em 1869, como meio de comunicação. O seu sucesso concorreu para que fosse adotado em outros países, entre os quais:

Alemanha, Espanha, Estados Unidos da América, Japão e Portugal (Guereña, 2005). Os primeiros postais foram feitos por fototipia, depois, também, em papel fotográfico (Palá, 2004). Na fotografia encontram um aliado, vindo a favorecer um comércio postal à escala nacional e internacional, potenciado pelos turistas. Eles divulgam, cidades, cenas típicas e aspetos culturais de distintos países. Segundo López (2013), o cartão postal consagrou-se como meio de difusão da imagem de culturas, cidades e monumentos; considerando-se como um objeto de valor cultural e patrimonial que simboliza a sua época, continuando a usar-se atualmente, como meio de comunicação postal, apesar dos novos sistemas.

O universo controverso e paradoxal que é o mundo da imagem, apenas ganha importância para a construção de um novo conhecimento quando é organizado para que a ele se possa aceder e recuperar a informação. No seu tratamento técnico ganha significativa importância a indexação. Lopes (2006) refere que a indexação acrescenta um valor informativo e documental numa base de dados ou num banco de imagens. A necessidade premente da representação e acesso por assunto a este valioso recurso de informação, como manifestação e preservação do património e da memória cultural, levou a que estes recursos informativos fossem indexados através de linguagens documentais que representassem o mais próximo possível a realidade neles expressa, para tal recorreu-se às linguagens documentais controladas. As linguagens documentais foram desde sempre um meio imprescindível para a representação e recuperação da informação. Gil Urdiciain (2004) refere que as linguagens documentais atuam em dois segmentos do processo documental: na representação da informação e na sua recuperação. Na literatura, é frequente encontrar-se a definição de linguagem controlada em oposição à definição de linguagens livres. De acordo com Lancaster (2004) entende-se por linguagem livre as palavras ou expressões que se encontram nos textos, enquanto a linguagem controlada é artificial (Courrier, 1976; Gil Urdiciain, 2004) derivando da linguagem natural (Sayão, 1996). Na recuperação da informação, este tipo de linguagem assume-se como mediador privilegiado entre a linguagem natural do utilizador e a do sistema de recuperação de informação que usa uma linguagem controlada. A sua função é controlar a terminologia de um domínio do conhecimento, com o fim de diminuir a ambiguidade concetual, de modo a organizar e a recuperar a informação de um modo pertinente e preciso (Rodríguez-Bravo, 2011). A linguagem controlada é constituída por termos autorizados: simples e compostos (NF Z 47-200, 1985; ANSI/NISO Z39.19-2005, rev. 2010) os quais foram sujeitos a um controlo, que se efetua a três dimensões: morfológico (género e número), semântico (controlo da sinonímia e polissemia) e sintático (composição dos termos compostos) (Simões, 2008). Das relações que se estabelecem entre os termos salientam-se as relações semânticas (Lancaster, 2004). Embora, as Normas de construção de vocabulários controlados preconizarem o uso do termo simples e sempre que possível a decomposição sintática e morfológica do termo composto (NF Z 47-200, 1985; ANSI/NISO Z39.19-2005, rev. 2010; ISO 25964-1, 2011), porém, em alguns casos, admitem o recurso ao termo composto (NF Z 47-200, 1985, ponto 1.4; ANSI/NISO Z39.19-2005, rev. 2010, ponto 7.3; ISO 25964-1, 2011, ponto 7.3). Tal circunstância prende-se com o facto de, em determinados casos, o termo simples carregar consigo uma carga significativa de ambiguidade semântica, que concorre para falsos resultados na pesquisa. Neste sentido, a decomposição morfológica ou semântica do termo composto, poderá eliminar o conceito que é representado por um grupo de palavras (Gomes & Campos, 2004). No esforço natural de aproximação das linguagens controladas às linguagens naturais, de modo a completarem-se entre si (Muddamalle, 1988), o termo composto, em particular o que se apresenta pela ordem direta, ganhou uma maior relevância. Para esta circunstância contribui, entre outros fatores, a autonomia nas pesquisas por parte do utilizador alavancadas nas Tecnologias de informação e comunicação (TIC). O recurso ao termo composto forma invertida, prática privilegiada nas listas de encabeçamentos de matérias, começou a dar lugar ao termo composto ordem direta. Esta alteração de práticas teve a ver

com a adoção, por parte das listas de encabeçamentos de matérias, de algumas orientações metodológicas relativas à elaboração dos tesouros que postula o uso do termo composto ordem direta, não admitindo o uso do termo invertido (recurso à vírgula ou ao hífen, para relacionar os termos). A disseminação e aceitação generalizada do tesouro, associada à autonomia na pesquisa por parte do utilizador, veio consolidar o uso deste tipo de termo quer na representação quer na recuperação da informação. Assim, a prática do termo composto pela ordem direta, que emergiu de modo discreto, na indexação em ambientes informacionais tradicionais, veio a consolidar-se na indexação em ambientes informacionais informatizados, em particular no que se refere aos ambientes de inteligência artificial, sobretudo na indexação social. A necessidade de representar assuntos compostos e/ou complexos, de desambiguar o sentido de alguns termos, de precisar conceitos, aliados ao facto de se aproximar da linguagem natural, o termo composto por ordem direta, são motivos para justificar a sua plena aplicação (Simões, 2008). Por isso, a recomendação ao seu uso aparece regulada num conjunto significativo de normativas nacionais e internacionais, que orientam as práticas da indexação e, das quais se destacam as já mencionadas, assim como em textos teórico-metodológicos, entre as quais se salienta a obra seminal *Rules for a dictionary catalog* (1876) de Charles Cutter, na qual já se insinua o seu uso, e a partir da qual foi elaborada, a *Library of Congress Subject Headings* (LCSH). Com o intuito de contribuir para o estudo do termo composto na indexação de imagens, propõe-se, com este estudo, identificar e analisar o uso do termo composto na indexação de uma coleção de postais ilustrados datados entre 1940-1960, da Biblioteca Geral da Universidade de Coimbra, comparando essa prática com as formas consideradas na NF Z 47-200 (1985) relativo a este tipo de termo. Constituíram-se como objetivos específicos: identificar e selecionar os termos compostos usados na coleção considerada; analisar os princípios que a NF Z 47-200 (1985) apresenta para o uso do termo composto (pontos 1, 2, 3 e 4); comparar as formas usadas no caso empírico e as orientações referidas na dita Norma, para o seu uso.

METODOLOGIA

Para a concretização dos objetivos optou-se por uma abordagem metodológica qualitativa exploratória, com o propósito de se familiarizar com a questão central do estudo salientando-a enquanto problema real Gil (2002). Esta abordagem assentou em dois segmentos metodológicos, o primeiro, na revisão da literatura e, o segundo no estudo empírico. Para a análise dos dados que resultaram da recolha, recorreu-se às técnicas de análise de conteúdo e da análise comparativa, dado a relevância que estas técnicas metodológicas assumem na interpretação dos dados, em contexto qualitativo. A revisão de literatura apoiou-se em artigos científicos e normas relativas ao controlo da linguagem. Esta teve como propósito apresentar e refletir sobre a indexação de imagens e de termo composto, assim como de alavancar teoricamente o estudo empírico. O estudo empírico consistiu na recolha e análise do corpus do estudo (termos compostos), à luz das orientações da NF Z 47-200 (1985). Numa primeira fase procedeu-se à caracterização da coleção de postais ilustrados, que constituíram o objeto de estudo, com base num estudo prévio (Mendes, 1974), assim como da Norma NF Z 47-200 (1985), com particular incidência para os pontos: 2, 3 e 4. Seguiu-se a determinação da amostra (349) postais num universo de 688.

Os critérios da recolha de dados basearam-se nos pontos (1.4; 1.5) da NF Z 47-200 (1985) que se referem à existência de termos compostos numa lista de encabeçamentos de matérias e, o ponto 2, que refere o uso de termos compostos por ordem direta e/ou invertida. Para a recolha dos dados foram construídas tabelas e os gráficos respetivos, nos quais constam as percentagens e as suas modalidades de aplicação.

Seguiu-se a definição das categorias de análise e os respetivos critérios de seleção. Para a análise dos termos compostos, partiu-se de duas categorias de análise: estrutura e conteúdo (NF Z 47-200, 1985). Quanto à estrutura analisaram-se os termos quanto à sua composição, por ordem direta (termo prepositivo, adjetivo ou outro) e por ordem invertida, tendo estes últimos sido analisados de acordo com a pontuação e construção sintática (ponto 4 da Norma). Quanto ao conteúdo, com base no ponto 1.5 e nos pontos 2 e 3, os termos foram analisados de acordo com o assunto que representavam, tendo sido classificados em: onomásticos, geográficos, temáticos, cronológicos e formais. Para a recolha destes elementos foram construídas tabelas similares à usada na apresentação dos termos compostos. Elaboram-se tabelas para recolher dos dados relativos à Estrutura (composição) e ao Conteúdo.

APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

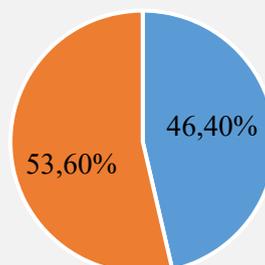
Neste ponto apresentam-se os resultados e a respetiva discussão, articulando-se a os resultados obtidos no estudo empírico e a revisão da literatura. Inicia-se esta apresentação pela constituição dos termos compostos por ordem direta.

1. APRESENTAÇÃO DOS TERMOS COMPOSTOS

Os termos compostos numa lista de encabeçamentos de matérias assumem dois tipos de apresentação: por ordem direta e por ordem invertida (pontos; 2 e 3), tal como se pode observar no quadro que se segue, relativamente ao levantamento efetuado no catálogo de assuntos da coleção de postais ilustrados (Tabela 1).

Tabela 1: Apresentação dos termos compostos

Tipologia	Exemplo	Número	Percentagem
Termo composto ordem direta	Casa típica	162	46,4%
Termo composto ordem invertida	Figo - colheita	187	53,6%
Total		349	100%



■ Ordem direta ■ Ordem inversa

Gráfico 1. Percentagem dos termos compostos

O gráfico 1 mostra que os termos compostos por ordem inversa, na amostra considerada, representam 53,6%, sendo mais de metade dos termos em análise. Apresentam-se com exemplos: *Figo – colheita*, *Costumes – roupa a secar* e *Corgo – seca do milho*. Relativamente aos termos compostos ordem direta

assumem o valor percentual de 46,40%. Referem-se como exemplos: *Casa típica*, *Danças regionais* e *Tipos indígenas*.

1.1. CATEGORIAS DE ANÁLISE E RESPECTIVOS CRITÉRIOS DE SELEÇÃO

Para a análise dos termos compostos, com base na Norma, partiu-se de duas variáveis: a sua estrutura e o seu conteúdo. Quanto à estrutura analisaram-se os termos quanto à sua composição, os que se encontram por ordem direta (termo prepositivo, adjetivo ou outro) e os que se encontram por ordem invertida analisaram-se de acordo com a pontuação e construção sintática (ponto 4).

Quanto ao conteúdo ainda com base na dita Norma e em particular com o ponto 1.5 e o desenvolvimento do ponto 2 e 3, foram analisados de acordo com o assunto que representam, no caso foram classificados em: onomásticos, geográficos, temáticos, cronológicos e formais.

1.1.1. Estrutura

- a) Composição dos termos por ordem direta: no que se refere à estrutura dos termos por ordem direta na sua constituição eles podem recorrer a uma preposição, a um adjetivo ou a uma partícula de ligação (ponto 2.2).

Tabela 2: Constituição dos termos compostos por ordem direta

Tipologia	Exemplo	Número	Percentagem
Termo composto prepositivo	Lagar do vinho	84	52%
Termo composto adjetivo	Trajes regionais	59	36,4%
Outro	Capote e capelo	19	11,6%
Total		162	100%

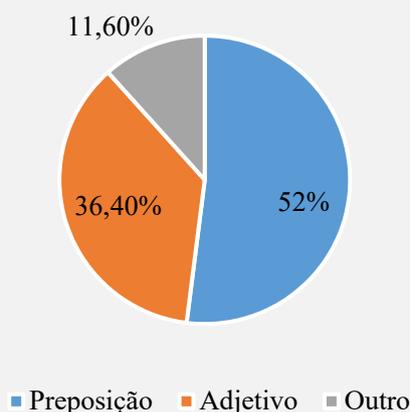


Gráfico 2. Percentagem dos termos compostos por ordem direta

O gráfico 2 apresenta o termo composto prepositivo com 52%, tal como se pode observar através dos exemplos *Amendoeiras em flor*, *Barcos de pescadores*, *Ceifa do arroz* ou *Lagar do vinho* (Tabela 2). O termo composto adjetivo assume um valor percentual de 36,40%, tal como demonstram os exemplos: *Amendoeiras floridas*, *Casa rústica* ou *Trajes regionais*. Com menor frequência utilizam-se outros tipos de forma de termos como a copulativa, dos quais se apresenta o exemplo *Capote e capelo* (Tabela 2).

- a) Relativamente à composição dos termos por ordem invertida, recorre-se a duas vias: pontuação (ponto 4.1) e à construção (ponto 4.2). Por esta circunstância, e dado relevo que este tipo de termos assume na lista de encabeçamentos de matérias, procedeu-se ao levantamento deste tipo de termos para uma análise *posterior*.

Tabela 3: Composição dos termos por ordem invertida

Tabela 3.1: Pontuação

Tipologia	Exemplo	Número	Percentagem
Hífen	ANGOLA – Tipos indígenas – Noivas	170	85%
Vírgula	ANTÓNIO, Santo; ALMOUROL, castelo	14	7%
Ponto	-----	0	0%
Parenteses	ASTAIRE (Fred) Actor	16	8%
Total		200 ¹	100%

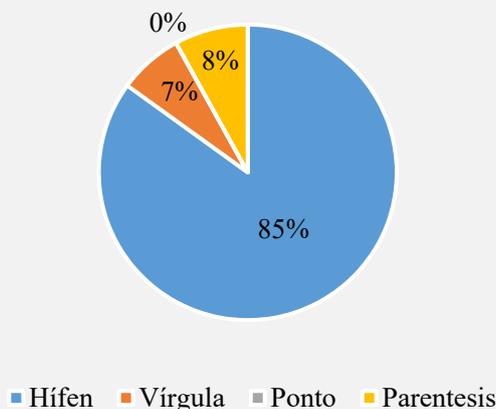


Gráfico 3. Percentagem da pontuação

De acordo com o gráfico 3, em 85% dos casos recorrem ao hífen. Em nenhum termo foi usado o ponto. Outras formas do pontuação utilizadas na composição dos termos por ordem invertida são os parêntesis 8%, como por exemplo: *ALBERGARIA (Lopo Soares d')*, *ALCOBAÇA (Album)* e *ASTAIRE (Fred) Actor*. O uso dos parêntesis oferecem uma fixação do significado dos termos (qualificador), como é o caso de um nome de um personagem ou de uma tipologia documental. Os termos compostos separadaos por vírgula assumem um valor percentual de 7%, como se observa nos exemplos: *ALGARVE, Costumes* ou *ALMOUROL, Castelo* (Tabela 3.1).

Tabela 3.2: Construção (ordem de registo dos elementos)

Disposição dos elementos	Exemplo	Número	Percentagem
Geográfico+ temático	ALCÁCER DO SAL-Vindima	207	66,8%
Geográfico+ temático+ temático	AÇORES-Trajos regionais-capote e capelo	88	28,4%

¹ Refere-se que o número é superior a 187 (v. Tabela 1) dado o facto de alguns termos apresentarem mais do que um tipo de pontuação

Principal+ temático	ANJO-Vasco Fernandes	6	2%
Principal+ temático+ temático	ALA ARRIBA – Filme – “Tia Saramago”	8	2,5%
Principal+ temático+ geográfico	ANJO-aparição-Fátima	1	0,3%
Total		310	100%

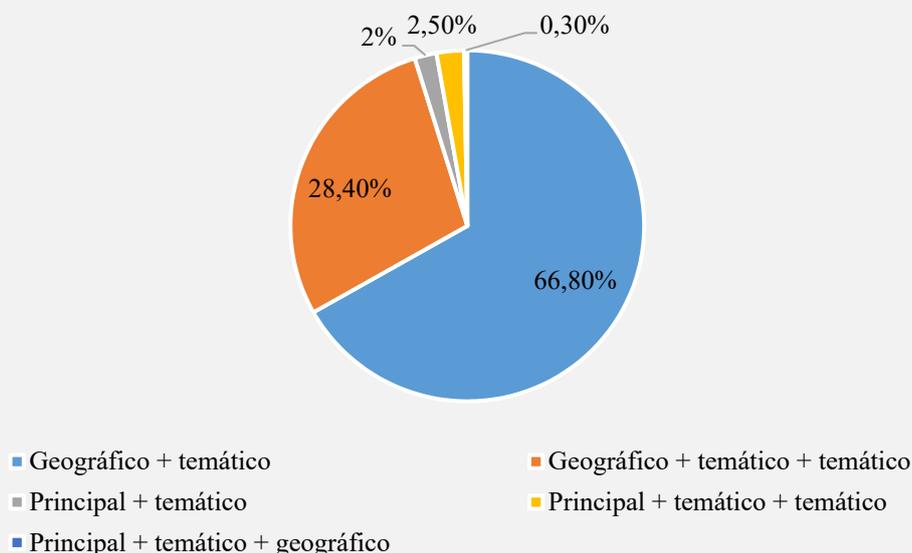


Gráfico 4. Ordem dos elementos

No gráfico 4, observou-se que a disposição de termos com maior percentagem é o geográfico, seguido do temático, 66,80% dos casos, como se observa nos exemplos: *ALBUFEIRA – Mercado*, *ALCÁCER DO SAL-Vindima* ou *ALVOR – Praia*, apenas um apresenta o geográfico em último lugar (0,30%) como *ANJO-aparição-Fátima*. Com menor frequência, o temático apresenta-se na disposição do termo composto em primeiro lugar em 2% dos casos, como pode observar-se no exemplo: *ANJO-Vasco Fernandes*.

1.1.2. CONTEÚDO:

Quanto ao conteúdo e de acordo com os pontos 1.5 e 2.2 da NF Z 47-200 (1985), (quando se refere às subdivisões) este apresenta uma tipologia diversificada, a qual se apresenta no quadro. Justifica o seu levantamento o facto de ser através dela que se pode observar as temáticas e a percentagem que registam nesta coleção.

Tabela 4: Classificação dos termos quanto ao conteúdo

Tipologia	Exemplo	Número	Percentagem
Onomásticos	Afonso Henriques (D.) – Rei de Portugal	37	10%
Geográficos	África portuguesa	315	85,1%
Temáticos	Barcos de pescadores	17	4,6%
Cronológicos	Aguardente – 1920	1	0,3%
Formais	-----	0	0%
Total		370	100%

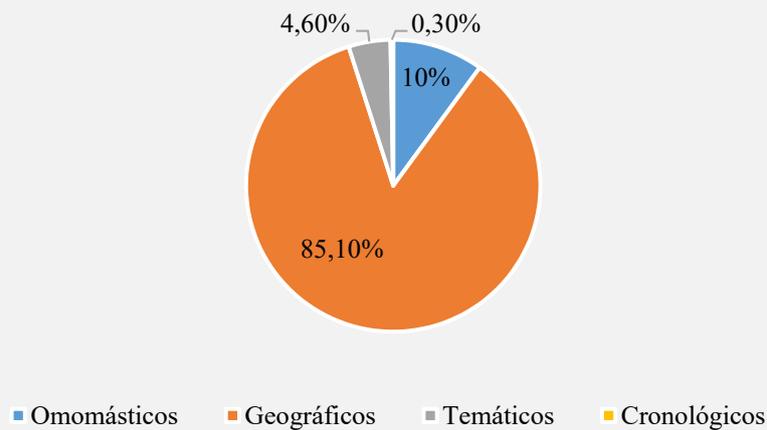


Gráfico 5. Termos-conteúdo

Quanto à nomenclatura dos termos compostos, com base no conteúdo, o gráfico 5 mostra que o geográfico é predominante (85,10%), dado o facto de os postais ilustrados se referirem a lugares. Apenas 10% recorrem ao onomásticos, exemplo: *Afonso Henriques (D.) – Rei de Portugal*. O recurso ao termo composto temático assume 4,60%. Relativamente ao cronológico apenas existe um termo (0,30%). Salienta-se que não se observa nenhum formal.

CONCLUSÕES

A partir dos resultados, concluiu-se que o uso do termo composto na indexação de postais ilustrados se encontra, de um modo geral, de acordo com os princípios teórico-metodológicos recomendados pela NF Z 47-200 (1985). O termo composto ordem invertida aplica-se com maior frequência do que o termo composto ordem direta, prática muito próxima de uma lista de encabeçamentos de matérias. Também se concluiu alguma incoerência e inconsistência na aplicação do termo composto quanto à estrutura, isto é: há conceitos similares que são representados tanto por um termo composto prepositivo como por um termo adjetivo, o mesmo acontece no caso de conceitos similares serem representados, indiferentemente, por termos compostos ordem direta e por termos compostos ordem invertida. A mesma incoerência e inconsistência observa-se no que se relaciona com a pontuação e com a sua construção.

O uso do termo composto é um recurso utilizado na indexação de imagens, em particular nos postais ilustrados, independentemente da sua forma: invertida ou por ordem direta, pelo facto deste tipo de termos permitir traduzir o específico, que é muito significativo nos postais, dado que representam pormenores. A aplicação da NF Z 47-200 (1985), de uma forma geral, é evidente na indexação dos postais ilustrados, ponto que se entende muito positivo, dado o facto de se caracterizarem por um conjunto diverso de particularidades, que deverá ser harmonizado com o recurso a normativas.

Considera-se esta investigação original e pioneira em Portugal, abrindo portas à reflexão, investigação e sensibilizando os profissionais para a indexação deste tipo de recurso de informação, pois são um meio indiscutível de difusão e preservação do património cultural. Este estudo será um contributo muito relevante para a área da organização do conhecimento em particular no que se refere ao uso do termo composto na indexação de imagens, no caso de postais ilustrados.

Encontram-se limitações no que respeita ao assunto, na medida em que a letra A (correspondente a duas gavetas [1068 fichas]) na generalidade refere-se a áreas geográficas, pelo que limita a riqueza dos resultados. Dada esta limitação os resultados consideram-se probabilísticos e representativos de uma possível tendência de aplicação do termo composto na indexação de postais ilustrados relativos a locais e a temas com eles associados. Assim, para a confirmação desta ideia, sugerem-se outros trabalhos futuros.

REFERÊNCIAS BIBLIOGRÁFICAS

ANSI/NISO Z39.19-2005 (Rev. 2010). *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*. Bethesda, Maryland: NISO.

Courrier, Y. (1976). Analyse et langage documentaires. *Documentaliste*, 13(5-6), 178-189.

Cutter, C. (1876). *Rules for a dictionary catalog*. Washington: Government Printing Office. Disponível em: <https://archive.org/stream/rulesforadictio06cuttgoog#page/n4/mode/2up/search/compound>.

Gil, A. C. (2002). *Como elaborar projetos de pesquisa*. São Paulo: Atlas.

Gil Urdiciain, B. (2004). *Manual de lenguajes documentales*, 2ª ed. Revisada y ampliada. Gijón: Trea.

Gomes, H. E. & Campos, M. L. de A. (2004). Tesouro e normalização terminológica: o termo como base para intercâmbio de informações. *DataGramaZero - Revista de Ciência da Informação*, 5(6). Disponível em: http://www.brapci.ufpr.br/brapci/_repositorio/2011/11/pdf_98b6a82906_0019456.pdf.

Guereña, J-L. (2005). Imagen y memoria: La tarjeta postal a finales del siglo XIX y principios del siglo XX. *Berceo*, (149), 35-58. Disponível em <https://dialnet.unirioja.es/descarga/articulo/2229424.pdf>.

ISO 25964-1 (2011). *Information and documentation - Thesauri and interoperability with other vocabularies - Part 1: Thesauri for information retrieval*. Geneva: ISO.

Lancaster, F. W. (2004). *Indexação e resumos: teoria e prática*. Brasília: Briquet de Lemos.

Lopes, I. L. (2006). Diretrizes para uma política de indexação de fotografias. Em A. Miranda y E. Simeão (Eds.). *Alfabetização digital e acesso ao conhecimento* (pp. 199-214). Brasília: Universidade de Brasília.

López Hurtado, M. (2013). La tarjeta postal en España: usos y tendencias. *Revista General de Información y Documentación*, 23(2), 437-453. doi: 10.5209/rev_RGID.2013.v23.n2.43136.

Maillet, L. (1991). *Subject control of film and video: a comparison of three methods*. Chicago: American Library Association.

Manini, M. P. (2002). *Análise documentária de fotografias: um referencial de leitura de imagens fotográficas para fins documentários*. São Paulo: Universidade de São Paulo.

Manini, M. P., Lima-Marques, M. & Miranda, A. S. S. (2007). Ontologias: indexação e recuperação de fotografias baseadas na técnica fotográfica e no conteúdo da imagem. *VIII ENANCIB – Encontro*

Nacional de Pesquisa em Ciência da Informação. Salvador (Bahia). Disponível em: <http://www.enancib.ppgci.ufba.br/artigos/GT2--151.pdf>.

Mendes, M. T. P. (1974). Fundos especiais da Biblioteca Geral da Universidade. *Separata do Boletim da Biblioteca da Universidade de Coimbra*, 31.

Muddamalle, M. R. (1998). Natural language versus controlled vocabulary in information retrieval: A case study in soil mechanics. *Journal of the American Society for Information Science*, 49(10), 881-887.

NF Z 47-200 (1985). *Documentación: Liste d'autorité de matières: structure et règles de l'emploi*. Paris: AFNOR.

Palá Laguna, F. (2004). La tarjeta postal ilustrada. Em F. Palá Laguna y W. Rincón García (Eds.), *Los sitios de Zaragoza en la tarjeta ilustrada* (pp. 33-63). Zaragoza: Fundación 2008.

Rodrigues, R. C. (2007). Análise e tematização da imagem fotográfica. *Ciência da Informação*, 36(3), 67-76. doi: 10.1590/S0100-19652007000300008.

Rodríguez-Bravo, B. (2011). *Apuntes sobre representación y organización de la información*. Gijón: Trea.

Sánchez Vigil, J. M. (2006). *El documento fotográfico: Historia, usos y aplicaciones*. Gijón: Trea.

Sayão, L. S. F. (1997). Bases de dados: a metáfora da memória científica. *Ciência da Informação*, 26(3), 314-318. Disponível em: <http://www.brapci.ufpr.br/brapci/v/a/9500>.

Simões, M. da G. (2008). *Da abstração à complexidade formações conceptuais num tesouro*. Coimbra: Almedina, 2008.

Smit, J. W. (1996). A representação da imagem. *Informare: Cadernos do Programa de Pós-Graduação em Ciência da Informação*, 2(2), 28-36.

Valle Gastaminza, F. d. (1999). El análisis documental de la fotografía. Em F. Valle Gastamiza (Ed.), *Manual de documentación fotográfica* (pp.113-132). Madrid: Síntesis.



O LEGADO DO MÉTODO DIPLOMÁTICO E A IDENTIFICAÇÃO DOCUMENTAL: CONTRIBUIÇÕES PARA A ORGANIZAÇÃO DO CONHECIMENTO

Gabrieli Aparecida da Fonseca¹, Sonia Troitinho²

¹Unesp-Marília, 0000-0002-1785-9896; gabrieli.arq@gmail.com

²Unesp-Marília, 0000-0002-7204-3283, smtr@marilia.unesp.br

RESUMO O presente artigo busca refletir a relação entre Diplomática e Identificação Documental e os benefícios que esta resulta para a Organização do Conhecimento. Assim, apresenta-se um histórico das transformações ocorridas nos conceitos de Diplomática e de documento, uma vez que estas ocorreram concomitantemente. Perpassamos pela Diplomática Científica, Histórica, Forense e Contemporânea compreendendo a relação de cada uma delas com a falsificação, de modo que ao avançar de suas transformações essa relação se torna cada vez menor, já que se volta ao documento arquivístico, o qual tem seu enfoque na proveniência e organicidade. Buscou-se também, averiguar quais os aspectos que influenciam a gênese e a contextualização dos documentos arquivísticos na atualidade, considerando a influência das novas tecnologias para a produção documental. Considera-se então, que a Identificação Documental, enquanto função arquivística voltada justamente para lidar com os novos desafios lançados a produção documental, pode encontrar na diplomática uma aliada. Contudo, para a eficácia dessa relação é essencial que a produção documental compreenda as formalidades impostas pela diplomática. Além disso, conclui-se que a Identificação Documental pode potencializar os efeitos das funções arquivísticas que são posteriores a ela, o que é de suma importância para que haja organização e acesso nos acervos arquivísticos.

PALAVRAS-CHAVE *Diplomática, Documento, Arquivística, Identificação Documental, Organização do Conhecimento*

ABSTRACT This article seeks to reflect the relationship between Diplomats and Identification and the benefits that this results for Knowledge Organization. Thus, presents history of the transformations occurred in the concepts of Diplomats and of record, since these occurred concomitantly. We pass through the Scientific, Historical, Forensic and Contemporary Diplomats, understanding the relation of each of them to the falsification, so that in advancing its transformations this relation becomes smaller and smaller, since it returns to the archival document, which has its approaches on provenance and archival bond. It was also sought to ascertain the aspects that influence the genesis and contextualization of archival documents nowadays, considering the influence of new technologies for documentary production. It is considered, therefore, that the Archival Identification, as an archival function turned precisely to deal with the new challenges launched the documentary production, can find in the diplomat an ally. However, for the effectiveness of this relationship it is essential that the documentary production understands the formalities imposed by the diplomats. In addition, it is concluded that the Document Identification can enhance the effects of the archival functions that are later to it, which is of paramount importance for there to be organization and access in the archival collections.

KEYWORDS *Diplomatics, Record, Archival Science, Archival Identification, Knowledge Organization*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Pretende-se aqui discorrer sobre as relações entre Diplomática e a Identificação Documental para a Organização do Conhecimento, considerando que a criação, uso, organização e acesso aos documentos estão sendo cada vez mais influenciados pelos fluxos e articulações da informação na contemporaneidade. Conforme Birger Hjørland (2008, p.86) a Organização do Conhecimento é um campo de estudo que concentra a natureza e a qualidade dos processos e sistemas de organização do conhecimento. Desse modo, a Diplomática e Identificação Documental permeiam esse campo, já que se voltam para a gênese documental.

Mario Barité (2001, p. 42) afirma que a primeira premissa básica da Organização do Conhecimento, é o entendimento do conhecimento como produto e necessidade social, de modo que os sistemas sociais em que nos encontramos funcionam em torno da informação e do conhecimento. Nesse sentido, a Diplomática, enquanto método que reflete a realidade dos sistemas jurídicos e sociais nos quais os documentos são criados, configura-se como um campo de relevância para os estudos de Organização do Conhecimento.

Com essa perspectiva, apresentaremos aqui a Diplomática em suas diversas variações: Científica, Histórica, Forense e Contemporânea, posto cada uma dela exercer parcela de contribuição para os avanços da Arquivística e da Organização do Conhecimento. Contudo, no âmbito da Identificação Documental, a Diplomática Científica e a Contemporânea são as influências mais significativas desse método. Essa influência se deve ao fato das variações da Diplomática serem as que melhor retratam aspectos formais da produção documental e dos tipos documentais, são os objetos de pesquisa da Identificação Documental.

Diante do contexto atual, acelerado pelo emprego de novas tecnologias, a evidenciação dos condicionantes de criação, da constituição dos tipos documentais, dos diversos fatores de sistematização de dados, assim como do uso da informação se faz primordial e a Identificação Documental, principalmente pelo estudo circunstanciado da proveniência que proporciona, se revela uma potente ferramenta de organização do conhecimento – inclusive, por caracterizar-se como uma etapa preliminar às funções arquivísticas, como a classificação e a descrição, essenciais para a organização e acesso ao patrimônio documental.

Assim, a contribuição tanto da Diplomática Científica quanto da Diplomática Arquivística para a Identificação Documental, transcende a questão da verificação da autenticidade e genuinidade dos documentos, de modo que parte em direção da reconstituição de seu contexto de produção, a partir da observação de elementos como a gênese documental (*actio* e *conscriptio*), tradição documental, estrutura formal, caracteres externos e internos, entre outros, a fim de compreender a função e relação dos documentos com o conjunto ao qual pertencem.

Desse modo, a relação da Diplomática com a Identificação Documental volta-se para uma das principais características da Arquivística, a contextualização dos documentos, que está enraizada nos Princípios de Proveniência e da Organicidade.

Sobre a relação do Princípio de Proveniência com a Organização do Conhecimento, Birger Hjørland considera que:

[...] In some cases documents are ordered by the so-called 'principle of provenance' which requires that documents remain in the collections and the order in which they were originally organised. This requires an insight into the organisation that existed when the collection was established. Ordering of documents and knowledge is always based on particular premises, world views, assumptions. Knowledge of these premises is often necessary to obtain a satisfactory return from descriptions of subjects. The necessary degree of interpretation depends upon the extent to which the subject description has anticipated and met the needs of the user. In the principle of provenance only a low degree of anticipation is attempted because the principle makes no attempt to consider the current user's context.[...] (HJORLAND, 1992, p. 185-186).

Dessa forma, à luz da vontade de averiguar quais os aspectos que influenciam a gênese e a contextualização dos documentos arquivísticos na atualidade, bem como tais aspectos podem influenciar na organização e acesso ao patrimônio documental, os elementos que envolvem essa questão serão apresentados e discutidos criticamente no presente trabalho.

METODOLOGIA

O referido artigo caracteriza-se como um recorte de pesquisa realizado a partir da dissertação de mestrado apresentada ao Programa de Pós-graduação em Ciência da Informação da Unesp- Marília, intitulada “Identificação Documental em arquivos pessoais: possibilidades, convergências e desafios”. Trata-se de uma revisão de literatura a respeito da temática abordada, ou seja, trata-se de uma pesquisa bibliográfica explicativa, já que apresenta um histórico das transformações ocorridas nos conceitos de diplomática e de documento. Assim, perpassando pela Diplomática Científica, Histórica, Forense e Contemporânea, buscou-se traçar a relação da Diplomática com a falsificação documental. Permitindo assim, reunir informações que pudessem contribuir para averiguar quais os aspectos que influenciam a gênese e a contextualização dos documentos arquivísticos na atualidade, considerando a influência das novas tecnologias para a produção documental e conseqüentemente, refletir de que maneira a Diplomática pode contribuir para a aplicação da Identificação Documental, assim como para Organização do Conhecimento – adota-se a perspectiva do entendimento da Identificação Documental como um importante instrumento para viabilizar o acesso à pesquisa em arquivos, visto que potencializa a representação do conhecimento, por meio da classificação e a descrição, poupando-se tempo e assim permitindo que o acesso ao patrimônio documental seja mais rápido e eficaz.

RESULTADOS

A Diplomática desenvolveu-se em concomitância ao documento com finalidade probatória, de maneira que suas alterações - as quais foram posteriores a seu momento inicial, na Idade Média - ocorreram em função do mesmo. Contudo, nesse período medieval o objeto de estudo da Diplomática é o documento diplomático, aquele de natureza jurídica, escrito e redigido segundo formas padronizadas. Tendo a Diplomática se desenvolvido e assumido diversas abordagens ao longo do tempo – Diplomática Científica, Diplomática Histórica, Diplomática Jurídica, Diplomática Arquivística - seu objeto também acompanhou tal evolução. Segundo Georges Tessier (1961), o objeto de análise da Diplomática transcende barreiras culturais e temporais, inclusive de suportes, formatos e gêneros.

O alcance da análise diplomática no tempo e no espaço, podendo ser aplicável a documentos das mais diversas localidades e épocas, deve ser problematizado, pois, ainda que documentos jurídicos possuam uma estrutura que os caracteriza, esta pode ser diversa de acordo com a sociedade na qual se encontra, já que os sistemas jurídicos se diferem de cultura para cultura.

Visando minimizar essas variantes nos sistemas jurídicos, as quais incluem diferenças nos tramites e fluxos documentais, decorrentes das especificidades da burocracia e do volume documental produzido, a Diplomática avançou ampliando seu campo, rumo a outras perspectivas – conforme será especificado melhor mais abaixo. Embora tais variações da Diplomática tenham sido insuficientes para sanar as dificuldades impostas pelas diferenças culturais, de certa forma, foram bastante significativas para refletir sobre o conceito de documento de arquivo. Assim, a respeito do documento de arquivo, entende-se que:

El concepto de documento archivístico es más amplio que el de documento diplomático, pues el primero incluye no sólo éste sino también otros que carecen de ese carácter como cartas o peticiones que no tienen por qué engendrar derechos ni obligaciones. (GALENE DÍAZ; GARCÍA RUIPÉREZ, 2003, p. 24).

Ou seja, embora se tratem de conceitos diferentes, o conceito de documento diplomático é abrangido pelo conceito de documento de arquivo, já que todo documento diplomático também será arquivístico. Contudo, o conceito de documento de arquivo é maior e pode abranger até mesmo documentos de cunho informal. Além disso, os princípios que norteiam os documentos arquivísticos também os distinguem claramente dos documentos diplomáticos. Sendo os princípios de proveniência e organicidade os principais deles, até pelo fato de considerarem as relações entre os documentos de mesmo conjunto documental, o que dificulta a aplicação da Diplomática desenvolvida por Jean Mabillon, já que esta analisa os documentos individualmente.

A respeito dos documentos de arquivo, Heloísa Bellotto (2004, p.37) considera que “os documentos de arquivo são produzidos por uma entidade pública ou privada ou por uma família ou pessoa no transcurso das funções que justificam sua existência como tal, guardando essas documentos relações orgânicas entre si”. Contudo, com a ampliação do conceito de documento arquivístico, abriu-se uma enorme variedade de tipos documentais. A maioria dos tipos emergentes é caracterizada como informais, não tradicionais. De modo que documentos como fotografias e realias, que antes não eram considerados como arquivísticos ou com valor jurídico inerente, passam a agregar esta qualidade, conforme aborda Ana Maria de Almeida Camargo (2015) em “Sobre espécies e tipos documentais”, trabalho no qual discute sobre a grande diversidade de linguagens e formatos de documentos no âmbito de arquivos presidenciais, assim como a metodologia empregada para a identificação de documentos e designação de nomes, durante o processo de reconhecimento das unidades documentais.

Da mesma forma que os documentos se distinguem em função de suas características, com a Diplomática ocorre o mesmo, pois tem a necessidade de acompanhar o desenvolvimento do documento. Manuel Romero Tallafigo (1994) apresenta e distingue as diferentes abordagens que a diplomática assumiu ao longo de seu desenvolvimento: Diplomática Científica, Diplomática Histórica, Diplomática Forense e Diplomática Contemporânea.

Como Diplomática Científica, também chamada de Diplomática Clássica ou Diplomática Tradicional, compreende-se a abordagem tradicional desenvolvida a partir dos preceitos traçados por Jean Mabillon. Seu objeto de estudo é a estrutura formal do documento, de forma que consiste na análise das partes

que o constitui. Apesar de seu método ser aplicável a documentos contemporâneos, este fora desenvolvido pensando nos diplomas medievais, o que faz com que alguns itens da estrutura sugerida por tal método não sejam compatíveis com a estrutura dos documentos contemporâneos. De acordo com Bellotto (2002, p.39-40), o método analítico comparativo da Diplomática Científica divide o documento em três partes: protocolo inicial; texto; protocolo final/escatocolo.

A Diplomática Histórica, por sua vez, se caracteriza pela análise da veracidade das informações contidas no documento. Desse modo, possui a finalidade de compreender as condições de produção e origem dos documentos. De acordo com Juan Carlos Galene Díaz e Mariano García Ruipérez essa abordagem histórica da Diplomática foi inicialmente apresentada por Julius Ficker e Teodoro von Sickel, dessa forma:

[...] Ficker considera la Diplomática como auxiliar de la Historia del Derecho, perfilando con nitidez las dos fases principales del documento diplomático: la «actio» y la «conscriptio». A la Diplomática —según él— le interesa, además de las formas, el contenido del documento en función de los datos jurídicos e institucionales que puede aportar (GALENE DÍAZ; GARCÍA RUIPÉREZ, 2003, p.12).

A Diplomática Forense ou Jurídica caracteriza-se antes de qualquer coisa, pela análise da legalidade dos documentos, pois seu conjunto de regras distinguem os documentos legítimos e genuínos dos demais. De acordo com Manuel Romero Tallafígo:

Es el carácter jurídico y forense de la Diplomática, que yo creo que no puede olvidar, aunque esté incardinada dignamente y con magníficos frutos en las Areas de Conocimiento de Historia o de Ciencias y Técnicas Historiográficas. Ese carácter la convierte siempre en actual, en ciência aplicada y técnica jurídica, y la desviste de los tonos medievalistas y positivistas con que más se presenta en las bibliografías y donde, sin duda, há alcanzado sus cotas más famosas (ROMERO TALLAFIGO,1994, p.23).

Dessa forma, é sua relação com o direito e aspectos jurídicos formais que fazem com que a Diplomática Forense perpetue sua permanência no tempo e a distingue da Diplomática Científica.

Diante do surgimento do documento eletrônico, tradicionais formas de análise documental foram reinventadas para atender a novas configurações de registro. Segundo Corinne Rogers (2015), a Diplomática Digital desenvolveu uma refinada teoria, assim como metodologia específica a partir da diplomática tradicional e da arquivística, com o objetivo de fornecer uma estrutura adequada para avaliar a autenticidade de documentos no sistema digital. Ainda, conforme a autora,

A convergence of perspectives and methods of the digital forensic investigator and the digital archivist is gaining momentum, and may be referred to as digital records forensics, or archival forensics. Shared theoretical perspectives include: (1) authorship and identity (authenticity of origin and forgery). (2) information pattern and change over time (reconstruction and relationships among extant traces and objects), (3) evidential reliability (provenance and integrity), (4) digital materiality and ornament (contextual detail and interpretation). (ROGERS, 2015, p.166-167)

A Diplomática Contemporânea, também conhecida como Diplomática Arquivística ou Tipologia Documental, é uma das abordagens mais recentes sobre Diplomática e volta-se para a análise de documentos de arquivo, especificamente dos tipos documentais. É no campo da Diplomática Contemporânea que se insere a já referida Diplomática Digital, ramo da Diplomática Contemporânea que se volta aos documentos digitais. A Diplomática contemporânea é um estudo voltado muito mais à gênese documental do que simplesmente a autenticidade formal dos documentos. Partindo dessa

abordagem, Luciana Duranti (1996) considera que o tipo documental pode revelar e perpetuar as funções de um documento, daí a importância dos estudos de Tipologia Documental, especialmente como aliada à Identificação Documental.

Uma vez que tem sua origem voltada especificamente para o documento arquivístico, a Tipologia Documental pode também se voltar para o estudo de documentos não diplomáticos, independentemente do formato, suporte ou gênero. Porém, seu principal objeto ainda é o documento escrito e formal. Segundo Heloísa Liberalli Bellotto, a análise tipológica pode ser realizada tanto a partir da perspectiva Diplomática quanto da Arquivística.

Essas diferenças em relação à abordagem da análise tipológica se devem ao fato da arquivística considerar os documentos em seu contexto orgânico de produção, enquanto a Diplomática Científica considera-os isoladamente.

Dessa forma, cada vertente da Diplomática analisa aspectos distintos relativos aos documentos, perpassando pela verificação da autenticidade, genuinidade e falsidade dos mesmos – que assumem um significado distinto para cada tipo de análise diplomática.

Duranti (1996, p. 29) ressalta que “La *autenticidad diplomática* no coincide con la *autenticidad legal*, aun cuando ambas puedan conducirnos a la atribución de una *autenticidad histórica*, en una disputa judicial.” Ainda segunda a autora, em relação aos documentos, os termos autêntico e genuíno se distinguem:

Un documento es “autentico” cuando presenta todos los elementos que se han estipulado para proveerlo de autenticidad. Un documento es “genuino” cuando es verdaderamente lo que se propone ser. Por lo tanto, una sentencia es legalmente auténtica cuando está firmada por el magistrado y es también genuina si la firma no está falsificada. De acuerdo a esto, un privilegio que pretende haber sido emitido por una chancillería imperial es diplomáticamente auténtico, cuando todas sus formas corresponden perfectamente a las que prescriben las reglas de esa chancillería y es también genuino, si realmente ha sido emitido por esa chancillería.

Sin embargo, la distinción entre autenticidad y genuinidad no es válida en sentido histórico. En realidad, el derecho y la diplomática evalúan separadamente las formas del documento y sus autores, de modo tal que podemos tener un documento auténtico que no es genuino y vice-versa. Por el contrario, la historia evalúa solamente el contenido del documento, de modo tal que, desde el punto de vista histórico, auténtico es sinónimo de genuíno (DURANTI, 1996, p. 30).

É importante ressaltar que autenticidade e genuinidade atestadas aos documentos são embasadas na confiabilidade de seu produtor, pois é isto antes de qualquer coisa, que as assegura.

De fato, a verdade do documento é algo que vai além dos critérios das diversas abordagens diplomáticas, pois se trata de uma atividade humana, de modo que não raramente a falsidade documental ultrapassa os mecanismos de criação do documento ao derivar da intencionalidade humana relacionada aos interesses envolvidos. De acordo com Le Goff (1992, p.34), “não existe documento-verdade. Todo documento é mentira. Cabe ao pesquisador não fazer papel de ingênuo”.

Com o grande aumento na produção de documentos em meio digital, a produção documental passa a ter que redobrar seus cuidados com as questões de autenticidade, pois o ambiente digital é mais suscetível a fraudes.

Assim, a Arquivística passa a valer-se de métodos diplomáticos para livrar os documentos digitais de possíveis falsificações. Contudo, a aplicação da diplomática em meio digital é um desafio, de forma que esta deve adaptar-se e buscar amparo na tecnologia. Na verdade, aplicação da Diplomática atualmente tem afastado muito de seu propósito original, de averiguar apenas a verdade contida nos documentos. Isso se deve não apenas ao fato de as condições para essa investigação ter se apresentado de forma distinta nas diferentes variações do método diplomático. Mas principalmente pelo fato de a Diplomática ter conquistado seu espaço junto a Arquivística.

Os documentos arquivísticos se diferem dos diplomas medievais em diversos aspectos, de modo que a verdade documental é tratada de forma diferenciada. A Arquivística considera a questão da falsificação, porém, aspectos como a proveniência e organicidade dos documentos são sua questão central, as quais podem inclusive elucidar a verdade documental.

Assim, a Identificação Documental busca por meio do estudo da produção e dos tipos documentais, estabelecer a proveniência e organicidade dos documentos. Aliás, a Identificação Documental trata-se de uma alternativa para tratar as necessidades do período atual, onde há uma variação cada vez maior de tipos documentais, documentos produzidos em meio digital e aumento da emissão de documentos. Dessa forma, encontra-se entre os métodos que foram fortemente influenciados pela Diplomática.

Uma vez que a Identificação Documental é um método aplicado a documentos de arquivo, a análise diplomática contribui muito para a composição de diversos aspectos de sua metodologia. A Diplomática, mesmo considerando os documentos individualmente, se faz essencial para a contextualização da informação.

A Diplomática vem a relacionar-se com a Identificação exatamente através da produção documental, que é um elemento que possui destaque em ambos estudos. Pois da mesma forma que a diplomática oferece a base para aspectos primordiais da produção documental, como tradição documental, e categorização dos documentos, a Identificação busca elucidar tais aspectos.

Consequentemente, o aporte da Diplomática é essencial para que a Identificação Documental possa contextualizar os documentos de forma coerente e desvendar sua organicidade, até pelo fato de a organicidade ser um dos principais elementos que distinguem o documento de arquivo. Nesse aspecto, a Diplomática contribui com a Identificação Documental por meio da explicitação do contexto de criação e de uso dos documentos, facilitando o revelar das conexões existentes entre informações e documentos, o que é essencial à Organização do Conhecimento.

CONCLUSÕES

De fato, tanto o documento quanto as suas áreas de estudo, com destaque para a Diplomática, tem apresentado transformações desde o momento de sua criação até os dias de hoje, demonstrando o forte elo que os mantêm. Dessa forma, a aplicação da Diplomática atualmente vem ampliando o seu propósito original de averiguar a verdade contida nos documentos. Isso se deve não apenas ao fato de as condições para essa investigação ter se apresentado de forma distinta nas diferentes variações do método diplomático. Mas principalmente pelo fato de a Diplomática ser aplicada hoje pela Arquivística. Os documentos arquivísticos se diferem dos diplomas medievais em diversos aspectos, de modo que a verdade documental é tratada de forma diferenciada. A Arquivística considera a questão da falsificação,

porém, aspectos como a proveniência e organicidade dos documentos são sua questão central, as quais podem inclusive elucidar a verdade documental.

Assim, a Identificação Documental busca por meio do estudo da produção e dos tipos documentais, estabelecer a proveniência e organicidade dos documentos. Aliás, a Identificação trata-se de uma alternativa para tratar as necessidades do período atual, onde há uma variação cada vez maior de tipos documentais, documentos produzidos em meio digital e aumento da emissão de documentos. Nesse sentido, a Diplomática contribui para a análise da Identificação Documental. Desse modo, a união de ambas as áreas de estudo pode colaborar para maior contextualização dos documentos mediante a grande diversidade de possibilidades de sua formulação nos dias de hoje. Corroborando com essa posição, Luciana Duranti afirma que

Agora, mais do que nunca, os arquivistas necessitam de princípios sólidos, permanentes e universais sobre os quais se apoiar, e de conceitos estabelecidos e claramente definidos contra os quais formas novas, ou aparentemente novas, possam ser lançadas, como em uma câmara de neblina, e com os quais possam ser comparadas. Esses princípios e conceitos estão enraizados na ciência da diplomática: é essencial colocar os elementos de seus velhos padrões em contato com os novos padrões determinados pelas tecnologias da informação e fazer novas conexões de modo que as várias partes do sistema de prova documental se reorganizem num todo novo. O trabalho que temos pela frente requer que abramos nossas mentes. (DURANTI, 1994, p. 13).

Enfim, a Diplomática tem muito a proporcionar para que consigamos lidar com o contexto nos quais os documentos estão sendo gerados na atualidade, porém é preciso que saibamos transitar entre o antigo e o novo, aproveitando o que há de melhor, de acordo com a realidade vivenciada, atentando-se para as necessidades de uma eficaz organização e acesso ao patrimônio documental, viabilizando a Organização do Conhecimento.

REFERÊNCIAS BIBLIOGRÁFICAS

Barité, M. G. (2001). Organización del conocimiento: un nuevo marco teórico-conceptual en bibliotecología y documentación. In: Carrara, Kester (org.). **Educação, universidade e pesquisa**. III Simpósio em Filosofia e Ciências Marília. São Paulo: Unesp-Marília-Publicações, Fapesp.

Bellotto, H. L. (2004). **Arquivos permanentes: tratamento documental**. FGV Editora.

Bellotto, H. L. (2002). **Como fazer análise diplomática e análise tipológica de documento de arquivo**. São Paulo: Arquivo do Estado.

Camargo, A. M. A. (2015). Sobre espécies e tipos documentais. In: **Dar nome aos documentos**. São Paulo: Instituto Fernando Henrique Cardoso.

Duranti, L. (1996). Diplomática: usos nuevos para una antigua ciencia. Carmona (Sevilla): **Asociación de Archiveros de Andalucía**.

Duranti, L. (1994). Registros documentais contemporâneos como provas de ação. Trad. Adelina Novaes e Cruz. **Estudos históricos**. Rio de Janeiro, v.7, nº13, p.49-64, jan./jun.

Galene Diaz, J. C.; Garcia Ruiperez, M. (2003). El concepto de documento desde una perspectiva interdisciplinar: de la diplomática a la archivística. **Revista General de Información y documentación**, 13, n 2, p.7-35.

Hjørland, B. (1992). The concept of 'subject' in information science. **Journal of Documentation**. vol.48. n 2, p. 172-20.

Hjørland, B.(2008). What is knowledge organization? **Knowledge Organization**, v.35, n.2/3, p.86-101

Le Goff, J. (1992). Documento. In: **História e Memória**. Campinas: Unicamp.

Romero Tallafigo, M. (1994). Ayer y hoy de La Diplomática: ciencia de La autenticidad de los documentos. In. **Archivística y Archivos: diplomática de ahora mismo**. Carmona: Asociación de Archiveros de Andalucía, p. 13-49.

Rogers, C. (2015). Digital Records Forensics. In DURANTI, L. e FRANKS, P. (eds.) **Encyclopedia of Archival Science**. London: Rowman & Littlefield. p. 166-169

Starn, R. (2002). Truths in the archives. **Common Knowledge**: Duke University, v. 8, n. 2, p. 387-401. Recuperado em setembro, 2015 em:< <http://muse.jhu.edu/journals/ckn/summary/v008/8.2starn.html> >.

Tessier, G. (1952) **La diplomatique**, Paris: Presses Universitaires de France.



ARQUIVOS PESSOAIS E SEUS TIPOS DOCUMENTAIS: A PERSPECTIVA DA REPRESENTAÇÃO

Lucia Maria Velloso de Oliveira¹, Camilla Campoi de Sobral²

¹Fundação Casa de Rui Barbosa (FCRB/Minc), 0000-0003-0546-2602, luciamaria@rb.gov.br

²Universidade Federal Fluminense (PPGCI/UFF), 0000-0002-6725-2929, camillacampoi@id.uff.br

RESUMO O presente trabalho apresenta o desenvolvimento e resultados parciais dos projetos de pesquisas sobre tipologia documental desenvolvidos na Fundação Casa de Rui Barbosa. As discussões teóricas e metodológicas dos projetos consideram o tipo documental como produto social e representação das ações do produtor e de suas interrelações sociais. Nesse sentido, o estudo e uso da tipologia documental busca, por meio da identificação dos vínculos e atividades que deram origem ao documento, garantir uma nomeação adequada dos tipos documentais, possibilitando a potencialização do acesso e difusão ao oferecer aos usuários um mecanismo de acesso que condensa a forma que o documento se apresenta, assim como as atividades que originaram a sua criação.

PALAVRAS-CHAVE *acesso, patrimônio cultural, organização do conhecimento, tipologia documental, arquivos pessoais.*

ABSTRACT This paper presents the development and partial results of research projects on document typology developed at the Fundação Casa de Rui Barbosa. The theoretical and methodological discussions of the projects consider the document type as a social product and representation of the producer's actions and their social interrelations. In this sense, the study and use of the documentary typology seeks, through the identification of the links and activities that gave origin to the document, to guarantee an adequate naming of the documentary types, making it possible to increase access and diffusion by offering users an access mechanism that condenses the form that the document presents, as well as the activities that originated its creation.

KEYWORDS *access, cultural heritage, knowledge organization, document typology, personal archives.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Estudos e pesquisas em arquivos pessoais ainda ocupam um lugar periférico no campo da pesquisa em Arquivologia e na produção bibliográfica na área se compararmos com a produção de conhecimento sobre arquivos públicos ou gestão de documentos. Mas, paulatinamente, vêm ganhando espaço na agenda de pesquisa, na medida em que suscitam reflexões e direcionamentos na elaboração de abordagens que contemplem a potencialização que esses conjuntos documentais representam, em especial para a compreensão da sociedade em seus diferentes momentos históricos. O reconhecimento, pelos usuários, dos arquivos pessoais como patrimônio cultural e campo de interesse proporcionou a

valorização desses conjuntos frente à teoria arquivística, que, por muito tempo, dedicou-se principalmente aos arquivos institucionais e ao tratamento documental desses conjuntos.

Com o crescente interesse pelas análises de tipo biográfico e pelas edições de correspondência, os arquivos pessoais, que sempre despertaram curiosidade por representarem o imaginário social e o acesso à intimidade, passaram a constituir um campo de valiosas fontes de pesquisa.

Esses arquivos, produzidos no âmbito da vida privada de um indivíduo, inseridos numa época específica, possuem especial relevância na construção de uma memória coletiva, necessitando de uma ação de pesquisa para remontar os contextos do período de sua produção. Para isso, é necessário reconhecer o contexto arquivístico como inserido em uma conjuntura social. É importante ressaltar que entendemos contexto arquivístico como o conjunto de ambientes e fatores internos e externos que influenciam a produção, manutenção e uso dos arquivos. Nessa perspectiva, por pressuposto, o documento de arquivo é um produto social, e o produtor do arquivo, um sujeito social que produz documentos para se instrumentalizar diante da necessidade de comprovação de suas ações e para se lembrar. Nesse sentido, estamos diante de duas funções para a existência dos arquivos pessoais: a produção de provas e de memória individual.

Produzidos a partir das necessidades, desejos e preferências de seus titulares, os arquivos pessoais congregam não só as diferentes esferas de relacionamento do produtor com o Estado e a sociedade, mas também a sua intimidade. Esses múltiplos aspectos e facetas da vida dos titulares são refletidos nos contextos que permeiam produção documental e representam também o produtor como sujeito social.

Assim, os arquivos produzidos na esfera pessoal revelam parâmetros da produção documental de um determinado período histórico e do segmento social ao qual o seu produtor pertence. Os elementos reguladores que regem a produção desses registros sociais incluem legislação e normas institucionais, assim como as convenções sociais e práticas culturais do grupo social que o produtor se insere são registradas em secretários, manuais de etiqueta e nas tradições.

O trabalho nos arquivos pessoais deve considerar as características acima expostas no processo de organização do conhecimento, pois cada uma delas possui um papel definido na construção do contexto arquivístico. Segundo Theo Thomassen (2006, p. 11), o contexto arquivísticos extrapola o contexto de produção, pois considera os contextos sócio-político, cultural e econômico que determinam diretamente o conteúdo, a forma e a estrutura dos documentos arquivísticos, além dos contextos de manutenção e uso dos arquivos. Assim, o contexto possibilita a correta compreensão do arquivo. De acordo com a autora Angelika Menne-Haritz “os arquivistas são os únicos especialistas que possuem as ferramentas teóricas e metodológicas para tornar a evidência acessível e assim revelar o contexto exploratório da informação” (Menne-Haritz, 2005, p. 34). A autora alemã, destaca formas de relações contextuais e a sua importância como respostas para se compreender os documentos. Assim, ao falarmos do relacionamento entre documentos, do lugar que um documento ocupa em um conjunto e do lugar de um conjunto dentro de uma organização, estamos relacionando diferentes contextos que permeiam os documentos e que podem ser acionados em conjunto ou isoladamente.

Essa necessidade de compreender os arquivos considerando as múltiplas camadas que influenciam a sua formação configura uma nova demanda da agenda de pesquisas da área e o desenvolvimento de estudos e metodologias de organização sobre arquivos pessoais, permitiram que o Serviço de Arquivo Histórico e Institucional da Fundação Casa de Rui Barbosa, no ano de 2010, desenvolvesse o projeto de pesquisa *Análise tipológica dos documentos em arquivos pessoais: uma representação do código*

social, que tinha por objetivo a identificação de tipos documentais em quatro grandes eventos: nascimento, casamento, grandes viagens e morte. O projeto identificou 41 tipos documentais e apontou para a necessidade de ampliar a pesquisa em torno da tipologia e dos arquivos pessoais. Dessa forma, em 2013, deu-se início ao projeto de pesquisa *Tipologia documental na Família Barbosa de Oliveira*, coordenado pela Dra. Lucia Maria Velloso de Oliveira, e que tem por objetivo principal a identificação dos tipos documentais produzidos no ambiente privado, analisando os costumes dos séculos XVIII e XIX. Articulado ao projeto mencionado anteriormente, o projeto *Tipologia documental na Família Barbosa de Oliveira: definição dos tipos documentais* tem por objetivo a formulação de conceitos dos tipos documentais identificados no projeto citado, organizados até o momento na forma de um glossário.

Ambos os projetos estão intrinsecamente relacionados e promovem o estudo tipológico em arquivos pessoais com o intuito de aprofundar o conhecimento das ações e atividades que dão origem ao documento em si, suscitando desta maneira a potencialização do acesso aos usuários. O estudo da tipologia documental pode ser uma importante ferramenta teórico-metodológica para os arquivos pessoais, pois busca encontrar a ligação existente entre os documentos e os contextos que os geraram. Esses vínculos nem sempre se estabelecem de maneira aparente, podendo apresentar-se de formas sutis. Cabe ao arquivista representá-los e compreendê-los nesse processo.

Segundo a pesquisadora e professora Ana Maria de Almeida Camargo (2015, p. 14), tem sido bastante negligenciada a nomeação adequada dos documentos em relação à organização e descrição dos arquivos. Muitos profissionais se eximem de identificá-los no processo de organização do conhecimento sobre o arquivo objeto de processamento, supondo que, sendo todos da mesma espécie, basta reconhecê-los pela função ou assunto que tratam.

Na contramão desse contexto, os estudos e usos da tipologia documental em arquivos pessoais representam esse esforço significativo de conferir ao documento o nome mais adequado de acordo com as características físicas e a atividade que o gerou dentro de um determinado contexto.

Barité (2001, p. 42-53) destaca que o conhecimento, enquanto produto e necessidade do dínamo social, realiza-se a partir da informação que, ao socializar-se novamente, transforma-se, constituindo-se numa estrutura e num processo de comunicação abertos. Nesse sentido, os esforços nos princípios, métodos e instrumentos para a gestão do conhecimento humano podem ser vistos sob uma tripla perspectiva: sua representação, sua organização e sua comunicação documental.

Assim, o estudo e o uso da tipologia documental buscam a identificação dos vínculos e atividades que deram origem ao documento, assim como a nomeação adequada dos tipos documentais representa, no processo de organização do conhecimento, um dispositivo de potencialização do acesso ao usuário. A discussão da representação no âmbito dos arquivos carece de mais estudos. Discute-se de forma tímida a representação no arranjo e na descrição. O projeto - na medida em que reconhece o tipo documental como produto social e como representação das ações do produtor e de suas interações sociais, bem entendido, contemplando o Estado - oferece uma oportunidade de ampliação do escopo de estudos e pesquisas.

Elizabeth Yakel (2003), em seu artigo *Archival representation*, de 2003 dedica-se ao tema da representação arquivística. Para a autora, a representação arquivística refere-se tanto ao processo de arranjo documental, respeitando ou não a ordem original, como a descrição arquivística e a criação de guias, catálogos, inventários, repertórios etc. Ela também contempla a criação de sistemas, bases de dados e informações arquivísticas estruturadas. De forma analógica entendemos o tipo documental

como uma forma de representação arquivística que, ao identificar as ações e atividades que dão origem aos documentos, possibilitam a compreensão dos contextos que permeiam a gênese documental, evidenciando as relações existentes entre o produtor como sujeito social e as suas formas de registros. A tipologia documental como representação oferece aos usuários amplas possibilidades de abordagens, e, também, permite o reconhecimento das formas de produção documental pelo indivíduo em sociedade.

Busca-se, assim, apresentar o desenvolvimento e os resultados parciais dos projetos de pesquisa em torno da tipologia documental desenvolvidos na Fundação Casa de Rui Barbosa.

Para isso, centraremos as nossas análises no reconhecimento do tipo documental como forma de representação das ações e efeitos da interação entre sujeito e sociedade. Destacamos a definição de Geoffrey Yeo (2007) para documento de arquivo. O autor inova ao definir que “os documentos de arquivo são representações persistentes de atividades criadas por participantes ou observadores dessas atividades ou por seus representantes”(2007, p. 337).

Por extensão do pensamento do autor, o tipo documental identifica as diferentes atividades (aqui compreendidas como todas as ações, transações, manifestações e acordos passíveis de registro), que o indivíduo assume no decurso de seus diferentes papéis sociais. O estudo do tipo documental é fundamental não só para a compreensão do arquivo objeto de estudo, mas também para a descrição arquivística, na medida em que oferece o meio para a identificação do relacionamento entre a forma que os documentos assumem, as ações que lhes dão origem e as funções do produtor.

A descrição arquivística, como representação acurada do arquivo disponibiliza para o usuário as informações substanciais para sua pesquisa. Nenhuma representação expressa em totalidade o seu objeto ou as circunstâncias de sua produção, no entanto é a descrição como resultado de um trabalho de pesquisa sobre o arquivo, que possibilita o entendimento dos elementos constituidores do conjunto documental, do contexto arquivístico e conseqüentemente, faculta as múltiplas interpretações e usos dos usuários.

O estudo do tipo documental evidencia as diferentes formas de registro de atos com cunho jurídico, administrativo e pessoal, incluindo a vida íntima do produtor do arquivo, contribuindo para o conhecimento das formas de relacionamento em sociedade em suas diferentes matizes. Apresentaremos, a título de exemplificação do processo de pesquisa, alguns dos tipos documentais identificados em relação às esferas familiar, social e burocrática.

METODOLOGIA

Como já mencionado, o projeto *Tipologia documental na Família Barbosa de Oliveira*, iniciado em 2013 e sob a coordenação da Dr^a Lucia Maria Velloso de Oliveira, tem como objetivo central a identificação de tipos documentais produzidos no ambiente privado, analisando os costumes dos séculos XVIII e XIX, a partir da *Coleção Família Barbosa de Oliveira* (CFBO) que funciona como campo empírico do projeto.

A *Coleção Família Barbosa de Oliveira* compreende o período do fim do século XVIII até meados do século XX. Seus documentos registram as formas de viver e se relacionar por meio de correspondências e de documentos que comprovam a relação entre o indivíduo e o Estado, e também entre os membros de diferentes famílias que conviviam, casavam e negociavam entre si. Estas famílias uniram-se por

meio de relações matrimoniais e de negócios, e os documentos que compõem a coleção registram o cotidiano familiar desses grupos por meio de vasta correspondência. Os documentos expressam as transações econômicas e as relações sociais, familiares e amorosas (Oliveira, 2014).

A metodologia empregada no projeto de pesquisa citado anteriormente, *Análise tipológica dos documentos em arquivos pessoais: uma representação do código social*, realizado entre 2010 e 2013, contribui para o desenvolvimento do projeto de pesquisa atual. Destaca-se a utilização de instrumento que sistematiza as informações de forma, conteúdo e de vinculação dos documentos, além de informações cunhadas nos manuais de etiqueta, secretários, legislação e outras fontes utilizadas para identificação de um padrão social ou legal. O presente projeto ampliou a pesquisa bibliográfica e utiliza dicionários das áreas da Arquivologia, do Direito, da Sociologia, da Arquitetura, da Comunicação, da Literatura, da História, a legislação da época e também os manuais de costumes e secretários, não apenas para a identificação da forma e da organização de conteúdo dos documentos, como também para subsidiar a elaboração de verbetes que orientaram a conceitualização dos tipos documentais identificados. Os glossários da área também estão incluídos no escopo bibliográfico. Além disso, incluiu-se uma análise comparativa entre as práticas de definição do tipo documental em bases de dados internacionais. A pesquisa em bases internacionais (*Library of Congress/EUA*, *National Archives/Reino Unido* e *Archives Nationales/França*) tem sido fundamental para identificação da prática da área para nomear os tipos documentais em seus instrumentos de recuperação da informação *on-line*.

Ressaltamos que, no âmbito da pesquisa, a necessidade de conceitualização dos tipos documentais surgiu com o fim do projeto *Análise Tipológica dos documentos em arquivos pessoais: uma representação do código social*, que tinha por objetivo a identificação de tipos documentais em quatro grandes eventos: nascimento, casamento, grandes viagens e morte. O projeto finalizado havia evidenciado a importância da elaboração de conceitos que permitissem que os tipos documentais identificados fossem capazes de contemplar não somente a *Coleção Família Barbosa de Oliveira*, mas também outras coleções ou arquivos custodiados por outras instituições de preservação de acervos pessoais. Foram revisitados, apenas no âmbito da conceitualização, os tipos documentais do projeto anterior.

Atuam no projeto uma arquivista doutora (coordenadora do projeto), duas arquivistas mestres e uma bolsista de iniciação científica¹. O projeto conta com a colaboração de uma renomada pesquisadora² da área, vinculada à Universidade de São Paulo (USP), cuja experiência e produção contribuem significativamente para as discussões teóricas e metodológicas da pesquisa e também para a definição dos tipos documentais.

O projeto concentra duas frentes de trabalho, que se relacionam de forma dinâmica. A primeira concentra-se na análise, pesquisa em fontes, identificação e validação dos tipos documentais. A segunda, em um levantamento quantitativo das ocorrências de termos relativos às espécies e tipos documentais nos instrumentos de pesquisas das bases de dados internacionais: arquivos da *Library of Congress* (EUA), *National Archives* (Reino Unido) e *Archives Nationales* (França). Os critérios adotados no

¹ A equipe de pesquisadores dos projetos mencionados é composta pela Dr^a Lucia Maria Velloso de Oliveira, M^a Bianca Therezinha de Carvalho Panisset, M^a Isabel Cristina Borges de Oliveira, M^a Tatiane Lopes dos Santos e Vitória de Freitas Machado.

² A pesquisadora que colabora com a nossa pesquisa é a Prof^a Dr^a Ana Maria de Almeida Camargo.

levantamento priorizaram os arquivos ou coleções pessoais ou familiares, com datas-limite entre os séculos XVIII e XIX.

Quanto à identificação, validação e conceituação dos tipos documentais a partir da análise tipológica (que compreende o reconhecimento da atividade que justifica a produção do documento, a forma registrada que este adquire e o conceito que traduz essa relação entre atividade e forma), é realizada a formulação de hipótese quanto à identificação do tipo documental. Em seguida, após a pesquisa em fontes, a identificação é ratificada ou alterada e é feito o registro das fontes utilizadas. A validação ocorre em reunião da equipe dos dois projetos com a professora que vem contribuindo para a pesquisa. Posteriormente, o tipo documental é conceituado com o auxílio de glossários, manuais de costumes e legislação da época, além de fundamentação na Teoria dos Conceitos.

O levantamento realizado nas bases de dados internacionais contribui para o projeto como método balizador dos termos e ocorrências relacionados aos tipos documentais e possibilita que adotemos termos que apresentam eco nos dados obtidos a partir do levantamento quantitativo. Nessa etapa foi possível reiterar o uso dos tipos documentais utilizados tradicionalmente na área como: carta de amor, carta de negócios, relatório de negócios, carta de condolências, carta de recomendação, anúncio fúnebre, cardápio, entre outros tipos. Nesse levantamento ficou evidenciado que a área não se dedica à identificação do tipo documental de forma aprofundada já que nas bases de dados encontramos referências à gênero, espécie, técnica, assunto e tipo documental em um mesmo campo de informação. Isso por outro lado, valoriza a pesquisa na medida em que indica a necessidade de estudos sobre tipologia documental como forma de produzir mecanismos de recuperação da informação mais precisos para os usuários.

Os termos sistematizados neste levantamento são também indicativos de possibilidades em relação a espécies e tipos. E a sua contribuição é a compreensão do posicionamento das instituições sobre os tipos documentais. Mais precisamente, sobre as lacunas e ausências de termos que sejam tipos documentais e que, também, são especialmente significativas nessa compreensão (Sobral, 2016).

RESULTADOS

A tipologia documental como método de análise em arquivos pessoais permite oferecer ao usuário uma fórmula condensada que exprime a forma com que o documento se apresenta mais a sua atividade geradora.

O estudo do tipo documental também permite analisar aspectos não explicitados na estrutura do documento, fornecendo elementos para a compreensão do comportamento da sociedade, das formas de viver e se relacionar e, também, das normatizações e legislações vigentes.

Essa compreensão e sensibilidade são fundamentais no trabalho de descrição em arquivos pessoais, o qual deve, segundo Oliveira (2010), fundamentar-se na reconstrução dos contextos arquivísticos e na representação dos acervos, de forma a assegurar a expressão dos papéis sociais dos titulares e na facilitação do atendimento ao usuário.

O desenvolvimento dos projetos *Tipologia documental na Família Barbosa de Oliveira* e *Tipologia documental na Família Barbosa de Oliveira: definição dos tipos documentais* identificaram, até o presente momento, 203 espécies e tipos documentais em 649 documentos do universo empírico

analisado.³ A conceituação dos tipos documentais encontra-se em processo de elaboração e, até o momento, já foram conceituados 175 tipos.

CONCLUSÕES

O trabalho realizado no projeto combina a identificação espécies documentais e tipológicas produzidas nos séculos XVIII e XIX e o estudo dos contextos de normas e padrões que englobam a produção do registro documental. Por tratar-se de um universo empírico que compreende a vida privada e suas relações, estas atividades são extremamente delicadas, evidenciando a necessidade de uma atenção aos padrões de comportamento e conduta vigentes no período da produção documental. Além disso, introduz a questão da representação em arquivos através de uma abordagem distinta do arranjo e da descrição.

Assim, a tipologia documental não se apresenta somente como método de produção de conhecimento sobre o arquivo ou uma forma de potencializar o acesso ao usuário, na medida em que permite a identificação dos vínculos presentes nos documentos de arquivo. Trata-se, essencialmente, de uma forma de retratar a forma de viver em um momento histórico. Essa abordagem torna-se mais reveladora como forma de representação da informação e do conhecimento, ao oferecer ao usuário uma fórmula nuclear das ações e atividades que originaram o documento e também ao oferecer, na conceituação do tipo documental, as formas de registro e o contexto histórico-social a que o produtor se inseria no momento da produção documental.

Ao final do projeto, será possível disponibilizar ao usuário os tipos documentais identificados no universo empírico. Estes servirão como um dispositivo de acesso e difusão, na medida em que fornecem ao usuário as atividades que deram origem ao registro documental. Também será possível a elaboração de uma bibliografia de referência que contribua no que diz respeito à tipologia documental em arquivos pessoais.

REFERÊNCIAS BIBLIOGRÁFICAS

Barité, M. (2001). Organización del conocimiento: un nuevo marco teórico-conceptual en Bibliotecología y Documentación. *Educação, universidade e pesquisa* (pp. 35-60). Marília: Unesp-Marília-Publicações.

Bellotto, H. L. (2008). *Diplomática e tipologia documental em arquivos*. Brasília: Briquet de Lemos/Livros.

Camargo, A. M. de A. (2009). Arquivos pessoais são arquivos. *Revista do Arquivo Público Mineiro*, 2 (pp. 26-39). Belo Horizonte.

_____. (2015). Sobre espécies e tipos documentais. In Camargo, A. M. de A. et al (org.). *Dar nome aos documentos* (pp. 14-31). São Paulo: Instituto Fernando Henrique Cardoso.

³ Resultados parciais.

Menne –Haritz. A. (2005). Avaliação ou Documentação: podemos avaliar arquivos através da seleção de conteúdos? *Registro: Revista do Arquivo Público Municipal de Indaiatuba*, 4(4), (pp. 18-34). São Paulo.

Oliveira, L. M. V. de. (2010) *Análise tipológica dos documentos em arquivos pessoais: uma representação do código social*. Disponível em: http://www.casaruibarbosa.gov.br/dados/DOC/bolsistas/2010/FCRB_Selecao_de_Bolsistas_2010_Analise_tipologica_dos_documentos.pdf. Acesso em: 18 maio 2015.

_____. (2012). *Descrição e pesquisa: reflexões em torno dos Arquivos Pessoais*. Rio de Janeiro: Móbile.

_____. (2016) Os tipos documentais: a forma de registro de um modo de viver. In *Seminário Hispano Brasileiro de Investigación en Información, Documentación y Sociedad*, 5, Madrid. Inédito.

_____. (2014) *Tipologia documental na família Barbosa de Oliveira*. Disponível em: http://www.casaruibarbosa.gov.br/arquivos/file/Bolsistas13/aa%20_%20Tipologia%20na%20Familia%20Barbosa%20de%20Oliveira.pdf. Acesso em: 10 maio 2015.

_____. (2015). *Tipologia Documental na Família Barbosa de Oliveira: Definição dos Tipos Documentais*. Fundação Casa de Rui Barbosa.

Oliveira, L. M. V. de & Oliveira, I. C. B. (2016). *Glossário de tipos documentais em arquivos pessoais* (versão preliminar).

Sobral, C. C. de. (2016). *Relatório parcial de atividades*. Fundação Casa de Rui Barbosa.

Thomassen, T. (2006). Uma primeira introdução à arquivologia. *Arquivo & Administração*, 5(1), (pp. 5-16). Rio de Janeiro.

Yakel, E. (2003). Archival representation. *Archival Science*, 3(1), (pp. 1-25).

Yeo, G. (2007). Concepts of record (1): evidence, information, and persistent representations. *The American Archivist*, 70(2), 315-343.



REFLEXÕES SOBRE UMA ORGANIZAÇÃO DO CONHECIMENTO FOTOGRÁFICO SEGUNDO SEU CONTEXTO DE PRODUÇÃO DOCUMENTAL

Bruno Henrique Machado¹, Rafael Semidão², Telma Campanha de Carvalho Madio³

¹Universidade Estadual Paulista (UNESP), Faculdade de Filosofia e Ciências, Marília, 0000-0002-3034-3122, machadobrunohenrique@gmail.com

²Universidade Federal do Rio Grande (FURG), Rio Grande. Bolsista do convênio FAPESP/CAPES processo nº 2015/05677-4, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) 0000-0002-9715-3702, rafaelsemidao@gmail.com

³Universidade Estadual Paulista (UNESP), Faculdade de Filosofia e Ciências, Marília, 0000-0002-7031-2371, telmacarvalho@marilia.unesp.br

RESUMO Buscamos oferecer uma reflexão acerca da perspectiva arquivística diante da organização do conhecimento fotográfico de acervos institucionais. Recorrendo a uma abordagem bibliográfica, chamamos ao diálogo alguns autores da área para tecer um argumento a respeito daquilo que consideramos ser a mais condizente e rica abordagem de organização do conhecimento fotográfico em termos arquivísticos, a identificação do lastro de elementos que vão da fotografia, na condição de índice, até a sua gênese documental. Consideramos que tal perspectiva, por mais árdua que seja a sua operacionalização, permitiria um resgate histórico mais profundo do quadro de conhecimentos que os documentos fotográficos podem revelar. Como resultado das reflexões propomos, em primeiro lugar, não a introdução de novas técnicas de processamento, mas uma mudança de postura dos arquivistas, no sentido de se adotar um posicionamento investigativo de teor historiográfico e interpretativo que deixe de tomar a imagem fotográfica veiculada no suporte como insumo exclusivo do conhecimento, e que procure recuperar a via que leva da imagem (como índice) até o seu contexto de produção documental, proporcionando assim uma organização do conhecimento fotográfico apta a refletir a complexidade das ações, propósitos e concepções que condicionaram a produção da fotografia e o momento histórico que ela registra.

PALAVRAS-CHAVE *Documentos fotográficos, Gênese documental, Organização do Conhecimento, Arquivística.*

ABSTRACT It was sought to offer a reflection on the archival perspective in front of the organization of the photographic knowledge of institutional collections. Using a bibliographical approach, some authors of the area were called to the dialogue to make an argument about what is considered to be a more consistent and rich approach to the organization of photographic knowledge in archival terms, the identification of items that go from photography, in Index condition, to its documentary genesis. Believing that such perspective, however arduous it may be, would allow for a deeper historical retrieval of the framework of knowledge that photographic documents can reveal. As a result of the reflections it is proposed, at first, not an introduction of new processing techniques, but a change of position of the archivists, in the sense of adopting an investigative position of historiographic and interpretative content that stops taking a photographic image conveyed in the support as an exclusive input of knowledge, and that seeks to retrieve the image (as an index) to its context of documentary production, thus providing an organization of photographic knowledge able to reflect a complexity of actions, purposes and conceptions that conditioned the production of photography and the historical moment it records.

KEYWORDS *Photographic documents, Documental Genesis, Knowledge Organization, Archival Science*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Com o processo contínuo de consolidação das técnicas fotográficas a partir do século XIX, a fotografia foi, *pari passu*, recebendo um espectro de funções documentais que abarcou desde o registro efemérides de âmbito local e regional, como o registro do desenvolvimento de cidades e indústrias, até o registro de produções artísticas pela atribuição do qualificativo de obra de arte às imagens fotográficas. Nesse sentido, Tagg (2007, p.81) apresenta as configurações funcionais que a fotografia adquiriu socialmente, informando que foi adotada por “uma série de instituições reformadas ou emergentes, de tipo médico, legal e municipal, nas quais as fotografias funcionavam como meio de arquivo e como fonte de prova”.

Nesse processo sócio documental assinalado por Tagg, a fotografia se insere no âmbito documental da esfera teórica que Guimarães e Tognoli (2015) denominam de *archival knowledge organization* e que aqui traduzimos como organização do conhecimento arquivístico, cujo processamento teórico e técnico, essencialmente, procura seguir o lastro de funções e atividades que produziram o documento, adotando uma abordagem de contextualização administrativa, histórica e social, que se expressa na seriação orgânica dos documentos de mesma proveniência.

Uma vez que se insira a fotografia nesse âmbito, ele passa a ser compreendido como um construto institucional (evado de intenções, propósitos, ideários etc.), cuja fotografia veiculada no suporte figura como um indício da presença de elementos contextuais que necessitariam ser identificados para proporcionar uma melhor compreensão a respeito dos conhecimentos administrativos e históricos que a fotografia encerra em si. Sendo considerada nessa perspectiva, a fotografia, arquivisticamente processada, poderá proporcionar conhecimentos mais condizentes com a complexidade dos fatos administrativos e históricos.

METODOLOGIA

Dessa maneira, busca-se propor uma discussão teórica que postula a inserção da fotografia no interior do universo da organização do conhecimento arquivístico, desde uma perspectiva que a considere como índice de um lastro de intenções, ações institucionais e procedimentos técnicos e químicos e hoje digitais que deve remontar até a sua gênese documental, a despeito de uma abordagem que se restrinja a descrever a imagem veiculada enquanto expressão exata de uma realidade histórica. Tal proposta, que poderíamos denominar organização do conhecimento fotográfico segundo seu contexto de produção documental, tem como objetivo chamar à discussão e, desse modo, contribuir com os estudos que pretendem o entendimento da fotografia dentro do âmbito teórico da Arquivística.

FUNÇÃO DOCUMENTAL DA FOTOGRAFIA NO ÂMBITO ARQUIVÍSTICO

A evolução das técnicas fotográficas se deu em grande parte no período entre o final século XIX e início do século XX. Nesse diapasão temporal inicial, a imagem fotográfica era considerada um espelho do

real, um meio pelo qual se alcançaria uma relação direta dos elementos da realidade por meio de processos mecânicos. Para Dubois (1993, p. 25) as concepções de fotografia desse período a compreendiam “como uma espécie de prova, ao mesmo tempo necessária e suficiente que atesta indubitavelmente a existência daquilo que mostra”, isto é, como uma expressão direta dos fatos, um espelho que, por refletir a realidade, tem na imagem que veicula o conteúdo total dos fatos, bastando descrevê-lo como aparece para que se alcance uma representação fidedigna, sem considerar necessariamente nenhum pressuposto, intenção, objetivo por detrás da imagem.

Inversamente a essa perspectiva inicial do movimento de consolidação da fotografia em função documental – captura do espelho da realidade - a organização do conhecimento arquivístico prioriza a dimensão da fotografia como índice, pois compreende o índice como o registro da execução de uma atividade administrativa. O arquivista busca a integração da atividade administrativa geradora com o documento, sua gênese documental. Neste caso, a informação fotográfica do referente que se mostra é apenas mais uma característica. A caracterização do documento fotográfico como índice implicaria que ele necessariamente representasse algo a ser identificado pelo arquivista (Lopez, 2008). De modo que o conteúdo imagético caracterizado pelo índice é a materialização de uma função administrativa que foi concretizada.

Assim, toma a definição de documentos fotográficos apresentada por Rezende e Lopez (2014) como a norteadora do artigo, pois os autores compreendem de maneira mais aglutinadora, pela justificativa de ainda encontrarmos muita diversidade da técnica fotográfica que se refletem nos acervos documentais, e apresentam a definição ampla, que engloba todos os processos técnicos de constituição da fotografia do analógico ao digital

Es comprendido en su sentido más amplio, abarcando la diversidad de los procesos técnicos, siendo preferente el uso de término “documentos fotográficos” para los documentos imagéticos generados por todos los procesos de captación fotográfica de imágenes, incluyéndose los digitales. Sin embargo, existen otros investigadores que defienden el uso de la expresión “documentos fotográficos” apenas para los que resultan de los procesos fotoquímicos. Aquellos que resultan de procesos digitales no pueden llamarse fotografías, sino “imágenes digitales”, tanto por su naturaleza numérica como por la cuestión de la ruptura referencial. Dicha corriente de pensamiento considera que la imagen digital no mantiene un “vínculo umbilical” con el referente, así como lo tiene la imagen fotográfica (calco fotoquímico del cuerpo referenciado), por basarse, o ser alterada, por cálculos algebraicos (Rezende e Lopez, 2014 p. 172).

Deixando evidente que a natureza de um documento de arquivo não é determinada pela sua forma, suporte ou linguagem, Delmas (2010) afirma que a ideia de documento de arquivo deve ser completada pela noção de informação orgânica. Frisando que o constitui tal tipo de documento não são os suportes, mas sim as informações que correspondem às construções, linguagens e procedimentos.

Portanto, a exegese do documento de arquivo parte de se o mesmo é produzido e/ou recebido a partir de uma atividade administrativa, não importando a maneira pela qual a comunicação é transmitida nem o gênero documental. Assim, todos documentos devem ser contemplados pelos princípios norteadores da Arquivística.

Joan Schwartz (1995) considera que os arquivistas ainda têm perpetuado o analfabetismo visual, promovendo na maioria das vezes a descontextualização, que visa apenas em termos de seu valor informativo, acessível pelo nome ou lugar, pessoas, entre outros, ratificando que os documentos fotográficos são retirados de seu contexto funcional e do poder comunicativo.

Nesse sentido, Heredia Herrera (2016, p. 3) afirma que “Hoje, não cabe nenhuma discussão sobre o papel documental — e, como tal, informativo — que a fotografia tem para a memória coletiva da sociedade, portanto, reivindica o direito à preservação e ao tratamento.”

Quanto à introdução do documento fotográfico nos arquivos Heredia Herrera (2016) afirma que permanecem diferentes aspectos que necessitam ser esclarecidos, tais como os elementos das esferas terminológicas, conservacional, das possibilidades informacionais, dos modos de organização, do tratamento processual, e da classificação e análise.

A autora propõe como principal critério a vinculação institucional como o que a diferencia dos diversos suportes. Apenas com estabelecimento desse lastro orgânico seria possível tratar arquivisticamente o documento fotográfico, tanto no que tange à sua produção, à sua conservação e armazenamento e às suas propriedades informacionais e possibilidades de recuperação.

Para a autora, as terminologias “arquivos fotográficos” e “fundos fotográficos” são denominações que entram em conflito com a terminologia arquivística, pois, “arquivos fotográficos” se apresentam como o local de custódia para os documentos fotográficos. De toda maneira, o conceito de arquivo faz referências ao seu produtor, no caso apresentado pela autora como arquivos pessoais - de um fotógrafo por exemplo. A alusão de “fundo fotográfico”, sob o ponto de vista terminológico, é equivocada. A denominação “fundo” diz respeito a uma estreita relação da produção documental de uma instituição, já “fundo fotográfico” vem sendo utilizada, equivocadamente como sinônimo de coleção fotográfica .

Esse apontamento vem de encontro às afirmações de Lacerda (2008), a qual questiona a ausência de metodologia de processamento arquivístico dos documentos fotográficos, dando assim legitimidade para uma abordagem que privilegie questões relativas à preservação dos suportes e ao conteúdo visual da imagem. Tal pensamento é reafirmado por Dámian Cervantes (2008)

Agruparlos por el apoyo supondría alejarse de su origen, y desvincular la relación entre el contenido de la información de los documentos: en materiales magnéticos pueden tener la misma información textual, iconográfica, de audio o audiovisual. Sin embargo, en un archivo fotográfico, independientemente de soportes magnéticos, películas, papel o plástico, entre otros existentes, habrá siempre seguirá siendo el mismo lenguaje para responder a las necesidades de gestión de la entidad que produce. (Damian Cervantes, 2008, p. 59).

Entretanto, esse agrupamento apresentando por Damian Cervantes quebra a cadeia de produção documental, especificamente o princípio da organicidade. Conforme apresenta por Luciana Duranti (1997) que a organicidade é a configuração que reflete as relações entre os diversos documentos produzidos por uma instituição que compõem um fundo de arquivo, criados em razão das funções e atividades de uma instituição ou pessoa. Estas relações orgânicas dos documentos evidenciam a relação entre o arquivo e seu organismo produtor.

Por seu turno, Lopez (2011) argumenta que os documentos fotográficos de arquivo, ao receberem a organização individualizada e a criação de coleções reduzem a compreensão do significado global pelo qual os documentos foram produzidos.

A inserção dos documentos fotográficos no ambiente de arquivo é tratada por Boadas i Raset (2008), de quatro formas/situações da entrada dos documentos fotográficos no processo de organização arquivística institucional: primeira, as fotografias geradas pela própria entidade que seguem um processo administrativo; segunda, as fotografias encontradas em reportagem pela própria instituição,

relacionadas às atividades culturais ou promocionais; terceira, pela aquisição ou compra do material fotográfico para uso institucional; e quarta, por doações de acervos fotográficos eivados de valor histórico para a instituição.

Nessa linha, segundo Iglésias Franch (2009, p. 1), conhecer o contexto de produção desses documentos fotográficos é fundamental, para “a identificação destes entes pode ser um dado apreciável para a tomada de decisões com respeito à conservação, além de proporcionar uma informação essencial para a organização de fundos”. O autor relaciona os documentos fotográficos aos outros documentos arquivísticos que são produzidos pela instituição.

Madio (2012), por sua vez, insiste que os procedimentos arquivísticos (identificação, classificação, avaliação, descarte e guarda permanente) sejam aplicados aos documentos fotográficos e que desde a sua produção seja incorporada no fluxo documental, considerando-a como produzida com um fim específico e com funções definidas e estabelecidas, para com isso evidenciar ações realizadas pela instituição, ou seja, a manutenção que a organização do conhecimento arquivístico propõe.

Com isso, fica sugerido que o contexto de produção dos documentos é essencial para propiciar aos pesquisadores a compreensão sobre os motivos que levaram à criação do mesmo, os quais ficam impossibilitados de serem identificados somente através do conteúdo da imagem fotográfica e, ademais, a atribuição de prazos de guarda e destinação final fica prejudicada pela identificação equivocada dos documentos (Lopez e Carvalho, 2013).

Desse modo, seria necessário que os princípios de proveniência (contexto de produção institucional do documento) e de organicidade/vínculos arquivísticos (as relações administrativas que se refletem no documento) e a gênese documental (a origem, a ação) sejam garantidos, pois mais do que um mero recurso ilustrativo, a fotografia produzida em decorrência de funções administrativas preestabelecidas atinge o patamar de documento de arquivo e como tal, requer a aplicação de todos os processos arquivísticos.

Essas características são apreciados por Cook (1992) a apresentar o conceito de fundo documental, que contempla qualquer documento arquivístico produzido pela instituição, assim:

The fonds, therefore, should be viewed primarily as “an intellectual construct.”The fonds is not so much a physical entity in archives as it is the conceptual summary of descriptions of physical entities at the series level or lower, and descriptions of the administrative, historical and functional character of the records creator(s) - as well as descriptions of the records-creating processes (metadata). The fonds is thus the conceptual "whole" that reflects an organic process in which a records creator produces or accumulates series of records which themselves exhibit a natural unity based on shared function, activity, form or use. It is at the heart of this process or relationship linking the creator to the records that the essence of provenance or respect des fonds can be found and must be protected. It is at this functional heart, moreover, that archival descriptive *systems should be aimed, structured and standardized* .(Cook, 1992, p. 33).

Assim, as imagens veiculadas pelos documentos fotográficos seriam o ponto de confluência de um lastro de elementos historicamente contextualizados que necessitaria ser tomado em conta quando da aplicação dos procedimentos técnicos de organização.

Deixando fora de dúvida que a introdução da fotografia no âmbito dos arquivos, a síntese das acepções dos autores a que se recorreu, expressa a clara ideia de que os documentos fotográficos precisam receber um tratamento técnico que considere os princípios arquivísticos (sem se resumir a mera descrição da

imagem) e estar sempre relacionado aos demais documentos arquivísticos para a identificação do contexto institucional de produção e, ao menos em parte dos autores a quem se recorreu, expressa também o fator construtivista do conhecimento registrado pelos documentos fotográficos uma vez que recomenda a consideração do lastro de atividades e funções que produziram a fotografia até elementos da situação de sua produção, como a atuação do fotógrafo, as técnicas de fixação da imagem, os contratantes, entre outros e a instituição.

PARA UMA ORGANIZAÇÃO DO CONHECIMENTO FOTOGRÁFICO SEGUNDO SEU CONTEXTO DE PRODUÇÃO DOCUMENTAL.

Consideramos a organização do conhecimento fotográfico uma vertente da esfera teórica da organização do conhecimento arquivístico que, sendo um campo independente, faz ao mesmo tempo parte da organização do conhecimento, conforme postula Hjørland (2016)

Knowledge organization in archives should, however, also be considered a part of KO as defined in the beginning of this article. Archives may contain official records, business records, images, letters, diploma etc. The most important specific principle of organization for this domain is the principle of provenance (Hjørland, 2016, p. 480).

Partindo da perspectiva dual de Hjørland (2008) sobre a natureza da organização do conhecimento, acreditamos que a organização do conhecimento arquivístico e, dentro dela, a organização do conhecimento fotográfico, compõem o sentido “narrow” de organização do conhecimento, isto é, aquele em que se enquadram as atividades de “document description, indexing and classification performed in libraries, bibliographical databases, archives and other kinds of “memory institutions” by librarians, archivists, information specialists . . .” (Hjørland, 2008, p. 86).

Dessa forma, como visto por meio do recurso a alguns autores da Arquivística, a compreensão de que o documento fotográfico requer ser processada de acordo com o *modus operandi* arquivístico, que observa o quadro de atividades e funções que contextualizam organicamente os documentos, está lançada, apesar de se reconhecer todas as dificuldades inerentes a tal processamento conforme já mencionado por Machado e Madio (2015).

O que se propõe como meio para se alcançar uma organização do conhecimento fotográfico segundo seu contexto de produção documental, é que se conscientize da natureza indicial, advindo de todos os recursos técnicos, químicos, óticos e de construção da imagem, do documento fotográfico, por detrás do qual está latente o lastro de atividades, funções, intenções e objetivos institucionais que constroem a “realidade” retratada. Tratar-se-ia, nessa medida, de uma concepção de construtivismo histórico, na base mesma do conhecimento registrado pelos documentos fotográficos. Conhecimento este que, como tal, requer ser organizado e representado observando-se o mesmo lastro de elementos que está presente como condicionante de seu contexto de produção e que encontra na imagem veiculada um índice o indicativo para um processo investigativo. Com isso, o que se propõe não são novas técnicas de processamento, mas sim uma guinada quanto à postura do arquivista.

Seria necessário assumir uma perspectiva de investigação na produção documental, até certo ponto, em que se busca identificar o lastro de elementos latentes na gênese documental. Partindo dessas premissas, a origem do documento está relacionada a uma origem corporativa, na qual se encontra uma rotina a ser cumprida; daí a ação que provoca a elaboração para produção do documento, ou conforme

mencionado por Duranti, Eastwood e MacNeil (2002), o contexto jurídico-administrativo que a atividade está inserida.

A gênese documental é materializada no documento ou, conforme a diplomática, a junção do *actio* (ação, fato, ato) com o *conscriptio* (sua transferência, passagem para a escrita do documento). Portanto, a gênese documental é dividida em etapas: nasce de uma iniciativa, depois de uma deliberação (momento da ação)e, dessa deliberação, o documento é produzido, implicando na escolha do tipo documental adequado e, posteriormente, no momento da execução da documentação (Bellotto, 2014).

Postulado que compreende o conhecimento registrado pelos documentos fotográficos como a confluência de várias linhas causais de um construtivismo que, notadamente não dá margem para que ingenuamente se compreenda tal conhecimento como expressão direta de fatos da realidade, e sim um documento arquivístico.

Há certamente nisso uma troca de papéis, no sentido de adotar a historiografia como técnica auxiliar da Arquivística, na medida simetricamente inversa da relação que no passado ligou a Arquivística à História. Essa postura, sem dúvida, acarreta inúmeras dificuldades e envolve a adoção de novos critérios, mas, ao mesmo tempo, possibilita um processamento técnico cujo resultado proporcionaria uma organização mais condizente com os meandros históricos inerentes à produção documental dos documentos fotográficos, e com isso poderia oportunizar aos usuários um resgate mais fidedigno de informações a respeito dos objetos por eles estudados.

Sem isso, se não no todo, ao menos em parte, o que nortearia a organização das fotografias seria a concepção de que a imagem veiculada reflete diretamente a realidade, bastando, dessa maneira, analisá-la e descrevê-la em si mesma, tolhendo assim aos pesquisadores importantes informações a respeito dos muitos condicionantes institucionais que influenciaram na construção do conhecimento veiculado pelas fotografias.

Propõe-se, dessa forma, uma postura que tome a imagem captada pelo documento fotográfico como índice de um conjunto de elementos historicamente condicionantes do conhecimento resultante de sua organização e que adote a abordagem historiográfica interpretativa como técnica privilegiada para a recuperação do lastro de produção do documento fotográfico a ser arquivisticamente processado.

REFERÊNCIAS BIBLIOGRÁFICAS

Bellotto, H. L. (2014). *Arquivologia: estudos e reflexões*. Belo Horizonte: Editora UFMG.

Boadas I Raset, J. (2008). Patrimônio Fotográfico: Estrategias para su gestión. *Revista de los museos de Andalucía*, n. 9, pp. 28-31.

Cook, T. (1992). The Concept of the Archival Fonds in the Post-Custodial Era: Theory, Problems and Solutions. *Archivaria*, number 35. pp.24-37.

Dámian Cervantes, G. (2008). *Los documentos especiales en el contexto de la Archivística*. México DF: Edición electrónica.

Delmas, B. (2010). *Arquivos para quê?: textos escolhidos*. São Paulo: Instituto Fernando Henrique Cardoso.

- Dubois, P. (1993). *O Ato Fotográfico e Outros Ensaio*s. Campinas: Papirus.
- Duranti, L. (1997). The Archival Bond. *Archives and Museum Informatics*, n.11, pp. 213-218.
- Duranti, L. ; Eastwood, T.; MacNeil, H. (2002) *Preservation of the Integrity of Electronic Records*. Vol. 2, The Archivist's Library. Dordrecht: Kluwer.
- Guimarães, J. A. C., Tognoli, N. B. (2015) Provenance as a Domain Analysis Approach in Achival Knowledge Organization. *Knowledge Organization*, 8:40, pp. 562-569.
- Heredía Herrera, A. (2016). A fotografia e os arquivos. *Revista Photo & Documento*. Num. 2. Acedido em 15 de maio de 2017 em: <https://goo.gl/8FM22m>
- Hjorland, B. (2008). What is Knowledge Organization (KO)? *Knowledge Organization*, 2:35, pp. 86-101.
- Hjorland, B. (2016) Knowledge Organization (KO). *Knowledge Organization*, 43:6, pp. 475-484.
- Iglésias Franch, D. (2009). Materiales Fotográficos: Conocer, Analizar y Preservar. In. *IV Jornada Provincial de Archiveros*. Cordoba. Acedido em 12 de janeiro de 2017 em: : www.girona.cat/sgdap/docs/materiales_fotograficos.pdf
- Lacerda, A. L. de. (2008) **A fotografia nos arquivos**: a produção dos documentos fotográficos na Fundação Rockefeller durante o combate à febre amarela no Brasil. São Paulo, Tese de Doutorado – Pós-Graduação em História Social.
- Lopez , A. P. A.(2008). *O contexto arquivístico como diretriz para a gestão documental de materiais fotográficos de arquivo*. Acedido em 12 de Janeiro de 2017 em: <http://repositorio.unb.br/handle/10482/303>
- Lopez , A. P. A. (2011).Contextualización archivística de documentos fotográficos *Alexandria: Revista de Ciencias de la Información*, 8: enero-diciembre. Acedido em 12 de Janeiro em 2017 em: <http://revistas.pucp.edu.pe/index.php/alexandria/article/view/213>
- Lopez , A. P. A., Carvalho, P. D. S. (2013). A Classificação Arquivística por assunto em documentos fotográficos: o exemplo do Arquivo Público do Distrito Federal. *Perspectivas em Gestão & Conhecimento*, João Pessoa, 3:2. pp. 271-279. Acedido em: 10 de Junho de 2016 em: <http://periodicos.ufpb.br/ojs2/index.php/pgc/article/view/17470>
- Machado, B. H., Madio, T. C. de C. (2015). *Documentos fotográficos de arquivo em museu*. In Desafios y oportunidades de las Ciencias de la Información y la Documentación en la era digital: actas del VII Encuentro Ibérico EDICIC 2015 (Madrid, 16 y 17 de noviembre de 2015). Universidad Complutense de Madrid, Madrid. Acedido em 12 de Janeiro de 2017m:<http://eprints.sim.ucm.es/34715/>
- Madio, T. C de C. (2012). Uma Discussão dos Documentos Fotográficos em Ambiente de Arquivo. In. Valentim. M. L. P. (Org.), *Estudos Avançados em Arquivologia*. (pp. 55- 68) Marília: Oficina Universitária. ; São Paulo: Cultura Acadêmica.
- Rezende, D. S. Lopez, A. P. A. Adecuación de la descripción archivística de documentos fotográficos a estándares internacionales. In: *Girona 2014: Archivos e Industrias Culturales*. AMGi/ICA. Acedido em 20 Junho de 2016 em: <http://www.girona.cat/sgdap/docs/qo4xhr0id164.pdf>

Schwartz, J. M. (1995). We make our tolls and our tools make us. Lessons from photography for the practice, politics and poetics of diplomacy. *Archivaria*, n. 40, pp. 40-74. Acedido em 12 de Junho de 2016 em: <http://archivaria.ca/index.php/archivaria/article/view/12096/13082>

Tagg, J. (2007) *El peso de la representación: ensayos sobre fotografías e historias*. Barcelona: Gustavo Gilli.



DESAFIOS À ORGANIZAÇÃO E ACESSO AO PATRIMÔNIO CULTURAL

DOCUMENTOS AUDIOVISUAIS, INFORMAÇÃO E MEMÓRIA: IDENTIFICAÇÃO DE ACERVOS FOTOGRÁFICOS E FÍLMICOS DAS REGIÕES ADMINISTRATIVAS DO DISTRITO FEDERAL DO BRASIL

Eliane Braga de Oliveira¹, Miriam Paula Manini², Sérgio Peçanha da Silva Coletto³

Universidade de Brasília

¹elianebo@unb.br, ²mpmanini@uol.com.br, ³sergiocoletto@gmail.com

RESUMO Andamento do projeto “Documentos audiovisuais, informação e memória: identificação de acervos fotográficos e fílmicos no Distrito Federal – 2ª Etapa”, que tem como objetivo mapear – qualificar e quantificar – a produção e a acumulação de registros audiovisuais (fotografias e filmes) em algumas cidades do Distrito Federal (DF) do Brasil. Trata-se da 2ª Etapa da pesquisa, que contempla o levantamento de arquivos fotográficos e fílmicos em instituições públicas do governo federal e distrital, além de acervos particulares de fotógrafos e produtores audiovisuais, abrangendo pela primeira vez as Regiões Administrativas (RAs). Enquanto a 1ª Etapa contemplou a pesquisa de campo em instituições localizadas na cidade de Brasília, a 2ª Etapa objetiva a expansão da iniciativa para as seguintes RAs do DF: Brazlândia, Ceilândia, Gama, Núcleo Bandeirante, Planaltina, Sobradinho e Taguatinga, escolhidas por sua tradição histórica e densidade populacional.

PALAVRAS-CHAVE *Brasília, Documentos audiovisuais, Informação, Memória, Regiões Administrativas.*

ABSTRACT Progress of the project “Audiovisual documents, information and memory: identification of photographic and film collections in the Federal District - Step 2” that aims to map – qualify and quantify – the production and accumulation of audiovisual records in cities of Brazil Federal District. This is the second stage of the research, which includes the collection of photographic and film archives in public institutions of the federal and district government, as well as private collections of photographers and audiovisual producers, covering for the first time the Administrative Regions. While a first stage includes a field research in institutions located in the city of Brasília, a second stage of implementation is at the Administrative Regions of the Federal District (Brazil): Brazlândia, Ceilândia, Gama, Núcleo Bandeirante, Planaltina, Sobradinho and Taguatinga, choices due its historical tradition and population density.

KEYWORDS *Administrative regions, Audiovisual documents, Brasília city, Information, Memory.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Este trabalho apresenta o andamento do projeto “Documentos audiovisuais, informação e memória: identificação de acervos fotográficos e fílmicos no Distrito Federal – 2ª Etapa”, que tem como objetivo

mapear – qualificar e quantificar – a produção e a acumulação de registros audiovisuais (fotografias e filmes) em algumas Regiões Administrativas (RAs) – antigas cidades satélites – do Distrito Federal (DF) do Brasil. Trata-se da 2ª Etapa da pesquisa, que contempla o levantamento de arquivos fotográficos e filmicos em instituições públicas do governo federal e distrital, além de acervos particulares de fotógrafos e produtores audiovisuais, abrangendo pela primeira vez as RAs. Enquanto a 1ª Etapa contemplou a pesquisa de campo em instituições localizadas na cidade de Brasília, a 2ª Etapa objetiva a expansão da iniciativa para as seguintes cidades: Brazlândia, Ceilândia, Gama, Núcleo Bandeirante, Planaltina, Sobradinho e Taguatinga, escolhidas por sua tradição histórica e densidade populacional.

Trata-se de uma iniciativa do Grupo de Pesquisa Imagem, Memória e Informação (IMI), cadastrado no Repositório de Grupos de Pesquisa do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)¹, composto por docentes, pesquisadores e estudantes de graduação e pós-graduação da Faculdade de Ciência da Informação da Universidade de Brasília (FCI/UnB).

Acervos de documentos fotográficos e filmicos custodiados em arquivos, museus, bibliotecas e centros de documentação são considerados objetos e veículos de memória individual e coletiva de uma sociedade. A reconhecida heterogeneidade cultural da população do DF, resultado da emigração de brasileiros de todas as regiões do país, encontra-se refletida de diversas maneiras nestes acervos, contribuindo para o fortalecimento da identidade regional e, conseqüentemente, nacional.

Nosso propósito é realizar um levantamento de grande abrangência, resultando num mapa o mais completo possível de acervos de imagens fixas e em movimento da região pesquisada, visto a necessidade e a importância destas informações para pesquisadores, acadêmicos, cineastas, fotógrafos e o público em geral; e seu ineditismo, obviamente.

A 1ª Etapa do Projeto aconteceu de novembro de 2013 a outubro de 2016 e mapeou a produção e a acumulação de registros audiovisuais (fotografias e filmes) em instituições públicas federais e distritais bem como acervos particulares de Brasília. As RAs do DF e cidades do entorno, por restrições orçamentárias, foram reservadas para etapas posteriores, como esta.

A fundamentação teórica de todas estas atividades de caráter prático revela os caminhos que se trilha até que se possa deparar com fotografias, filmes e a memória das pessoas e das localidades estudadas.

O primeiro grande tema que estudamos dentro do projeto é **imagem**: abordamos seu conceito, suas características e possibilidades que, em desenvolvimento, coadunem-se com as imagens técnicas que são a fotografia e o cinema.

A intenção é munir os envolvidos de material que lhes permita ver, compreender, analisar e descrever imagens fixas e em movimento, percebendo qual a sintaxe dessas linguagens visuais quase como um novo idioma, que requer método, tempo e dedicação para sua boa absorção e correto uso.

Como uma primeira vertente dos estudos sobre a imagem em geral são lidos textos sobre **fotografia**. Embora seja uma mensagem sem código por ser contínua – nada se interpõe entre a fotografia e o espectador –, a fotografia traz conteúdos, que são proporcionados pelo autor – o fotógrafo – e lidos pelo

¹ Espelho do Grupo de Pesquisa Imagem, Memória e Informação no Diretório de Grupos de Pesquisa do Brasil na plataforma Lattes/CNPq (Disponível em: <<http://dgp.cnpq.br/dgp/espelhogrupo/5886291772035636>>. Acesso em: 13 jun. 2017).

receptor – o público em geral. Desta forma, é aqui importante desvendar produção e recepção fotográfica e tudo que ambas envolvem.

A definição mais antiga de fotografia – após o significado etimológico *escrita com a luz* – diz ser ela um recorte de espaço da realidade num determinado momento (tempo). Este objeto que carrega um fato, coisa ou pessoa do passado – e cada clique tem seu passado imediatamente criado – insere-se instantaneamente na categoria de objeto de memória.

Pessoas, grupos, sociedades, povos inteiros poderão reconhecer numa fotografia um referente aurático de sua própria história. Na fotografia doméstica, é a memória familiar; na fotografia do mundo do trabalho, é a memória institucional; no fotojornalismo, é a memória social e política; na fotografia documental, é a memória histórica.

Nada marca melhor a aura-memória da fotografia do que o “isto foi”. Na esteira dos teóricos da modernidade, Barthes (1984 e 1990) sempre apontou que o sentido da imagem é o fotografado, o objeto fotográfico, estando o fotógrafo (como operador) em segundo plano, e o meio fotográfico também: esta é a fotografia documental. O objeto é o referente real; o “isto foi”, ou seja, algo da ordem da memória.

Uma das principais premissas da imagem fotográfica está relacionada ao seu caráter análogo com relação ao referente: aquilo que é fotografado existe/existiu, aquilo foi. O assunto, aqui, faz parada na Semiótica, cujo arcabouço permite pensar na fotografia como espelho do real (ícone), transformação do real (símbolo) e como traço do real (a prova de existência do referente).

Como indaga Ricoeur (2007, p. 61), seria a lembrança uma imagem que se faz do passado? A fotografia, no caso, sendo exatamente uma imagem que se faz do passado, é um objeto que pressupõe rememoração.

A consciência íntima que temos da passagem de tempo acaba sendo abalada e certificada pelo testemunho do objeto fotográfico. O efeito da imagem fotográfica sobre a memória é devastador. No exercício historiográfico, quando confrontamos dados históricos textuais com fotografias podemos corrigir a memória escrita e reformular aquilo que já se conhecia.

O discurso fotográfico pode ter iniciado como arte – e os pintores fizeram alarde, drama e furor – mas logo mostrou sua outra faceta: o testemunho, a prova, o documento – e aqui encontramos seu caráter indicial incontestável. Seja como expressão (arte, ficção) ou como documento (registro, memória, testemunho, prova), a fotografia passou por fases: a ênfase no objeto, a ênfase na linguagem, a ênfase no autor e, agora, estamos na ênfase na tecnologia. Isto certamente mudou a forma de ver imagens fotográficas, ou seja, alterou a forma de olhar.

Entretanto, a necessidade de se conhecer o dispositivo, a técnica e suas possibilidades sempre esteve presente em cada uma dessas fases, importando isto para dar a conhecer, por meio de códigos de leitura, a retórica da fotografia.

Importante será indagar as formas de produção e recepção da sociedade civil, do homem comum, e da sua relação com a fotografia tanto no campo familiar e do trabalho quanto na recepção de caráter artístico (exposições) e documental (acervos públicos). Dos registros de festas infantis e viagens de um passado recente pré-digital – quando caixas de camisa e de sapato serviam de guarda para a memória afetiva imagética familiar – à satisfação de necessidades informacionais em buscas a cartórios, arquivos e até dioceses: a Sociologia trará entendimento sobre a relação da sociedade com a imagem fotográfica, as ocultações e as revelações da vida cotidiana.

Também será importante desvendar o desenvolvimento estético e semiótico da fotografia. No início do século XX, o principal objetivo da fotografia era exprimir o belo do mundo, da natureza e das pessoas, sem qualquer intervenção do fotógrafo.

A noção de realismo foi modificada por uma fotografia que não queria apenas registrar a realidade, mas ser a forma como as coisas parecem aos olhos do mundo. Sob este aspecto, o fotógrafo é relator da realidade e não seu intérprete, crítico ou denunciador.

Por último, mas não menos importante, será fundamental reconhecer o caráter reprodutível que a fotografia demonstrou em seus primórdios, quando se revelou a técnica da cópia infinita pelo uso da matriz negativa, emblema da industrialização e da mecanização e embrião da vindoura revolução digital.

Com a mudança do paradigma científico – e, por consequência, do paradigma informacional – inicia-se uma transição epistemológica e tecnológica; a epistemologia conceitual debilita-se, tornando-se o fotográfico um campo heterogêneo e concreto ontológico (século XXI) em constante transformação. A oposição sempre existente entre arte e documentação se desenvolve e se fortalece. O século XXI descortina uma grande fragilidade da fotografia em servir de documento: um novo caráter subjetivo da memória emerge; surge mais um ponto nevrálgico para reflexões em torno de como agir em termos informacionais.

Resultado tecnológico primeiro da evolução fotográfica, o **cinema** é abordado em nossas pesquisas nas mesmas proporções. A já observada similaridade entre a narrativa de cinema e a experiência onírica tem sido utilizada, inclusive, como cine-terapia. Cinema e psicanálise, além de serem contemporâneos – enquanto os irmãos Lumière faziam suas primeiras exposições do cinematógrafo, Freud publicava seus *Estudos sobre a histeria* –, aproximam-se, atualmente, em torno da configuração do sujeito.

O encadeamento de imagens, a logicidade temporal de princípio, meio e fim – mesmo com os *flashbacks* do cinema e as experimentações mais recentes de deslocamento temporal narrativo – e a possibilidade sempre presente de relacionar passagens do filme com nossa vida particular fazem do cinema uma arte psicossocial por excelência.

Nesta vivência algumas vezes catártica, algo provoca interrogações: por que nos emocionamos com a exibição de determinados filmes, chegando mesmo a chorar? Por que, às vezes, muitas pessoas se emocionam com a mesma cena ou sequência? A Neurociência explica, mas a atenção, aqui, deve recair sobre o alcance do cinema enquanto elaboração, construção e reconhecimento da memória pelo indivíduo.

Será necessário escolher entre vários manuais de análise fílmica para apreender algumas ferramentas necessárias à real aquisição de efeitos da narrativa cinematográfica: encadeamento narrativo temporal, movimento de câmera, efeitos especiais, interpretação dos atores, cenas, sequências, e outros tantos detalhes técnicos, semióticos e estéticos da arte cinematográfica.

A Ciência da Informação é acionada em seguida para os estudos sobre **análise documentária de imagens fixas (fotografias) e em movimento (cinema)**; da mesma forma, é de fundamental importância adquirir conhecimento sobre **conservação física dos suportes documentais fotográficos e fílmicos**. Para que a equipe seja suficientemente conhecedora de dados sobre a região estudada, conhecer a **história das antigas cidades satélites de Brasília**, hoje denominadas Regiões Administrativas, é igualmente importante; esta parte histórica está sendo estudada juntamente com as questões que envolvem **memória individual e memória coletiva**. Costurar toda esta aquisição – e

aplicação – de conhecimentos é tarefa da parte de **metodologia**, que envolve leituras sobre pesquisa com imagens e aplicação de História Oral em entrevistas.

Desta forma, os principais autores que estão sendo estudados são:

- Laurent Gerverau, Donis A. Dondis, Régis Debray, Annie Duprat e Jacques Aumont para os estudos sobre imagens;
- Roland Barthes, Philippe Dubois, André Rouillé, Jean-Marie Schaeffer, Martine Joly, José de Souza Martins, Susan Sontag e Walter Benjamin para a parte que de fotografia;
- Jacques Aumont, Francis Vanoye, Anne Goliot-Lété e Jean-Claude Carrière para a parte que trata de cinema;
- Johanna Smit, Peter Homulos, Elaine Svenonius, Sara Shatford Layne, James M. Turner, José Antonio Moreiro González, Jesús Robledano Arillo e Felix Del Valle Gastaminza para a parte sobre tratamento da informação imagética fixa;
- Rosa Inês de Novais Cordeiro, Jérôme Bourdon, Clive Cochrane, Donna M. Romer e Alfonso L. Yepes para a parte sobre tratamento da informação imagética em movimento;
- Paulo Bertran, Luiz Sérgio Duarte, L. Fernando Tamanini, Adirson Vasconcelos, Alexandre Nonato, Luiz Ricardo Magalhães, Marcel Gautherot, Samuel Titan Junior e Kenneth Frampton sobre a história de Brasília e do DF;
- Andreas Huyssen, Beatriz Sarlo, Maurice Halbwachs, Iván Izquierdo, Paul Ricoeur, Danilo Santos de Miranda, Ramon Alberch Fugueras, Pierre Nora e Michael Pollak para as questões sobre memória;
- Tânia M. P. Müller, Wivian Weller, Nicolle Pfaff, Martins W. Bauer, George Gaskell, Carol Couture e Alessandro Portelli para a parte de metodologia;
- João Sócrates Oliveira, Sérgio Burgi, Sandra Baruki, Adriana Cox Hollós, Peter Mustardo, Luís Pavão, Clóvis Molinari, Mônica A. Kornis e Cinemateca Brasileira para estudos sobre preservação de suportes fotográficos e filmicos.

METODOLOGIA

A equipe da 2ª Etapa do Projeto “Documentos Audiovisuais, Informação e Memória” é composta por duas professoras orientadoras da FCI/UnB e um orientador externo, todos integrantes do Grupo de Pesquisa IMI. Assim como na 1ª Etapa, têm atuado na capacitação dos pesquisadores de campo, no acompanhamento da pesquisa bem como na compilação e análise de resultados.

A equipe recrutada em março de 2017 é composta por sete bolsistas de nível superior, com formação em Arquivologia, Biblioteconomia, Ciências Sociais e Museologia. Será também recrutado um pesquisador com formação em Ciência da Computação ou área correlata que será encarregado do processamento e difusão dos dados como, por exemplo, atualizar informações no *site* do Grupo de Pesquisa².

² <https://pesquisaimi.wordpress.com/>



Figura 1: Reunião de bolsistas da Fundação de Amparo à Pesquisa do Distrito Federal (FAP/DF) e integrantes do Grupo de Pesquisa IMI envolvidos com a 2ª Etapa da pesquisa “Documentos Audiovisuais, Informação e Memória”.

Foto: Marjorie Guedes.

No que tange ao custeio das bolsas, os bolsistas estão enquadrados na modalidade Apoio Técnico (AT) de nível superior (NS) e cadastrados na FAP/DF. Após a fase de recrutamento destes sete pesquisadores de campo, os mesmos passaram por processo de formação para as atividades do projeto em oficinas internas elaboradas pelo grupo de pesquisa, tais como *Objetos Audiovisuais*, *Uso do SPSS* e *Pesquisa de Campo e História Oral*. Esta capacitação foi mais intensa no primeiro mês de trabalho, mas continua durante toda a participação no Projeto, mesmo durante a ida a campo, pois são muitas as informações e as frentes de atuação. As atividades incluem leituras e discussões sobre os temas da pesquisa³, além de treinamentos para identificação de suportes audiovisuais e tabulação de dados no *software SPSS Statistics*. A formação inicial é seguida pela pesquisa de campo, na qual os acervos públicos e privados das referidas RAs estão sendo visitados pelos pesquisadores de campo e os dados levantados com a ajuda do ICD. O último mês de atuação dos pesquisadores de campo está reservado para a inserção e tabulação de dados.



Figura 2: Oficina sobre Objetos Audiovisuais aberta ao público da Universidade de Brasília ministrada pelo museólogo David Capelo de Carvalho, integrante do IMI e estudante de Iniciação Científica da 1ª Etapa, organizada para treinamento dos bolsistas FAP/DF envolvidos com a 2ª Etapa da pesquisa “Documentos Audiovisuais, Informação e Memória”.

Foto: Miriam Manini.

³ Ao final deste trabalho, incluímos uma bibliografia básica destes estudos introdutórios.

Enquanto contrapartida institucional exigida pela agência de fomento se apresenta a dedicação de duas orientadoras doutoras, integrantes do quadro docente da FCI/UnB, e um orientador externo, durante todo o projeto. Tem-se também enquanto contrapartida o uso de equipamentos e *softwares* da FCI/UnB para registro e tabulação dos dados. Por fim, a FCI também cederá seu Auditório, um laboratório de informática, duas salas de aula com projetores para exibição de filmes e áreas comuns de suas instalações físicas para a realização do evento de encerramento da 2ª Etapa do Projeto, durante cinco dias, previsto para dezembro de 2017.

No que se refere ao processo de tratamento dos dados, ele será realizado em uma base de dados digital criada no *software IBM SPSS Statistics*, vastamente utilizado em pesquisas nas Ciências Sociais e nas Ciências Sociais Aplicadas. A versão utilizada neste Projeto é a de número 22, de 2013, traduzida para o Português do Brasil.

Os pesquisadores de campo responsáveis pela coleta de dados se utilizam de entrevista pré-estruturada realizada pessoalmente com responsáveis pelos acervos. Os entrevistados devem fornecer informações para que os pesquisadores de campo organizem a informação coletada. Esta ação é guiada pelo Instrumento de Coleta de Dados (ICD), um roteiro impresso composto de 19 perguntas de preenchimento manual sob inteira responsabilidade do pesquisador de campo durante e imediatamente após a entrevista.

As perguntas podem ser respondidas de cinco maneiras diferentes: 1. Com texto de até 500 caracteres; 2. Com sim OU não OU não sabe/não respondeu OU não se aplica; 3. Com uma sequência de até oito caracteres numéricos; 4. Com uma faixa de porcentagem: de 0 a 25% OU de 26 a 50% OU de 51 a 75% OU de 76 a 100% OU não sabe/ não respondeu OU não se aplica; 5. Com o estado de conservação de um determinado acervo: muito bom OU bom OU deteriorado OU desconhecido OU não sabe/não respondeu OU não se aplica.

Nome	Tipo	Largura	Decimais	Rótulo	Valores	Ausente	Colunas	Alinhar	Medir	Função
Q_6_Filmes	Númérico	8	0	Existem filmes na sua unidade administrativa?	(0, Sim)...	999	8	Direto	Ordinal	Entrada
Q_6.1_Filmes_PB	Númérico	8	0	Existem filmes preto-e-branco na sua unidade administrati.	(0, Sim)...	999	8	Direto	Ordinal	Entrada
Q_6.2_Filmes_cor	Númérico	8	0	Existem filmes coloridos na sua unidade administrativa?	(0, Sim)...	999	8	Direto	Ordinal	Entrada
Q_6.3_Películas_de_cinema	Númérico	8	0	Existem películas de cinema na sua unidade administrativa?	(0, Sim)...	999	8	Direto	Ordinal	Entrada
Q_6.4_Películas_35mm	Númérico	8	0	Existem películas 35mm na sua unidade administrativa?	(0, Sim)...	999	8	Direto	Ordinal	Entrada
Q_6.5_Películas_35mm_qtdde	Sequência	500	0	Caso positivo, especifique a quantidade	Nenhum	Nenhum	8	Esquerdo	Nominal	Entrada
Q_6.5.1_Películas_35mm_horas	Númérico	8	0	Qual a quantidade estimada de horas?	Nenhum	Nenhum	8	Direto	Ordinal	Entrada
Q_6.6_Películas_16mm	Númérico	8	0	Existem películas 16mm na sua unidade administrativa?	(0, Sim)...	999	8	Direto	Ordinal	Entrada
Q_6.7_Películas_16mm_qtdde	Sequência	500	0	Caso positivo, especifique a quantidade	Nenhum	Nenhum	8	Esquerdo	Nominal	Entrada
Q_6.7.1_Películas_16mm_horas	Númérico	8	0	Qual a quantidade estimada de horas?	Nenhum	Nenhum	8	Direto	Ordinal	Entrada
Q_6.8_Películas_8mm	Númérico	8	0	Existem películas 8mm na sua unidade administrativa?	(0, Sim)...	999	8	Direto	Ordinal	Entrada
Q_6.9_Películas_8mm_qtdde	Sequência	500	0	Caso positivo, especifique a quantidade	Nenhum	Nenhum	8	Esquerdo	Nominal	Entrada
Q_6.9.1_Películas_8mm_horas	Númérico	8	0	Qual a quantidade estimada de horas?	Nenhum	Nenhum	8	Direto	Ordinal	Entrada
Q_6.10_Fitas_de_vídeo	Númérico	8	0	Existem fitas de vídeo na sua unidade administrativa?	(0, Sim)...	999	8	Direto	Ordinal	Entrada
Q_6.11_Fitas_VHS	Númérico	8	0	Existem fitas de vídeo VHS na sua unidade administrativa?	(0, Sim)...	999	8	Direto	Ordinal	Entrada
Q_6.12_Fitas_VHS_qtdde	Sequência	500	0	Caso positivo, especifique a quantidade	Nenhum	Nenhum	8	Esquerdo	Nominal	Entrada
Q_6.12.1_Fitas_VHS_horas	Númérico	8	0	Qual a quantidade estimada de horas?	Nenhum	Nenhum	8	Direto	Ordinal	Entrada
Q_6.13_Fitas_Betamax	Númérico	8	0	Existem fitas de vídeo Betamax na sua unidade administr.	(0, Sim)...	999	8	Direto	Ordinal	Entrada
Q_6.14_Fitas_Betamax_qtdde	Sequência	500	0	Caso positivo, especifique a quantidade	Nenhum	Nenhum	8	Esquerdo	Nominal	Entrada
Q_6.14.1_Fitas_Betamax_horas	Númérico	8	0	Qual a quantidade estimada de horas?	Nenhum	Nenhum	8	Direto	Ordinal	Entrada
Q_6.15_Fitas_Umatic	Númérico	8	0	Existem fitas de vídeo U-Matic na sua unidade administrat.	(0, Sim)...	999	8	Direto	Ordinal	Entrada
Q_6.15.1_Fitas_Umatic_qtdde	Sequência	500	0	Caso positivo, especifique a quantidade	Nenhum	Nenhum	8	Esquerdo	Nominal	Entrada
Q_6.16.1_Fitas_Umatic_horas	Númérico	8	0	Qual a quantidade estimada de horas?	Nenhum	Nenhum	8	Direto	Ordinal	Entrada
Q_6.17_Fitas_Betacam	Númérico	8	0	Existem fitas de vídeo Betacam na sua unidade administr.	(0, Sim)...	999	8	Direto	Ordinal	Entrada
Q_6.18_Fitas_Betacam_qtdde	Sequência	500	0	Caso positivo, especifique a quantidade	Nenhum	Nenhum	8	Esquerdo	Nominal	Entrada
Q_6.18.1_Fitas_Betacam_horas	Númérico	8	0	Qual a quantidade estimada de horas?	Nenhum	Nenhum	8	Direto	Ordinal	Entrada
Q_6.19_Fitas_Otras	Númérico	8	0	Existem outros tipos de vídeos na sua unidade administrat.	(0, Sim)...	999	8	Direto	Ordinal	Entrada
Q_6.20_Fitas_Otras_tipos	Sequência	500	0	Caso positivo, especifique os tipos	Nenhum	Nenhum	8	Esquerdo	Nominal	Entrada
Q_6.21_Fitas_Otras_qtdde	Sequência	500	0	Caso positivo, especifique a(s) quantidade(s)	Nenhum	Nenhum	8	Esquerdo	Nominal	Entrada
Q_6.21.1_Fitas_Otras_horas	Númérico	8	0	Qual a quantidade estimada de horas?	Nenhum	Nenhum	8	Direto	Ordinal	Entrada
Q_6.22_Filmes_digitais	Númérico	8	0	Existem filmes em mídias digitais na sua unidade adminis.	(0, Sim)...	999	8	Direto	Ordinal	Entrada
Q_6.23_Filmes_digitais_DVD	Númérico	8	0	Existem filmes em DVD na sua unidade administrativa?	(0, Sim)...	999	8	Direto	Ordinal	Entrada
Q_6.24_Filmes_digitais_DVD_qtdde	Sequência	500	0	Caso positivo, especifique a quantidade	Nenhum	Nenhum	8	Esquerdo	Nominal	Entrada
Q_6.24.1_Filmes_digitais_DVD_horas	Númérico	8	0	Qual a quantidade estimada de horas?	Nenhum	Nenhum	8	Direto	Ordinal	Entrada
Q_6.25_Filmes_digitais_cartão	Númérico	8	0	Existem filmes em cartão de memória na sua unidade ad.	(0, Sim)...	999	8	Direto	Ordinal	Entrada
Q_6.26_Filmes_digitais_cartão_qtdde	Sequência	500	0	Caso positivo, especifique a quantidade	Nenhum	Nenhum	8	Esquerdo	Nominal	Entrada
Q_6.26.1_Filmes_digitais_cartão_horas	Númérico	8	0	Qual a quantidade estimada de horas?	Nenhum	Nenhum	8	Direto	Ordinal	Entrada
Q_6.27_Filmes_digitais_MinDV	Númérico	8	0	Existem filmes em fita MinDV na sua unidade administrati.	(0, Sim)...	999	8	Direto	Ordinal	Entrada
Q_6.28_Filmes_digitais_MinDV_qtdde	Sequência	500	0	Caso positivo, especifique a quantidade	Nenhum	999	8	Esquerdo	Nominal	Entrada

Figura 3: Interface do *software IBM Statistics* com partes das variáveis da questão nº 6.

Foto: Sérgio Coletto.

A criação da base de dados foi realizada por um dos orientadores do Projeto. Busca uma fidedignidade absoluta com o ICD ao migrar dados quantitativos e qualitativos. Assim também o é com o

processamento dos dados coletados. Para isso, as versões das bases de dados utilizadas são controladas. Nesta 2ª Etapa realizamos as primeiras melhorias.

A migração dos dados do ICD preenchido para a base de dados pode ser realizada pelo próprio pesquisador de campo – com revisão obrigatória de um orientador – ou diretamente por um orientador – em comunicação direta virtual ou presencial com o pesquisador de campo. Contribui-se, desta maneira, para a garantia da qualidade dos dados trabalhados, além de permitir que eventuais dificuldades de interpretação sejam facilmente resolvidas.

RESULTADOS

Este Projeto de Pesquisa tem como objetivo mapear – qualificar e quantificar – a produção e a acumulação de registros audiovisuais (fotografias e filmes) no DF. Trata-se da 2ª Etapa da pesquisa, que contempla o levantamento de arquivos fotográficos e filmicos em instituições públicas do governo federal e distrital, além de acervos particulares de fotógrafos e produtores audiovisuais, abrangendo pela primeira vez as RAs.

Enquanto a 1ª Etapa contemplou a pesquisa de campo em instituições localizadas na cidade de Brasília, a 2ª Etapa objetiva a expansão da iniciativa para as seguintes RAs do DF: Brazlândia, Ceilândia, Gama, Núcleo Bandeirante, Planaltina, Sobradinho e Taguatinga, escolhidas por sua tradição histórica e densidade populacional.

São considerados objetivos específicos do projeto:

- Proceder a um levantamento dos acervos fotográficos (fotos, negativos, diapositivos) e filmicos (cinema, vídeo, DVD) existentes em arquivos, bibliotecas, museus e centros de documentação, informação e memória – públicos e privados – da região escolhida;
- Identificar acervos particulares de documentos fotográficos e filmicos relacionados à memória local;
- Diagnosticar o estado físico de conservação bem como do tratamento do conteúdo informacional dos itens encontrados;
- Observar como os detentores de acervos e profissionais das instituições visitadas colaboram com a salvaguarda dos materiais mapeados.

Baseados nestes objetivos, esperamos os resultados elencados a seguir: enriquecimento de banco de dados sobre os acervos imagéticos no DF, originalmente criado na 1ª Etapa; ampliação do mapa dos acervos imagéticos do DF, também originalmente criado na 1ª Etapa; apresentação de trabalhos em eventos da área de Ciência da Informação, tal qual na 1ª Etapa; submissão de artigos científicos para publicação em periódicos, tal qual na 1ª Etapa; realização de evento acadêmico para divulgação de resultados, tal qual na 1ª Etapa; atualização de Instrumento de Coleta de Dados (ICD) e de banco de dados; elaboração de produto de informação e comunicação em formato digital e interativo, conforme padrão estabelecido pela agência de fomento (FAP/DF); apresentação dos resultados obtidos em Seminário de avaliação promovido pela agência de fomento; estabelecimento de um maior relacionamento com a Associação Brasileira de Preservação Audiovisual (ABPA); formalização de parceria com o Grupo de Estudos sobre Cultura, Representação e Informação Digitais (CRIDI), da

UFBA⁴; e formalização de parceria com a Universidade Complutense de Madri por meio de projeto conjunto para configurar a Rede Hispano-Brasileira de Patrimônio Audiovisual em Arquivos, Bibliotecas e Serviços Universitários.

CONCLUSÕES

O Projeto “Documentos audiovisuais, informação e memória: identificação de acervos fotográficos e filmicos no Distrito Federal – 2ª Etapa” promove a memória enquanto bem inalienável de uma sociedade. Para isso, utiliza-se da disseminação da informação e da geração de conhecimento sobre objetos que porventura estejam “esquecidos”.

A identificação dos acervos localizados nas RAs do DF ainda está em andamento. No entanto, já é possível perceber características específicas que apontam para a participação mais ativa dos grupos organizados da sociedade civil na custódia e preservação dos registros, ao contrário do que foi identificado na 1ª Etapa, na cidade de Brasília, onde as instituições públicas assumem o protagonismo nessas atividades.

Pessoas, grupos, sociedades, povos inteiros poderão reconhecer nos documentos audiovisuais um referencial de sua própria história, riqueza cultural incontestável de valor inestimável para o Distrito Federal e demais regiões do país.

REFERÊNCIAS BIBLIOGRÁFICAS

Barthes, R. (1984). *A câmara clara: nota sobre a fotografia*. Rio de Janeiro: Nova Fronteira.

Barthes, R. (1990). *O óbvio e o obtuso: ensaios críticos III*. Rio de Janeiro: Nova Fronteira.

Ricoeur, P. (2007). *A memória, a História, o esquecimento*. Campinas: UNICAMP.

BIBLIOGRAFIA

Benjamin, W. (1987). A obra de arte na era de sua reprodutibilidade técnica. *Obras escolhidas: magia e técnica, arte e política* (3ª ed) (p. 165-196). São Paulo: Brasiliense.

Bourdieu, P. (2010). *Un art moyen: essai sur les usages sociaux de la photographie* (2ª ed.). Paris: Minuit.

Draaisma, D. (2005). *Metáforas da memória: uma história das idéias sobre a mente*. Bauru: Edusc.

⁴ Espelho do Grupo de Estudos sobre Cultura, Representação e Informação Digitais no Diretório de Grupos de Pesquisa do Brasil na plataforma Lattes/CNPq. (Disponível em: <<http://dgp.cnpq.br/dgp/espelhogrupo/6505066926981542#identificacao>>. Acesso em: 16 jun. 2017).

ESTEVAM, L. *O tempo da transformação: estrutura e dinâmica da formação econômica de Goiás*. Goiânia: Editora UCG.

Freund, G. (1974). *Photographie et société*. Paris: Seuil.

Halbwachs, M. (2004). *A memória coletiva*. São Paulo: Centauro.

Magalhães, L. R. *Sertão planaltino: uma outra história de Brasília*. Curitiba: CRV.

Manini, M. P. (2011). Imagem, memória e informação: um tripé para o documento fotográfico. *Domínios da Imagem*, 4(8), 77-87. Disponível em <http://www.uel.br/revistas/uel/index.php/dominiosdaimagem/article/viewFile/23354/17054>

Manini, M. P. et al. (2016). Documentos audiovisuais, informação e memória: resultados da identificação de acervos fotográficos e filmicos em Brasília. *Encontro Nacional de Pesquisa em Ciência da Informação*, XVII. Salvador/BA. Disponível em <http://www.ufpb.br/evento/lti/ocs/index.php/enancib2016/enancib2016/paper/viewFile/4041/2367>

Manini, M. P.; Oliveira, E. B.; Carrijo, E. (2016). Acervos audiovisuais da universidade de Brasília: imagem, memória e informação. *Seminário Hispano Brasileiro de Pesquisa em Informação, Documentação e Sociedade*, V. Madri/Espanha.

Nora, P. (1993). Entre memória e história. A problemática dos lugares. *Revista do Programa de Estudos Pós-graduados em História e do Departamento de História*, 10, 7-28.

Oliveira, A. G. (2013). *Preservação de acervos filmicos do Distrito Federal* (Dissertação de Mestrado, Faculdade de Ciência da Informação, UnB). Disponível em http://repositorio.unb.br/bitstream/10482/13540/1/2013_Ang%C3%A9licaGasparottodeOliveira.pdf



MEMÓRIA E PATRIMÔNIO CULTURAL: O MUSEU DO CEARÁ COMO LUGAR DE MEMÓRIA

Fátima Portella Cysne¹, Maria Cleide Rodrigues Bernardino², Carla Façanha de Brito³,
Adriana Nóbrega Silva⁴, Denise Marques Rodrigues⁵

¹UNILAB, fatimaportela@unilab.edu.br

²UFCA, cleide.rodrigues@ufca.edu.br

³UFCA, carla.facanha@ufca.edu.br

⁴UFC, adriufc.nobrega@gmail.com

⁵UFC, denisemarques.r@gmail.com

RESUMO Aborda o Museu do Ceará como um dos principais lugares de memória presentes na cidade de Fortaleza. Apresenta como problema de pesquisa ao refletir o Museu do Ceará como lugar de memória, e suas lacunas como gestor e fomentador da história local. A partir desse questionamento, traçaram-se alguns objetivos: identificar as políticas do Museu do Ceará para incentivar na comunidade o interesse pela história local; verificar se existem ações propostas pelo Museu do Ceará que despertem o interesse pelo mesmo ou pelos lugares de memória; avaliar a eficácia das propostas de atração ao Museu do Ceará, se as mesmas existirem e verificar a importância dos lugares de memória para a construção da identidade e da cultura fortalezense. Usou-se a pesquisa exploratória e descritiva, estudo de caso e o questionário como instrumento de coleta de dados. Foi possível constatar nos resultados que o Museu do Ceará é um lugar que traz subsídios para compreensão da história local, mas apresenta deficiências no quesito atração de novos visitantes. Concluímos que o Museu do Ceará cumpre o seu papel ao fornecer subsídios necessários para a compreensão da história do Estado do Ceará, mais especificamente, da cidade de Fortaleza. Ele exerce todas as exigências determinadas pela Legislação que rege os museus; apesar da falha da disponibilização do profissional, guia de visitas. O museu é capaz de comunicar a história local de forma sintética e o seu próprio espaço físico é atrativo aos olhos curiosos, porém deixa a desejar no quesito autopromoção/divulgação.

PALAVRAS-CHAVE *Memória, Patrimônio Cultural, Museu do Ceará, História Local*

ABSTRACT It approaches the Museum of Ceará as one of the main places of memory present in the city of Fortaleza. It presents as a research problem when reflecting the Museum of Ceará as a place of memory, and its gaps as manager and promoter of local history. Based on this questioning, some objectives were drawn: to identify the policies of the Museum of Ceará to encourage the interest in local history in the community; To verify if there are actions proposed by the Museum of Ceará that arouse the interest for the same or the places of memory; To evaluate the effectiveness of the proposals of attraction to the Museum of Ceará, if they exist and to verify the importance of places of memory for the construction of the identity and the strengthen culture. We used exploratory and descriptive research, a case study and the questionnaire as an instrument for data collection. It was possible to verify in the results that the Museum of Ceará is a place that brings subsidies to understand the local history, but presents deficiencies in the attraction of new visitors. We conclude that the Ceará Museum fulfills its role by providing the necessary subsidies for understanding the history of the State of Ceará, more specifically, the city of Fortaleza. It exercises all the requirements determined by the Legislation governing museums; Despite the lack of availability of the professional guide of visits. The museum is able to communicate local history in a synthetic

way and its own physical space is attractive to the curious eyes, but it leaves something to be desired in terms of self-promotion.

KEY-WORDS *Memory, Cultural Heritage, Museum of Ceará, Local History.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A explosão informacional sentida na sociedade como um dos fenômenos do processo de globalização traz algumas consequências. Uma das consequências é a preservação da memória. Le Goff (2003) afirma que “a memória, como propriedade de conservar certas informações, remete-nos em primeiro lugar a um conjunto de funções psíquicas, graças às quais o homem pode atualizar impressões ou informações passadas, ou que ele representa como passadas” (p. 419). Tal conceito se perde atualmente em meio a quantidade de informações presentes que acabam por se tornar obsoletas com maior frequência perdendo muitas vezes o seu valor de retenção. Para tanto, tornam-se necessários à presença de lugares de memória (Le Goff, 2003), ou seja, ambientes físicos, materiais ou imateriais onde a memória de uma sociedade se presentifica.

Os lugares de memória são responsáveis pela construção da cultura, para a constituição do saber e para a formação de uma identidade.

O Museu do Ceará é um dos principais lugares de representação do patrimônio cultural da cidade de Fortaleza. Surgido a partir do Arquivo Público do Ceará, contém registros dos principais acontecimentos ocorridos desde os primórdios da existência do Estado. (Secretaria da Cultura, 2011). Neste sentido, apresenta-se o problema de pesquisa a partir do questionamento: O Museu do Ceará, como lugar de memória, exerce a função de espaço fomentador do patrimônio cultural não somente preservando a informação extrínseca e intrínseca dos objetos, mas comunicando aos seus visitantes a história local?

A relevância desta temática dá-se pela afirmação que os lugares de memória são verdadeiros preservadores da cultura local, além de auxiliar na construção de uma identidade do povo fortalezense. A pesquisa apresenta relevância acadêmica para os cursos de Biblioteconomia do estado do Ceará, uma vez que a mesma tem como objeto de estudo a informação registrada; e os museus, como lugares responsáveis por preservar e comunicar documentos apresentam campo propício ao estudo em lide.

Tem como objetivos identificar as políticas do Museu do Ceará para incentivar na comunidade o interesse pela história local; verificar se existem ações propostas pelo Museu do Ceará que despertem o interesse pelo mesmo ou pelos lugares de memória; avaliar a eficácia das propostas de atração ao Museu do Ceará, se as mesmas existirem e verificar a importância dos lugares de memória para a construção da identidade e da cultura fortalezense.

MEMÓRIA E PATRIMÔNIO CULTURAL: O MUSEU DO CEARÁ

A palavra memória tem sua origem do latim *memoria* e significa a capacidade de reter e/ou readquirir ideias, imagens, expressões e conhecimentos adquiridos anteriormente tendo relação com as lembranças.

De acordo com Le Goff (2003),

[...] a memória, como propriedade de conservar certas informações, remete-nos em primeiro lugar ao conjunto de funções psíquicas, graças às quais o homem pode atualizar impressões ou informações passadas, ou que ele representa como passadas (p. 419).

A memória não se trata apenas de um processo mental, mas também espacial e temporal, além de envolver atividades mecânicas e intelectuais. De acordo com Le Goff (2003) “certos aspectos do estudo da memória, [...], podem evocar, de forma metafórica ou concreta, traços e problemas da memória histórica e da memória social” (p. 420). E Pollak (1992, p. 2) afirma que, “a priori, a memória parece ser um fenômeno individual, algo relativamente íntimo, próprio da pessoa. [...] a memória deve ser entendida também, ou sobretudo, como um fenômeno coletivo e social, ou seja, como um fenômeno construído coletivamente e submetido a flutuações, transformações, mudanças constantes”.

Assim, no tocante a realidade apresentada no espaço referente a museus, tecemos aqui um diálogo com a memória e o papel dos testemunhos, elementos esses importantíssimos para entendermos o processo de rememoração, de construção e representação que os objetos musealizados carregam em sua essência. Pois “para confirmar ou recordar uma lembrança, não são necessários testemunhos no sentido literal da palavra, ou seja, indivíduos presentes sob uma forma material e sensível”. (Halbwachs, 2006, p. 31). Entendemos que todo esse processo de representação apoiado nas possibilidades trazidas pela memória - fonte única na qual se podem representar as coisas do passado- tem uma forte ligação com a instância dos museus, ao servir de espaço de interação e troca de lembranças que se transformam em memórias, primordialmente individuais, mas, que se tornam coletivas não só por intermédio do compartilhamento destas, ao se colocarem materializadas pelos objetos proposto pelos espaços de representação, como os museus, mas em sua própria construção interna, trazidas ao presente pelo ato de rememoração, dos fatos vividos.

A memória pode ser vista não só como uma ferramenta de guardar dados mnemônicos, mas, sobretudo, como uma capacidade de (re) significação das coisas e de si mesmo, trata-se de uma representação das coisas vistas, vivenciadas do passado. A esse processo de trazer ao presente lembranças dos lugares de nossa memória, de uma possível reconfiguração de fatos guardados na memória que são despertados pela rememoração, lugares esses, como afirma Pierre Nora (1993), que se estendem além do espaço minimizado pela materialização, relacionados a um plano abstrato.

Na afirmação cunhada por Aristóteles de que “a memória é do passado” contemplamos o desejo de reconhecimento de uma coisa ausente, esse ato de reconhecimento encarado pelo processo de rememoração que ao ser evocado traz ao presente representações de coisas ausentes que se configuram em testemunhos, imagens e objetos. Esse exercício da memória evoca a imagem de um passado, sendo esta, a verdadeira presentificação desse passado, assegurando o caráter legítimo da memória.

Deve-se, porém ter a convicção que tais lugares se fazem necessários, pois nossa capacidade cerebral não nos permite guardar de tudo o que vivemos, sendo estes lugares responsáveis por auxiliar nas recordações e na construção, principalmente de uma memória coletiva.

Nora (1993, p. 13) conclui que:

Os lugares de memória nascem e vivem do sentimento que não existe memória espontânea, que é preciso criar arquivos, que é preciso manter os aniversários, organizar as celebrações, pronunciar as honras fúnebres, estabelecer contratos, porque estas operações não são naturais [...]. Se vivêssemos verdadeiramente as lembranças que eles envolvem, eles seriam inúteis. E se em compensação, a história não se apoderasse deles para deformá-los, transformá-los, sová-los e petrificá-los eles não se tornariam lugares de memória. É este vai-e-vem que os constitui: momentos de história arrancados do movimento de história, mas que lhe são devolvidos [...].

Os lugares de memória são espaços que comunicam bens materiais e imateriais que auxiliam na construção de uma identidade social, uma vez que essa mesma identidade se valida por uma memória coletiva e individual, simultaneamente. Peter Burke *apud* Ruoso (2009) diz que “os museus são muito mais do que meros recipientes para os objetos nele exibidos [...] tem sua própria história e podem nos dizer muito sobre [uma] época [...]” (p. 14). Ainda nesse contexto Brito (2017) afirma que “Entre as instâncias produtoras de conhecimento destacam-se os Museus. E este espaço [...] é reconhecido como local de transformação, onde as lembranças se tornam representadas por figuras, sons, imagens e objetos, lugar de legitimação dessas memórias.” (p. 58, no prelo).

Ao longo da história da humanidade, os museus se constituem, por exemplo, como espaços de memória científica, sócio-cultural, religiosa da sociedade, independentemente de suas especificidades. Os museus são necessariamente espaços de comunicação, como bem diz, Horta (1994, p.10) em sua análise semiótica do museu. Para essa estudiosa, os museus não se constituem apenas instituições, porém, “[...] mas como um meio, um instrumento, um sistema de comunicação, com uma estrutura flexível e mutante como a da linguagem que se apóia em um novo conceito do objeto museal”. Essa estudiosa diz ainda que, o processo de comunicação dos museus “implica o uso de diferentes códigos e sistemas semióticos, que vão atuar simultaneamente sobre os receptores”.

Os museus devem fazer parte da construção do imaginário social e da memória coletiva uma vez que “podemos tecer compreensões de aspectos da vida social, do comportamento humano, das vontades de memória, das sensibilidades e das mentalidades de um lugar e de um tempo” (Ruosso, 2009, p. 25). Porém, não devemos esquecer que os museus também são instituições responsáveis pela educação, no processo de aprendizagem e construção de conhecimento, pois “a instituição museal tem sob sua responsabilidade um acervo e no exercício das suas competências deve garantir a sua salvaguarda, além de promover a produção de conhecimento” (Ruosso, 2009, p. 29).

Diante de um mundo cercado de mudanças que tornam as coisas obsoletas cada vez mais rápido e a informação que atinge um número cada vez maior de pessoas num curto período de tempo, devemos pensar os museus como, “instituições muitas vezes voltadas ainda para a preservação e a guarda de objetos retirados do tempo e do espaço que os originaram, ou mesmo o acervo museológico, silencioso, estranho e vinculado ao culto do que é autêntico e original” (Santos, 2006, p. 16).

O Museu do Ceará como uma instituição de memória possui um dos maiores acervos do Estado. Passou por grandes mudanças antes de se tornar o que é atualmente. Foram inúmeras sedes e nomes dados ao museu. Foi criado no ano 1932 e recebeu o nome de Museu Histórico do Ceará. Em 1955, devido às novas aquisições, assim, o museu recebeu o nome de Museu Histórico e Antropológico do Ceará.

O Museu, espaço de significações, é um lugar instigante, incitador e um estímulo para a construção e o registro da memória coletiva no permanente processo que lhe é peculiar, onde se manifestam os

testemunhos das lembranças, visto que o Museu, como território para criar e produzir sentidos, configura-se em local no qual os testemunhos, a voz dos fatos passados ou recriados se encontram em uma troca de saberes individuais e coletivos.

Em se tratando de Museu/Patrimônio, Memória Social, Tradição/Manifestação Cultural, o Museu do Ceará conta com um acervo de mais de 7 mil peças, adquiridas entre compras e doações. E constitui a história do Estado do Ceará, retratando os fatos importantes desde os primórdios do Estado, bem como da cidade de Fortaleza que situa a necessidade de Preservação e Comunicação no seio de uma sociedade carente de uma história local que represente a face de sua realidade.

METODOLOGIA

Tentamos apresentar nesta pesquisa um arcabouço teórico-empírico para embasar os questionamentos e os cruzamentos de dados que foram realizados a partir da pesquisa, na tentativa de obter resultados confiáveis para a análise do estudo de caso.

A metodologia utilizada foi caracterizada como exploratória e descritiva. Adotou como método científico o procedimento de estudo de caso aplicado na instituição Museu do Ceará.

As principais finalidades da pesquisa exploratória, de acordo com Tripodi, Fellin e Meyer (1981, p. 64), sintetizam em “desenvolver, esclarecer e modificar conceitos e ideias, a fim de fornecer hipóteses pesquisáveis para estudos posteriores”. Busca obter conhecimento e compreensão a respeito de algo pouco explorado quer seja referente a uma temática específica, ou pela maneira diferenciada em observar um determinado fenômeno, de modo a torná-lo mais evidente, e arquitetar o fundamento conceitual necessário ao desenvolvimento do modelo proposto pela investigação. E a descritiva tende a observar, registrar e analisar, e correlacionando fatos, fenômenos, ou mesmo situações, sem a pretensão de manipulá-los (TRIVIÑOS, 1987).

O delineamento da pesquisa foi um estudo de caso com uma pesquisa bibliográfica, onde foi realizado um levantamento da produção científica sobre o tema, em que foram selecionadas as produções relevantes para a pesquisa. O delineamento considera o ambiente no qual são coletados os dados e as formas de controle das variáveis envolvidas. A pesquisa apresentou o seguinte delineamento:

A população objeto da pesquisa foi constituída pelos funcionários do Museu do Ceará. De acordo com Minayo (2010), no estudo de uma população, a amostra é um recurso amplamente utilizado, pois os levantamentos por amostragem permitem a aplicação de procedimentos de inferência estatística, os quais propiciam que os dados analisados possam ser estendidos e validados com segurança para a população.

A definição da amostragem foi a seguinte: quatro funcionários que se submeteram à responder os questionários ligados ao setor administrativo da instituição.

Pelo fato de haveremos optado pelo método do estudo de caso e consoante o que aponta a literatura especializada, conseguimos evidências através da coleta de dados pelo o instrumento de pesquisa questionário.

A escolha do estudo de caso ocorreu por considerarmos, como indica Minayo (2010, p.102), partindo de uma perspectiva qualitativa, a preocupação se dá "menos com a generalização e mais com o

aprofundamento e abrangência da compreensão". “[...] fatos ou resumos narrativos de situações ocorridas em empresas, órgão público ou em outras instituições [...]”. Usado inicialmente como técnica na área de saúde, o estudo de caso é amplamente empregado pela Administração.

Na intelecção de Godoy (1995), o estudo de caso visa ao exame detalhado de um ambiente, de um sujeito ou de uma situação em particular, partindo de focos de interesse mais amplos, de tal forma que as abstrações e o quadro teórico vão se consolidando à medida que os dados são coletados e examinados.

O instrumento principal utilizado para a coleta de dados nessa pesquisa, foi o questionário, que de acordo com Gil (1991), é um instrumento de investigação que visa recolher informações baseando-se, geralmente, na indagação de um grupo representativo da população em estudo. Para tal, coloca-se uma série de questões que abrangem um tema de interesse para os investigadores, não havendo interação direta entre estes e os inquiridos.

Assim, os dados coletados foram analisados de forma qualitativa e quantitativa, buscando-se classificar as informações com base em aspectos previamente coletados, de acordo com a relevância dos questionamentos construídos mediante base teórico-empírica.

RESULTADOS

A fim de cumprir os objetivos da pesquisa, os questionários foram aplicados com os cargos de chefia ligados à administração e ações educativas do Museu do Ceará, por se tratarem das pessoas que realizam o planejamento e avaliação da instituição. O questionário foi composto de dez perguntas com perguntas fechadas de múltipla escolha.

Pode-se afirmar apesar das lacunas encontradas que o Museu do Ceará tem papel relevante como um agente social institucional que agrega conceito e característica aos objetos em exposição e atua mediando o processo cognitivo de seleção da memória social.

A análise conclui que independente da maneira como o Museu do Ceará se apresenta atualmente para a sociedade, ainda consegue atrair visitantes diariamente. O perfil do público visitante é em 75% composto por escolas de ensino fundamental e médio e universidades e 25% de um público variado.

Aponta como fator negativo a falta de políticas no que diz respeito às novas exposições e ao marketing promocional da instituição. Entretanto, os respondentes sinalizam várias ações realizadas no Museu e que contribui, de forma significativa, para o aumento do número de visitantes, o que se configura como uma estratégia de estabelecer um relacionamento com a comunidade usuária e contribuir para a educação do cearense no que diz respeito a história local.

A respeito do retorno dos visitantes, os respondentes afirmaram em 50% “não sabemos informar”, 25% que “a maioria visita apenas uma vez” e os outros 25% responderam “a maioria visita duas vezes pelo menos”, isso dificulta no cumprimento do Art. 36 da Subseção III da Legislação sobre museus: “As estatísticas de visitantes dos museus serão enviadas ao órgão ou entidade competente do poder público, na forma fixada pela perspectiva entidade, quando solicitadas” (2013, p. 33).

O que se percebe neste item é que não há uma preocupação do Museu em saber se os visitantes saem satisfeitos com a exposição. As respostas apresentadas confirmam uma imprecisão em saber se os visitantes desejam voltar ao museu ou não.

O Museu do Ceará é assertivo ao disponibilizar um guia de visitas, mas age de maneira errônea ao que diz respeito ao agendamento de uma visita orientada, pois mesmo que o museu conte com um público distinto que variam entre querer visitar o museu independentemente e querer visitar o museu com o auxílio de um guia, ele deve estar preparado para ocasionalidades e/ou imprevistos, como por exemplo, turistas estrangeiros, turistas nacionais, entre outros.

O Museu do Ceará é uma instituição memorialística necessária e eficiente a compreensão da história local. De acordo com as respostas, pode-se afirmar que o que o museu apresenta atualmente, em termos de documentos, materiais em geral que remetem a história da cidade de Fortaleza, bem como do Estado do Ceará, pode ser considerado suficiente para a compreensão da história do mesmo, mas isso não quer dizer que apenas uma visita, ainda que seja longa e sem a orientação de um guia, por exemplo, seja suficiente para que toda a história seja compreendida.

Tomando por base a questão anterior, é consensual à todos os pesquisados que o museu é suficiente para a compreensão da história de Fortaleza, entretanto, verifica-se que alguns deles se abstêm em responder apenas se o Museu é atrativo ou não a população fortalezense.

CONCLUSÕES

A ideia de pesquisar sobre o Museu do Ceará e sua função de comunicar o patrimônio cultural do estado do Ceará levou a constatar o seguinte quadro atual: o museu recebe bastantes visitantes, mas que estes mesmos visitantes, em sua maioria, são alunos de escolas e universidades, ou seja, oriundos de visitas feitas por estes mesmos alunos são feitas de forma obrigatória, através de atividades extracurriculares (aulas de campo) o que muitas vezes acabam sendo programadas e objetivas sem a oportunidade de discutir e refletir sobre a exposição.

Também concluímos que, além das exposições fixas do Museu do Ceará, não há uma periodicidade no que diz respeito à apresentação de novas exposições, este pode ser considerado um fator importante, pois pode atrair visitantes recorrentes, bem como visitantes que estão ali pela primeira vez. Isso nos faz menção à outra questão percebida nas pesquisas, que o Museu do Ceará não se preocupa em realizar ações de divulgação do Museu, não há um marketing direcionado especificamente ao Museu e quando há, esta se realiza apenas dentro do mesmo.

Ao mesmo tempo em que não se padronizam ações de autopromoção do Museu, ficou nítido que as poucas ações desenvolvidas têm seus objetivos alcançados, como a demanda de visitantes, mas ao mesmo tempo não se sabe ao certo quantas pessoas visitam o Museu diariamente, seja pela primeira vez ou se já são recorrentes.

O Museu do Ceará também disponibiliza guia de visita para os seus visitantes, mas o mesmo não fica disponível a qualquer momento, é preciso agendar uma visita orientada previamente, isto também não contribui para a valorização do museu, uma vez que, caso existam turistas, por exemplo, que não conhecem as normas do Museu, estes deverão se utilizar de seu conhecimento para compreender o que as exposições abordam sobre a história de Fortaleza e do Ceará.

Podemos afirmar que o Museu do Ceará cumpre o seu papel ao fornecer subsídios necessários para a compreensão da história do Estado do Ceará, mais especificamente, da cidade de Fortaleza. Ele exerce todas as exigências determinadas pela Legislação que rege os museus; apesar da falha da

disponibilização do profissional, guia de visitas. O museu é capaz de comunicar a história local de forma sintética e o seu próprio espaço físico é atrativo aos olhos curiosos, porém deixa a desejar no quesito autopromoção/divulgação.

Recomendamos mais pesquisas sobre os Museus com um espaço incitador e um estímulo para a construção e o registro da memória coletiva. Onde se manifestam os testemunhos das lembranças, visto que o Museu, como território para criar e produzir sentidos, configura-se em local no qual os testemunhos, a voz dos fatos passados ou recriados se encontram em uma troca de saberes individuais e coletivos.

REFERÊNCIAS BIBLIOGRÁFICAS

Brito, Carla Façanha de (2017). *Ex-Votos do Museu Vivo do Padre Cícero e Musealização: modelando a Informação Museológica do bem material integrando a presença intangível, simbólica, da memória coletiva*. [tese]. Rio de Janeiro: UNIRIO. No prelo.

Gil, Antonio Carlos. (1991). *Métodos e técnicas de pesquisa social*. 5. ed. São Paulo: Atlas.

Godoy, A. S. (1995). Introdução à pesquisa qualitativa e suas possibilidades. *Revista de Administração de Empresas*. São Paulo, v. 35, n. 2, mar./abr., p. 57-63.

Halbwachs, Maurice (2006). *A memória coletiva*. 2. ed. São Paulo: Centauro.

Horta, Maria de Lourdes Parreiras (1994). Semiótica e museu. *Cadernos de Ensaios: estudos de Museologia*, n. 2. Rio de Janeiro, IPHAN.

Legislação sobre museus. (2013). Brasília: Câmara dos Deputados. Recuperado em 18 março, 2017, de <http://bd.camara.gov.br/bd/handle/bdcamara/14599>.

Le Goof, J. (2003). Memória. In: Le Goof, J. (2003). *História e Memória*. (5th ed.). Campinas: Unicamp.

Minayo, M. C. S. *Pesquisa social: teoria, método e criatividade*. 24. ed. Petrópolis: Vozes, 2010.

Nora, P. (1993). Entre memória e história: a problemática dos lugares. *Projeto História*. 10, pp. 07-28. Recuperado em 18 março, 2017, de <http://revistas.pucsp.br/index.php/revph/article/viewFile/12101/8763>

Pollak, M. (1992). Memória e identidade social. *Estudos históricos*, 5 (10), pp. 200-212. Recuperado em 18 março, 2017, de <<http://bibliotecadigital.fgv.br/ojs/index.php/reh/article/view/1941/1080>>.

Ruoso, C. (2009). *O museu do ceará e a linguagem poética das coisas: (1971 – 1990)*. Fortaleza: Secult.

Santos, M. S. (2006). *A escrita em museus históricos*. Rio de Janeiro: Garamond.

TRIPODI, T.; FILLIN, P.; MEYER, H. (1981). *Análise da pesquisa social*. Rio de Janeiro: Francisco Alves.

TRIVIÑOS, A. N. S. (1987). *Introdução à pesquisa em ciências sociais: a pesquisa qualitativa em educação*. São Paulo: Atlas.

Yin, R. K. (2010). *Estudo de caso: planejamento e métodos*. 4. ed. Porto Alegre: Bookman.



COLECCIONES DE DATOS ABIERTOS ENLAZADOS: DE LA BÚSQUEDA AL DESCUBRIMIENTO DE INFORMACIÓN

María Luisa Alvite Díez

*Área de Biblioteconomía y Documentación. Universidad de León, 0000-0003-1490-8936,
luisa.alvite@unileon.es*

RESUMEN En el ámbito bibliotecario se abordan desde hace años proyectos dirigidos a la publicación de datos, metadatos y vocabularios aplicando el modelo de “Datos abiertos enlazados”. El acceso a estos datos se ha orientado, mayoritariamente, a usuarios expertos -desarrolladores- que utilizan SPARQL para interrogar estos silos de información. En este contexto, el estudio dirige su mirada a las soluciones concretas enfocadas a usuarios finales que han sido implementadas por instituciones inmersas en proyectos de datos abiertos enlazados. Se pretende realizar una aproximación sobre las características básicas de las interfaces de usuario en una muestra de portales de datos abiertos y extraer conclusiones sobre la potencia de búsqueda, la usabilidad y las tendencias observables en la visualización de la información en este tipo de proyectos. Para este fin, se han formalizado criterios preliminares para el análisis de las iniciativas seleccionadas. La metodología propuesta ha permitido realizar una primera aproximación sobre las funcionalidades básicas de las interfaces de búsqueda y visualización en estos sistemas. Los proyectos estudiados muestran, en conjunto, la potencialidad de las tecnologías semánticas frente a las colecciones que emplean un modelo de descripción plano. La publicación de datos abiertos enlazados mediante interfaces visuales intensifica el descubrimiento de información, potencia la contextualización de los resultados y mejora la experiencia de usuario, tanto por el incremento de la capacidad de navegación como por las interrelaciones entre los datos. No obstante, a pesar de las innovaciones, se concluye que el estadio de desarrollo actual de las plataformas observadas presenta debilidades, entre otras, destacamos las relacionadas con la visualización de la información -excesivo peso del texto y escasas innovaciones gráficas-, entornos poco intuitivos, o limitadas posibilidades de personalización y retroalimentación.

PALABRAS CLAVE *Datos abiertos enlazados, Experiencia de usuario, Usabilidad, Web semántica.*

ABSTRACT Projects for the publication of data, metadata and vocabularies have been carried out in the librarian field applying Linked Open Data model for some years. Mainly access to this data has been oriented to expert users - developers - who use SPARQL to query these repertories. In this context, the study focuses on specific solutions for end users that have been implemented by institutions immersed in linked open data projects. The aim is to make an approximation about the basic characteristics of the user interfaces in a sample of open data portals and draw conclusions about the search capability, usability and trends in the visualization of the information. For this purpose, preliminary criteria for the analysis of the selected initiatives have been formalized. The proposed methodology allowed us to make a first approximation about the basic functionalities of the search and visualization interfaces in these systems. The projects studied show the potential of semantic technologies versus collections that use a flat description model. The publication of linked open data by visual interfaces intensifies the discovery of information, enhances the contextualization of results and improves the user experience, both by increasing navigation capacity and by the interrelationships between data. However, in spite of the innovations, the present stage of development of the analysed platforms presents weaknesses, we highlight, the ones related to the information visualization -massive weight of text and few graphic innovations-, unintuitive environments, or customization and feedback limited.

KEYWORDS *Linked Open Data, Semantic Web, Usability, User Experience.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUCCIÓN

Bibliotecas, archivos y museos abordan desde hace años proyectos dirigidos a la publicación de datos, metadatos y vocabularios aplicando el modelo de “Datos abiertos enlazados”, una concreción específica del concepto de web semántica en la que se apuesta por la representación de información en RDF (Resource Description Framework), el empleo de URIs (Identificadores Uniformes de Recursos) y el uso del lenguaje de búsqueda en datos RDF, SPARQL.

Entre las ventajas de la aplicación de este modelo para las instituciones de la memoria destacan, como señala el informe del Library Linked Data Incubator Group (2011), el avance en la visibilidad de las colecciones, la reutilización de los datos, la mejora en los procesos de catalogación, la colaboración con otras comunidades generadoras de datos o la reducción de costes para las instituciones, entre otras.

Los datos publicados como datos enlazados están ideados para ser entendidos por máquinas y utilizados por aplicaciones que los necesiten para generar nuevos recursos y servicios, mediante descargas masivas o a través de consultas selectivas de datos empleando puntos de acceso SPARQL o utilizando otras tecnologías. Se trata de explotar los datos bibliotecarios mostrando su valor estratégico dentro de la web de datos interoperables. Igualmente relevante resulta la integración del paradigma de los datos enlazados con la iniciativa normalizadora Schema.org.

Paralelamente a los estándares procedentes de la web semántica han confluído alternativas procedentes del ámbito catalográfico, en concreto, RDA (Resource Description and Access) basada en el modelo conceptual FRBR-LRM (Functional Requirements for Bibliographic Records - Library Reference Model) acerca a las bibliotecas el escenario de datos enlazados. Un buen número de proyectos de datos enlazados ha utilizado FRBR como modelo conceptual de referencia y como modelo de datos. Especial mención merece BIBFRAME (Bibliographic Framework), concebido como modelo general para expresar y conectar datos bibliográficos en el entorno de los datos enlazados con el objetivo de que los datos procedentes de las bibliotecas pasen a formar parte de la web de datos.

El crecimiento de plataformas que implementan estas tecnologías semánticas en el ámbito bibliotecario es innegable, sin ser exhaustivo, el repertorio *The data Hub*¹ recopila 74 colecciones de datos y vocabularios. Con todo, el tránsito desde las iniciativas y prototipos experimentales a la implementación ordinaria habrá de venir de la mano de la generación y presentación de datos enlazados integrada en los programas de gestión de bibliotecas y en las herramientas de descubrimiento.

Atendiendo a este contexto, nuestro trabajo pretende indagar en las soluciones dirigidas a los usuarios finales que han sido adoptadas por instituciones inmersas en proyectos de datos abiertos enlazados. Entre las oportunidades y ventajas que la adopción de este modelo implica para el usuario final, la literatura (Hallo *et al.*, 2015; Konstantinou, Houssos, y Manta, 2014; Library Linked Data Incubator

¹ <https://datahub.io/es/dataset?q=lld>

Group, 2011. Mitchell, 2016; Simon *et al.*, 2013) detalla específicamente beneficios como los siguientes: mejor experiencia, contextualización de la información, dimensión multilingüe o avances en el rigor y la interdisciplinariedad en la actividad investigadora gracias al enriquecimiento de la información.

En concreto, este estudio persigue los objetivos siguientes:

- Efectuar una aproximación sobre las características básicas de las interfaces de usuario en una muestra de portales de datos abiertos.
- Formalizar criterios preliminares para el análisis de las iniciativas seleccionadas y observar la implementación de los mismos.
- Extraer conclusiones sobre la potencia de búsqueda, la usabilidad y las tendencias observables en la visualización de la información en este tipo de proyectos bibliográficos.

METODOLOGÍA

La metodología aplicada tiene en cuenta trabajos sobre evaluación de catálogos y bibliotecas digitales que presentan modelos de análisis que utilizaremos como referencia para sistematizar los indicadores aplicados (Alvite Díez, 2009; Sulé Duesa, Estivill Rius y Gascón García, 2011; Rodríguez Yunta, 2014). Se han establecido cuatro bloques de análisis:

- A) Colección
 - Descripción del proyecto
 - Volumen de datos
 - Licencia
 - Condiciones de uso
 - Publicaciones
 - Desarrollos
 - Descripción en otros idiomas
 - Contacto
- B) Tecnologías y estándares
 - Plataforma
 - Modelo conceptual
 - Modelos de datos
 - Vocabularios
 - Ontologías propias
 - Datos embebidos
 - Recuperación de datos
- C) Interfaz de usuario
 - Características generales
 - Navegación
 - Cajas de búsqueda
 - Sugerencia de términos
 - Presentación y visualización de los resultados
 - Refinamiento tras una búsqueda dada

- Sistemas de ayuda
- D) Funcionalidades de valor añadido
 - Vinculación semántica
 - Recuperación multilingüe
 - Mecanismos de retroalimentación e integración con herramientas sociales
 - Opciones de personalización
 - Espacios experimentales asociados

Como objeto de estudio se han seleccionado cuatro proyectos de entidad, volumen y tecnologías diversas que permiten observar tendencias en las soluciones previstas para el usuario final en el acceso a colecciones de datos enlazados:

- La Biblioteca Virtual de Polígrafos (BVP)² comenzó su andadura en 1999 con la intención de recoger la obra de pensadores españoles, hispanoamericanos, portugueses y brasileños. Agenjo Bullón y Hernández Carrascal (2010) señalan como clave del proyecto la constante adaptación del mismo a las transformaciones tecnológicas y de la web. La biblioteca se asienta en registros MARC 21/RDA, un repositorio OAI-PMH y la publicación de datos desde 2010 en formatos de la web semántica ajustados a la ontología de Europeana Data Model (EDM).
- El Catálogo de datos abiertos de la Biblioteca Virtual Miguel de Cervantes (DBVMC)³ se ha constituido tras un complejo proceso de transformación de los registros MARC 21 a una base de datos relacional basada en el modelo conceptual FRBR-FRAD. El contenido de la base de datos posteriormente fue migrado automáticamente a tripletas RDF utilizando el vocabulario RDA (Candela *et al.*, 2017).
- DataBnF⁴, la colección de datos abiertos enlazados de la Biblioteca Nacional de Francia (BnF) está disponible en la web desde julio de 2011, proporcionando diferentes visualizaciones de la misma información (Simon *et al.*, 2013). El proyecto supuso la transformación de los datos desde bases de datos no interoperables a datos estructurados e intercambiables compatibles con la web semántica.
- El Portal de datos bibliográficos de la Biblioteca Nacional de España (DBNE)⁵ da sus primeros pasos en 2010 y en diciembre de 2011 presentó la primera versión de datos.bne.es. La Biblioteca asume el reto de publicar los catálogos bibliográficos y de autoridades en formato RDF, conforme a los principios de los datos enlazados, partiendo de los modelos de la IFLA (Vila-Suero, Villazón-Terrazas y Gómez-Pérez, 2013)

Por lo que se refiere a la exploración de los portales seleccionados, esta se ha llevado a cabo a lo largo de los meses de febrero y marzo de 2017.

² <http://www.larramendi.es/fundacion/biblioteca-virtual-de-poligrafos/>

³ <http://data.cervantesvirtual.com/>

⁴ <http://data.bnf.fr/>

⁵ <http://datos.bne.es/>

RESULTADOS

Se presentan en este apartado los resultados del estudio atendiendo a los parámetros establecidos en la metodología.

1. COLECCIÓN

En la primera categoría de análisis se indaga en aspectos esenciales relacionados con la información proporcionada sobre el proyecto, volumen de la colección, objetivos, evolución, condiciones de uso o publicaciones, entre otros.

Los cuatro proyectos observados en este estudio suponen apuestas decididas por la publicación de catálogos como datos enlazados y cuentan con información bibliográfica sobre los mismos. Además, todos ellos se presentan como propuestas sujetas a mejoras continuas dentro de las cuales se integra el desarrollo de interfaces visuales innovadoras para la presentación de datos bibliográficos.

Por la propia naturaleza y entidad de las colecciones de origen, los proyectos de las dos bibliotecas nacionales son los más cercanos, sobresale, en nuestra opinión, la exhaustividad de la exposición del proyecto DataBnF recogida en la web. Con la excepción de la BVP, las licencias aplicadas a los datos, son licencias abiertas, libres de restricciones, que buscan favorecer la reutilización de los datos. No se han localizado estadísticas de uso de los portales.

2. TECNOLOGÍAS Y ESTÁNDARES

En este segundo parámetro se observan aspectos relacionados con las tecnologías que sustentan estas colecciones, modelos conceptuales, modelos de datos y vocabularios empleados, estándares embebidos o recuperación de datos.

La BVP está implementada en el programa de gestión digital DIGIBIB⁶ de la empresa DIGIBÍS. Paralelamente a la evolución de EDM se han incorporado sucesivas versiones a la colección; igualmente se han ido adaptando los campos del formato MARC 21 a las reglas de catalogación RDA. En el análisis funcional de esta colección se ha tenido en cuenta FRBR y también la iniciativa de la Library of Congress, BIBFRAME. Es posible la exportación de registros como Linked Open Data/EDM 5.2.7. La BVP dispone de un servidor SRU⁷ (Search/Retrieve via URL).

El Catálogo de DBVMC aplica el modelo FRBR a los registros existentes en MARC, utiliza la librería Apache Jena⁸ (Apache License, Version 2.0) para la transformación de los registros del catálogo a RDF, un marco de código abierto para desarrollar aplicaciones de web semántica. Sesame⁹ (en la actualidad Eclipse RDF4J) se usa como repositorio RDF para almacenar las tripletas resultado de la transformación. Para mejorar la interoperabilidad, se han utilizado vocabularios como Dublin Core, RDA, FOAF (Friend Of A Friend) o la ontología OWL-Time. En cuanto a datos embebidos, el proyecto utiliza

⁶ <http://www.digibis.com/software/digibib>

⁷ <http://www.larramendi.es/i18n/sru/sru.cmd>

⁸ <https://jena.apache.org/>

⁹ <http://rdf4j.org/>

Schema.org, Opendgraph Protocol (OG) ha sido utilizado para que las páginas puedan ser enlazadas en las principales redes sociales y Glyphicon Halflings para mostrar iconos. El punto de consulta SPARQL¹⁰ se ha creado empleando Sesame como motor de la base de datos y como servidor HTTP/SPARQL.

Por su parte, DataBnF emplea la plataforma de código abierto CubicWeb¹¹, un marco de aplicaciones semánticas disponible bajo licencia LGPL. El modelo de datos empleado es FRBR, además, la BnF hace uso de schema.org y Open Graph Protocol. El proyecto integra registros producidos fundamentalmente en formato InterMarc, XML-EAD y Dublin Core. Finalmente, los datos están estructurados en RDF atendiendo a diversas sintaxis. Para favorecer la interoperabilidad además emplea SKOS (Simple Knowledge Organization System), Dublin Core, FOAF, RDA y la ontología propia bnf-onto. En lo que se refiere a la recuperación de datos, prevé varias posibilidades, entre las que destacamos el empleo de SPARQL¹² y la descarga masiva mediante FTP.

En cuanto a las tecnologías empleadas en DBNE, la herramienta Marimba se ha usado para la transformación de los datos, se trata de una aplicación desarrollada por el Ontology Engineering Group que soporta el proceso de generación de RDF a partir de registros MARC 21 y su posterior enlazado. La datos se encuentran alojados en un servidor Virtuoso¹³, al que se accede mediante un terminal de SPARQL con una interfaz Linked Data construida con la herramienta gratuita Pubby¹⁴. FRBR ha sido utilizado como modelo conceptual de referencia y como modelo de datos. Además de propiedades de FRBR e ISBD, se han utilizado como referencia otras de otros vocabularios, como RDA, Dublin Core, SKOS, o BIBO (Bibliographic ontology). Para un mayor control del vocabulario se ha implementado una ontología propia BNE. Por lo que respecta a la recuperación de datos en DBNE, el portal cuenta con un punto de acceso SPARQL¹⁵, es posible la descarga de ficheros masivos en formato RDF: registros de autoridad, registros bibliográficos y encabezamientos de materias de la Biblioteca Nacional en SKOS.

En este parámetro de análisis se aprecia un predominio de tecnologías abiertas, basadas en estándares web, seguimiento del marco de referencia FRBR y estándares bibliográficos como RDA. Podemos hablar de la excepcionalidad de la BVP, construida con un programa propietario de la empresa DIGIBÍS y que no dispone de punto de consulta SPARQL. En el caso de DBNE el equipo de la Universidad Politécnica de Madrid ha desarrollado una herramienta *ad hoc* para el proceso de generación de RDF a partir de registros MARC 21 y su posterior enlazado. Destacamos que las dos bibliotecas nacionales aquí tratadas han elaborado ontologías propias. Por último, la apuesta por mejorar la indexación y búsqueda en los motores web se concreta particularmente en el uso de Schema.org por parte de DBVMC y DataBnF.

¹⁰ <http://data.cervantesvirtual.com/sparql>

¹¹ <https://www.cubicweb.org/>

¹² <http://data.bnf.fr/sparql>

¹³ <https://virtuoso.openlinksw.com/>

¹⁴ <http://wifo5-03.informatik.uni-mannheim.de/pubby/>

¹⁵ <http://datos.bne.es/sparql>

3. INTERFAZ DE USUARIO

Este bloque se dirige a observar las soluciones previstas en estos portales en cuanto a sus funcionalidades para la recuperación de información, nos hemos fijado en las opciones previstas para la navegación y la búsqueda explícita; se indaga igualmente en las capacidades de interacción con el usuario final, deteniéndonos en las soluciones para la visualización de resultados, las posibilidades de agrupación automática de los mismos, la exportación, etc., todo ello con la finalidad de indagar en aspectos comunes y en aportaciones singulares.

La mayor parte de los portales llaman la atención sobre el carácter experimental y en continuo desarrollo de la interfaz. Conviene subrayar que las acciones de búsqueda y navegación se entremezclan, es precisamente este punto el que marca una gran distancia con respecto a los catálogos tradicionales, esto es, los datos están federados alrededor de conceptos (Autores, Obras y Temas) y, además, los datos de autoridad juegan aquí un papel fundamental.

Salvo en el caso de la BVP, articulada en torno a una interfaz bibliográfica clásica, el resto se aproxima a las herramientas de descubrimiento focalizadas en proponer una caja de búsqueda única disponible para el usuario en cualquier momento del proceso de interacción con el sistema. Con la excepción de la BVP que desarrolla portales temáticos, el acceso directo por navegación prioritario se fija en presentar contenidos destacados en la página de inicio, explotando en algunos casos un mapamundi para navegar por localizaciones.

En lo que respecta a la presentación de resultados, la BVP despliega los mismos en formato abreviado acompañados de un marco para navegar por facetas; en el resto de los sistemas la solución se aleja de los catálogos clásicos y busca que el usuario seleccione los ítems de interés, identificando en el listado de resultados la categoría a la que pertenece, fundamentalmente: Autor, Obra y Tema. Entre los sistemas observados, se advierte en DBVMC el empleo de categorías para navegar como “Expresión” y “Manifestación” susceptibles de generar confusión en el usuario final. Cabe destacar en la presentación de los registros finales, las páginas de obras y de autores, estas últimas suponen, sin duda, una clara metamorfosis en la visualización de la información bibliográfica. DataBnF resulta admirable en el detalle prolijo empleado para mostrar los niveles de participación de los autores en las obras. Con todo, creemos necesarias mejoras en la visualización de los datos, entornos más atractivos e intuitivos, menos sobrecargados de texto y que requieran un empleo más moderado de la barra de desplazamiento lateral.

Por otro lado, no se han observado herramientas de corrección automática y solo los sistemas de las dos bibliotecas nacionales implementan sugerencias automáticas que valoramos muy positivamente, tal vez mejor resueltas en la solución de la BNE. La explotación de las autoridades de materias, en ambos portales merece igualmente ser destacada.

En general, salvo en la BVP, resulta muy limitada la interacción con el usuario, con escasas posibilidades de ordenación de resultados, limitaciones para guardar o reutilizar búsquedas, exportar o guardar varios ítems conjuntamente, etc. En ninguno de los sistemas está prevista la posible exportación a gestores bibliográficos.

Sería deseable contar con algún material de apoyo sobre el uso de los sistemas, con ejemplos prácticos y aclaraciones sobre los criterios de ordenación automática de los resultados, filtros propuestos, etc. En el momento de redacción es este trabajo, solamente DBNE dispone de una guía de ayuda.

4. FUNCIONALIDADES DE VALOR AÑADIDO

Este último parámetro se articula para tratar de valorar aspectos como la profundidad en la explotación de los enlaces, la capacidad de comunicación con los usuarios, la previsión de personalización o los posibles desarrollos de herramientas semánticas generadas a partir de la publicación de los datos de las colecciones consideradas.

La BVP emplea de forma sistemática VIAF, GeoLinkedData (Instituto Geográfico Nacional), GeoNames y Wikidata y apuesta por técnicas de reconciliación semántica para extender la selección de vocabularios con los que establecer vínculos: las bibliotecas nacionales de Alemania, España, Francia, la Library of Congress, el Catálogo del Système Universitaire de Documentation (SUDOC) o WorldCat. Se ha extendido la vinculación a las DBpedias en castellano, catalán, gallego, vasco y portugués, además de otras fuentes como Social Network and Archival Context (SNAC). Todos los encabezamientos de materia permiten una navegación entre conceptos gracias a la integración de la Lista de Encabezamientos de Materia publicada en SKOS por la Subdirección General de Coordinación Bibliotecaria. Estos registros SKOS están vinculados con otras listas de materias, como las de la Library of Congress, RAMEAU, FAST (Faceted Application of Subject Terminology) o las del Gemeinsame Normdatei.

DBVMC incorpora enlaces a VIAF, ISNI, Wikidata, Wikipedia para autores, idiomas y fechas, a la Biblioteca Nacional de Francia o a la Library of Congress.

Por su parte, DataBnF presenta vínculos con repositorios externos equivalentes: Library of Congress, Biblioteca Nacional de Alemania, VIAF, Geonames, Agrovoc y Thesaurus W (Tesoro para el control de autoridades de archivos franceses); enlaces al catálogo general de la BnF, Catalogue collectif de France, Archivos y manuscritos BnF, Centre national de la littérature pour la jeunesse - La Joie par les livres, Europeana, SUDOC, WorldCat; por último, vincula sus datos a Wikipedia vía DBpedia y Wikidata.

En DBNE los datos correspondientes a ciertos elementos, fundamentalmente autores, materias y obras, se enriquecen con conexiones a sus equivalentes en otras fuentes de datos. En el caso de autores, se proveen enlaces, si están disponibles, con Library of Congress, Bibliotecas nacionales de Alemania, Francia y Suecia, SUDOC, VIAF e ISNI. Los registros de materias, geográficos y género/forma tienen enlaces a Library of Congress.

Este cuarto parámetro dedicado a otras funcionalidades de los sistemas pone en valor la fortaleza de la publicación de datos abiertos como recurso extraordinario para el enriquecimiento de la información y la contextualización de la misma. Destaca la amplitud de conexiones a fuentes de datos de DataBnF y la integración de lenguajes documentales en la BVP. Con todo, las posibilidades de interconexión son susceptibles de ampliarse de modo formidable, datos de archivos, museos, datos estadísticos, pueden servir de ejemplo.

En este bloque se observan aspectos claramente mejorables en lo que se refiere a la recuperación multilingüe, a los mecanismos de comunicación y retroalimentación, como posibilidades de alertas según perfil del usuario, posibles cuestionarios sobre la colección y/o el servicio o formularios de sugerencias. En los casos de la BVP y DBNE no está previsto compartir páginas en redes sociales y solamente en la BVP y en la BVMC los usuarios pueden introducir anotaciones, valoraciones o

comentarios. Tampoco se han observado posibilidades de personalización por parte del usuario o preferencias de sesión.

Finalmente, consideramos muy alentadores los espacios dedicados a desarrollos asociados a los proyectos de datos abiertos enlazados, de modo especial en la BVMC y DataBnF. La BVP dispone de una herramienta denominada “Línea del tiempo¹⁶”, destinada a ubicar a los polígrafos temporal y espacialmente.

En DBVMC ocupa un lugar destacado el acceso a “Lab¹⁷”, donde se presentan un conjunto de herramientas mayoritariamente lingüísticas: Buscador geográfico, Analizador sintáctico automático, Resúmenes automáticos generados a partir de los textos de obras de la BVMC, Buscador diacrónico dentro de un corpus de obras de la BVMC, y Corpus de Sonetos del Siglo de Oro.

En el caso de la colección de DataBnF, Atelier¹⁸, es un espacio experimental para descubrir los datos de la biblioteca: frisos cronológicos, galerías de imágenes, mapas, etc., los datos podrían visualizarse de diferentes maneras como cronogramas, geográficamente o como un mapa de relaciones científicas entre obras, personas, instituciones o movimientos.

CONCLUSIONES

Si bien uno de los objetivos fundamentales de la publicación de las colecciones bibliográficas como datos abiertos enlazados se fija en que esta información esté disponible más allá del limitado entorno del catálogo, este reto puede conciliarse con portales que trascienden al catálogo para convertirse en nuevos ecosistemas de conocimiento que impulsen aproximaciones más transversales al saber.

La metodología de análisis propuesta nos ha permitido realizar una primera aproximación sobre las características básicas de las interfaces de búsqueda y visualización de colecciones de datos orientadas al usuario final. Los proyectos estudiados muestran, en conjunto, la potencialidad de las tecnologías semánticas frente a las colecciones que emplean un modelo de descripción plano. La publicación de datos abiertos enlazados mediante interfaces visuales intensifica el descubrimiento de información, potencia la contextualización de los resultados y mejora la experiencia de usuario, tanto por el incremento de la capacidad de navegación como por las interrelaciones entre los datos.

Con todo, a pesar de las innovaciones, el estadio de desarrollo actual de las plataformas observadas presenta carencias, entre otras, destacamos las relacionadas con la visualización de la información - excesivo peso del texto y escasas innovaciones gráficas-, uso de terminología bibliotecaria innecesaria, o limitadas posibilidades de personalización y retroalimentación.

Abogamos por sistemas de búsqueda para el usuario final más sofisticados y potentes, estudio de algoritmos de ordenación de resultados más eficaces y explotación de nuevas facetas de navegación que incrementen las posibilidades de descubrimiento y eviten el desbordamiento cognitivo del usuario. Resulta crucial potenciar sistemas más amigables, favorecer la recuperación multilingüe o impulsar el

¹⁶ <http://www.larramendi.es/es/timeline/autores.do>

¹⁷ <http://data.cervantesvirtual.com/blog/labs/>

¹⁸ <http://data.bnf.fr/atelier/>

desarrollo de herramientas experimentales para la explotación de subconjuntos de datos, con enormes posibilidades para la docencia e investigación, entre otras.

Por último, consideramos esencial estudiar los desarrollos de estos portales, valorar su adaptación al comportamiento real de los usuarios y su influencia en el uso de las colecciones de las bibliotecas. Creemos que para garantizar su aceptación por parte de la comunidad a la que se dirigen deviene imprescindible someter estos sistemas a procesos evaluativos rigurosos, adaptados a la complejidad de las colecciones y a la diversidad de usos y de usuarios.

REFERENCIAS BIBLIOGRÁFICAS

Agenjo Bullón, X., y Hernández Carrascal, F. (2010). Las Bibliotecas Virtuales FHL: fuente de información bibliográfica para el pensamiento iberoamericano en la Web 3.0. En *V Congreso Nacional de Bibliotecas Públicas, Gijón (España), 3, 4 y 5 de noviembre de 2010*. Recuperado de <http://hdl.handle.net/10760/14719>

Agenjo Bullón, X., y Hernández Carrascal, F. (2015). Cómo y qué consumir en el mundo Linked Open Data; cómo y qué producir en Linked Open Data”. En *XIV Jornadas Españolas de Documentación, Gijón 28, 29 y 30 de mayo de 2015* (pp. 169-195). Madrid: Fesabid. Recuperado de http://www.fesabid.org/sites/default/files/repositorio/actas_fesabid_2015_0.pdf

Alvite Díez, M. L. (2009). Interfaces y funcionalidades de Bibliotecas digitales. *Anales de Documentación*, 12, 7-23. Recuperado de <http://revistas.um.es/analesdoc/article/view/70221>

Berners-Lee, T. (2009). *Linked Data*. W3C. Recuperado de <http://w3.org/DesignIssues/LinkedData.html>

Candela G., Escobar, P., Carrasco, R.C., y Marco-Such, M. (2017). Migration of a Library Catalogue into RDA linked open data. *Semantic Web* (preprint), 1-11. Doi: 10.3233/SW-170274

Cole, T.W., Han, M-J. K., Sarol, M.J., Biel, M., y Maus, D. (2017). Using Linked Open Data to Enhance the Discoverability, Functionality & Impact of Emblematica Online. *Library Hi Tech*, 35(1), 159-178. doi:10.1108/LHT-11-2016-0126

Deliot, C. (2014). Publishing the British National Bibliography as Linked Open Data. *Catalogue & Index*, (174), 13-18. Recuperado de http://www.bl.uk/bibliographic/pdfs/publishing_bnb_as_lod.pdf

Hallo, M., Luján-Mora, S., Maté, A., y Trujillo, J. (2016). Current state of Linked Data in digital libraries. *Journal of Information Science*, 42(2), 117-127. Doi: 10.1177/0165551515594729

Hyland, B., Ateazing, G., y Villazón-Terrazas, B. (2014). Best Practices for Publishing Linked Data. *W3C Working Group Note*, January 9. Recuperado de <http://www.w3.org/TR/ld-bp/>

Konstantinou, N., Houssos, N., y Manta, A. (2014). Exposing Bibliographic Information as Linked Open Data using Standards-based Mappings: Methodology and Results. *Procedia: Social and Behavioral Sciences*, 147, 260-267. Doi: 10.1016/j.sbspro.2014.07.169

Libhub Initiative. <http://www.libhub.org/>

Library Linked Data Incubator Group (2011). *Library Linked Data Incubator Group Final Report*. W3C. <http://w3.org/2005/Incubator/llid/XGRllid-20111025>

Méndez, E., y Greenberg, J. (2012). Datos enlazados para vocabularios abiertos y marco general de HIVE. *El profesional de la información*, 21(3), 236-244. Recuperado de http://www.elprofesionaldelainformacion.com/contenidos/2012/mayo/03_esp.pdf

Mitchell, E. T. (2013). Building Blocks of Linked Open Data in Libraries. *Library Technology Reports*, 49 (5), 11–25. Recuperado de <https://journals.ala.org/index.php/ltr/article/view/4692/5584>

Mitchell, E. T. (2016). Library Linked Data: Early Activity and Development. *Library Technology Reports*, 52(1). Recuperado de <https://journals.ala.org/ltr/issue/download/534/290>

Rodríguez Yunta, L. (2014). Usabilidad y recuperación en Europeana: comparación con otras bibliotecas digitales y portales. En L. F. Ramos Simón, & R. Arquero Avilés (Coords.). *Europeana. La plataforma del patrimonio cultural europeo* (pp. 177-200). Gijón: Trea.

Simon, A., Wenz, R., Michel, V., y Di Mascio A. (2013) Publishing Bibliographic Records on the Web of Data: Opportunities for the BnF (French National Library). En Cimiano P., Corcho O., Presutti V., Hollink L., Rudolph S. (eds.) *The Semantic Web: Semantics and Big Data. ESWC 2013. Lecture Notes in Computer Science*, vol. 7882. Berlin: Springer. Doi: 10.1007/978-3-642-38288-8_38

Sulé, A., Centelles, M., Franganillo, J., y Gascón, J. (2016). Aplicación del modelo de datos RDF en las colecciones digitales de bibliotecas, archivos y museos de España. *Revista Española de Documentación Científica*, 39(1): e121. Doi: <http://dx.doi.org/10.3989/redc.2016.1.1268>

Sulé Duesa, A., Estivill Rius, A., y Gascón García, J. (2011). Interfaces de consulta en las colecciones digitales patrimoniales españolas. *Anales de Documentación*, 14(2). Recuperado de <http://revistas.um.es/analesdoc/article/view/113931>

Torre-Bastida, A. I., González-Rodríguez, M., y Villar-Rodríguez, E. (2015). Datos abiertos enlazados (LOD) y su implantación en bibliotecas: iniciativas y tecnologías. *El profesional de la información*, 24(2), 113-120. Doi: <http://dx.doi.org/10.3145/epi.2015.mar.04>

Vila-Suero, D., Villazón-Terrazas, B., y Gómez-Pérez, A. (2013). datos.bne.es: a Library Linked Data Dataset. *Semantic Web*, 4(3), 307-313. Doi: <http://dx.doi.org/10.3233/SW-120094>



LA VISIBILIDAD DEL TRADUCTOR EN CATÁLOGOS Y REPERTORIOS BIBLIOGRÁFICOS ESPAÑOLES (1970- 2010): POSIBILIDADES Y ESTRATEGIAS DE ANÁLISIS

Belén Santana López¹, Crispulo Travieso Rodríguez², Manuel de la Cruz Recio³

Universidad de Salamanca

¹0000-0003-0102-5309, bsantana@usal.es

²0000-0002-0774-0728, ctravieso@usal.es

³0000-0002-0686-6652, manueldelacruzrecio@usal.es

RESUMEN El objetivo de este estudio es el análisis de la representación de la figura del traductor en los catálogos bibliográficos, como ejemplos tradicionales de instrumentos de organización del conocimiento. Partiendo de la premisa de la estrecha relación entre ambas disciplinas (Traducción y Ciencias Documentales), se ha procedido a realizar una búsqueda exhaustiva de todas las obras traducidas del alemán al español en el período comprendido entre 1970 y 2010. Tras una exploración y evaluación de las fuentes de información posibles, se realizó una búsqueda documental intensiva en el catálogo de la Biblioteca Nacional de España. Fruto de ese proceso y de la consiguiente obtención de un corpus bibliográfico completo sobre estas traducciones, se han conseguido identificar determinadas claves intervinientes en la consideración de la labor del traductor y su visibilidad en uno de los recursos bibliográficos más importantes en el ámbito lingüístico del español.

PALABRAS CLAVE *Traducción, catálogos bibliográficos, visibilidad, alemán, español.*

ABSTRACT This contribution aims at analyzing the translator's visibility in bibliographic catalogues for being a traditional example of knowledge organisation resources. Starting from a close interaction between Translation Studies and Information Sciences, the corpus consists of all books translated from German into Spanish between 1970 and 2010. After a thorough evaluation of the sources of information available, an active search was conducted based on the catalogue of the Spanish National Library. This search resulted in a complete bibliographic corpus of translations which allows to identify specific practices as far as the consideration of the translator as a literary agent and his/her visibility in one of the most important bibliographic catalogues within the Spanish speaking world are concerned.

KEY WORDS *translation, bibliographic catalogues, visibility, German, Spanish*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUCCIÓN

La vinculación entre las disciplinas de Documentación y Traducción queda patente en el ejercicio profesional de los especialistas en cada materia de una forma constante, podríamos decir incluso que instintiva. Los traductores precisan el uso de fuentes documentales de todo tipo para su labor y los documentalistas necesitan de las traducciones y relaciones lingüísticas que los primeros realizan. Dicha

imbricación tiene también su reflejo en el ámbito académico, donde la inclusión en los planes de estudio de materias pertenecientes a la otra disciplina permite que surjan zonas de intersección que favorecen distintas sinergias, tanto a nivel investigador como docente (Santana y Travieso, 2013). Al margen de esta vertiente aplicada, pueden establecerse asimismo otros ámbitos de mutua dependencia: desde el punto de vista de la Organización del Conocimiento, los esquemas que las Ciencias de la Documentación establecen para facilitar la descripción y recuperación de la información también determinan el posicionamiento y la visibilidad de los propios agentes implicados en la creación, difusión e intercambio de dicha información. En definitiva, y como ocurre en cualquier intento por sistematizar la realidad a posteriori, el modo en que formulamos representaciones de los documentos con criterios biblioteconómicos refleja, intrínsecamente, una cierta visión del protagonismo o peso de la aportación de cada uno de los elementos intervinientes.

Precisamente el concepto de “(in)visibilidad” del traductor ha merecido especial atención en los últimos años, en particular dentro del ámbito anglosajón de los Estudios de Traducción (Venuti, 1995). Dicha (in)visibilidad muestra una triple vertiente, ya que puede aplicarse (1) a la persona del traductor y su condición de co-autor de un texto en tanto en cuanto esta es reconocida dentro del campo literario (editoriales, medios de comunicación, instituciones culturales, organismos encargados del fomento de la lectura, etc.); (2) a la plasmación concreta de la labor del traductor dentro de un texto o (3) a la actividad traductora como práctica y proceso cultural y sus correspondientes productos (Emmerich, 2013, p. 200). Es en la primera de las vertientes donde se sitúa este estudio prospectivo, el cual, partiendo de la estrecha relación entre las operaciones de gestión de colecciones bibliográficas y el grado de notoriedad que consecuentemente se concede a los profesionales implicados en las mismas, se propone analizar la visibilidad de la figura del traductor en los catálogos y repertorios bibliográficos españoles con el fin de extraer unas pautas iniciales que permitan abordar estudios experimentales de mayor calado. Así, inspirados por fuentes documentales específicas ya existentes (Calañas y Raposo, 2006) y las propias normas de descripción bibliográfica, se pretende ir un paso más allá y analizar de forma empírica en qué medida dichas fuentes son completas y hasta qué punto las reglas de catalogación se aplican de forma consistente. Dicho análisis permitirá extrapolar algunas conclusiones preliminares sobre la visibilidad del traductor como agente creador de conocimiento.

En cuanto a la conceptualización del objeto de estudio y sin perjuicio de realizar posteriores ajustes, se consideraron tres niveles iniciales de visibilidad, según los grados de menciones siguientes esperados a priori:

- 1) Visibilidad nula: en el registro no se menciona al traductor ni al hecho de que se trata de una obra traducida.
- 2) Visibilidad parcial: en el registro se menciona que se trata de una obra traducida del idioma X.
- 3) Visibilidad total: en el registro se menciona que se trata de una obra traducida del idioma X por el traductor Y.

METODOLOGÍA

La siguiente elección metodológica fue la selección –convertida, a su vez, en evaluación a la postre– de las fuentes de información bibliográfica que se utilizarían para el análisis. La dificultad para llevar a cabo esta labor venía dada porque la fuente debía cumplir varias condiciones:

- proporcionar una interfaz de búsqueda que permitiera consultas con un alto grado de especificidad;
- contener la mayor cantidad posible de registros potencialmente útiles, lo que incluía la extensión del período cronológico abarcado;
- disponer de una descripción catalográfica exhaustiva;
- facilitar la extracción de los resultados en formatos manejables;
- contar con una adecuada actualización.

Como se fue comprobando en este proceso inicial, no hubo una fuente ideal para realizar las búsquedas que pudiese tomarse como punto de partida infalible. Precisamente todo ese ejercicio de tanteo trajo consigo la identificación del primer bloque de conclusiones que pudieron extraerse de este trabajo. En todo caso, a continuación se enuncia la relación íntegra de fuentes que se barajaron para la investigación y se describen brevemente sus características

BIBLIOTECA NACIONAL DE ESPAÑA (BNE)

La Biblioteca Nacional de España, fundada en 1711, recibe y conserva ejemplares de todos los libros publicados en España. En 1986, mediante un real decreto se insiste en la obligatoriedad de entregar un ejemplar de toda obra impresa en España a la Biblioteca Nacional. En 1957 se aprueba el Reglamento de Depósito Legal, que se actualiza mediante una Ley del 30 de julio de 2011. Esta nueva ley, fruto de un gran acuerdo bibliotecario, responde a la necesidad de adaptar la recopilación del patrimonio bibliográfico –y así su conservación y difusión– a los cambios producidos en el mundo de la edición como consecuencia de las nuevas tecnologías y, especialmente, a las publicaciones en red. Son objeto de depósito legal, todas las obras bibliográficas, sonoras, visuales, audiovisuales y digitales, producidas o editadas en España, por cualquier procedimiento de producción, edición o difusión y distribuidas en cualquier soporte, tangible o intangible.

AGENCIA DEL ISBN Y DISTRIBUIDOR DE INFORMACIÓN DEL LIBRO ESPAÑOL EN VENTA (DILVE)

Para consultar los libros editados en España existe la plataforma *DILVE plus: Distribuidor de información del libro español en venta*. Esta plataforma es gestionada por la Federación de Gremios de Editores de España, una entidad de carácter privado que financian los editores y recibe subvenciones, entre otras del Centro Español de Derechos Reprográficos (CEDRO). La plataforma DILVE plus recoge la *Base de datos de libros editados en España* (MECD), gestionada por la Agencia del ISBN (ahora en manos de la Federación), y los datos que aportan las propias editoriales (DILVE). El acceso a DILVE plus y la extracción de registros tiene lugar en régimen de suscripción. Previa solicitud con fines investigadores, pudimos acceder a la plataforma y extraer los datos necesarios.¹ Según datos facilitados por el propio DILVE, la Agencia del ISBN incluye unas 17.000 editoriales, la mayoría de las cuales no están en activo. Otras editoriales no están suscritas a la plataforma DILVE, desde la que se han importado las antiguas bases del ISBN (originalmente en papel, con lo cual pueden faltar muchos datos

¹ En este punto los autores desean agradecer a los responsables de gestionar la plataforma DILVE las facilidades obtenidas y toda la información suministrada durante el proceso de elaboración de este artículo.

o no estar actualizados). La responsabilidad de cumplimentar y actualizar los datos en DILVE recae en cada editorial.

CATÁLOGO COLECTIVO REBIUN (RED DE BIBLIOTECAS UNIVERSITARIAS)

El catálogo colectivo REBIUN recoge todos los registros volcados por las distintas Bibliotecas Universitarias de España, públicas y privadas. Por su propia naturaleza, este catálogo da prioridad a obras de carácter académico.

WORLDCAT

WorldCat es la red más grande del mundo en contenido y servicios de bibliotecas. Dado que en esta primera fase nuestro estudio se ciñe al ámbito español y que son las propias bibliotecas nacionales las que vuelcan sus datos a este catálogo, consideramos que no es una fuente principal para esta investigación. Por otro lado, la documentación elaborada por el propio Worldcat detecta lagunas en la catalogación de los traductores (Smith-Yoshimura, 2014).

INDEX TRANSLATIONUM (IT)

El Index Translationum es un repertorio de obras traducidas en todo el mundo, es decir, una bibliografía internacional de traducciones, creada en 1932. La base de datos contiene una información bibliográfica acumulativa sobre las obras traducidas y publicadas en un centenar de Estados Miembros de la UNESCO entre 1979 y 2009. En el caso de España, el Index Translationum se gestionó desde la BNE, pero según información de la propia institución los datos no están actualizados. En la propia página del IT figura para España que el último año en el que se recibieron datos fue el 2012, mientras que el último año sobre el que se publicaron datos fue el 2008 ("2009, 2010, 2011 and 2012 currently being processed by the INDEX team"). Tras dirigir una consulta al respecto al portal oficial del IT, no obtuvimos respuesta.

GOETHE INSTITUT (GI)

El Instituto Goethe publica distintas bibliografías, siempre selectivas, sobre traducciones del alemán al español. Asimismo, pone a nuestra disposición varias fuentes electrónicas, también selectivas, bien referidas a un número determinado de autores (40), a un determinado género (literatura infantil y juvenil) o a las traducciones que han sido subvencionadas por la propia institución (cf. Bibliografía). Del mismo modo, nos ofrecen un extracto del catálogo de la biblioteca del GI de Madrid filtrado para obtener las obras traducidas al español y las obras en edición bilingüe.

PANORÁMICA DE LA EDICIÓN ESPAÑOLA DE LIBROS

La *Panorámica de la Edición Española de Libros* es una publicación del Observatorio de la Lectura y el Libro existente desde 1988, que recoge y analiza la estadística del sector del libro en España, elaborada por el Ministerio de Educación, Cultura y Deporte, a partir de los datos gestionados por la Agencia Española del ISBN. Los datos se pueden consultar en relación con dos periodos: Datos anuales

referidos al año anterior y Cuadros de evolución de los últimos cinco años. Los datos de Comercio Interior (del que forman parte los datos sobre Derechos de autor) y los de Comercio Exterior proceden de la Federación de Gremios de Editores de España y de la Federación de Cámaras del Libro, respectivamente. Los datos ofrecidos recogen los principales indicadores del sector, mientras que el análisis pormenorizado se encuentra en la publicación digital de la *Panorámica de la Edición Española de Libros*. La última panorámica completa corresponde al año 2015; asimismo se puede consultar un avance de los datos de 2016. En lo que respecta a la traducción, la panorámica recoge datos sobre: porcentaje de libros traducidos respecto a la producción total; porcentaje de lenguas traducidas sobre la producción total y porcentaje de lenguas sobre los libros traducidos. No se recogen datos específicos sobre la figura del traductor.

REPERTORIOS BIBLIOGRÁFICOS

Los repertorios bibliográficos y ensayos existentes aplicados a la combinación de lenguas alemán-español (Calañas y Raposo, 2006; Fortea, 2009; Maldonado, 2009; Möning, 1969; Stumpe, 1978) corresponden a un periodo que no se ajusta al objeto de estudio en esta primera fase. No obstante, creemos que pueden resultar de interés para posteriores investigaciones como fuente de consulta y herramienta de cotejo. En especial queremos destacar el estudio de Cuéllar Lázaro (2000) en lo que respecta a su rigor metodológico. En la primera parte de su tesis doctoral, la autora elabora un corpus de traducciones de literatura contemporánea del alemán al español realizadas entre 1945 y 1990. Aunque este estudio no se centra en la visibilidad del traductor, sí ofrece un análisis crítico de las fuentes consultadas y constata que ninguna de ellas es completa. Esta imprecisión convierte la elaboración del corpus en una tarea "paciente y ardua" (2000, pp. 41s.). Una vez obtenido el corpus, se apuntan varias perspectivas para un análisis pluridisciplinar que incluyen datos estadísticos y reseñas sobre la procedencia de los autores traducidos; los autores más traducidos por países; la práctica inexistencia de traducciones de la literatura de la RDA durante el franquismo; los géneros más traducidos y los traductores y las editoriales más prolíferos en España, si bien en este último punto, como la propia autora reconoce, la muestra no es lo suficientemente representativa.

Una vez analizadas todas estas fuentes y teniendo en cuenta los criterios de selección, se decidió centrar la búsqueda en el catálogo de la Biblioteca Nacional por ser la fuente que ofrece la mayor y más exhaustiva cobertura del periodo abarcado, en especial a partir de 1986, año en el que se insiste en la obligatoriedad del depósito legal de toda obra impresa. Asimismo, en 1987 entra en vigor la *Ley de propiedad intelectual*, que reconoce la condición de autor del traductor (artículo 46) y establece, en su artículo 64, que el editor está obligado a "Reproducir la obra en la forma convenida, sin introducir ninguna modificación que el autor no haya consentido y *haciendo constar en los ejemplares el nombre, firma o signo que lo identifique*" (cursiva propia). Por último, atendiendo al estudio de la visibilidad del traductor como transmisor del conocimiento editorial en la sociedad, se ha dado prioridad a una fuente de libre acceso.

En cuanto a la estrategia de búsqueda, se pretendía localizar todas las obras traducidas del alemán al español entre 1970 y 2010 recogidas en el catálogo de la BNE. El modo en que se trasladó la estrategia de búsqueda ya constituyó en sí mismo uno de los resultados más elocuentes de esta investigación. El número de documentos recuperados bajo los distintos campos del formato MARC accesibles en el modo de búsqueda avanzada (Figura 1) ya era un indicio sobre una mayor o menor visibilidad de la figura del traductor. Concretamente, se lanzaron consultas sobre los siguientes campos del formulario:

- Título. Equivalente al campo 245, de título y mención de responsabilidad.
- Todos los campos. Se trataba de una búsqueda inclusiva, pero previsiblemente inexacta para nuestro propósito, que abarcaba el campo 700 (Punto de acceso adicional – Nombre de persona) y los campos 5xx (relativos al área de notas).

Figura 1: Interfaz de búsqueda avanzada del catálogo de la BNE

Los resultados de esas búsquedas fueron bajos, especialmente en comparación con lo esperable según el porcentaje de obras traducidas del alemán que se declaraban en la *Panorámica de la edición española del libros* que edita el Ministerio de Cultura basándose en el ISBN (Gráfico 1). Según esa fuente, la media de traducciones del alemán en los últimos 20 años ha sido del 6,3% del total de libros traducidos, oscilando en un rango nunca menor al 5,1%. Sin embargo, las búsquedas ejecutadas mediante el OPAC web de la Biblioteca Nacional arrojaban unas cifras sensiblemente inferiores (2,66%).²

² Secuencia de búsqueda en el campo 245 (Título): trad\$ = 400.511 registros / trad\$ + alem\$ = 10.668 registros. Cuando la búsqueda se ejecutaba en todos los campos bibliográficos el porcentaje era incluso menor debido al ruido documental de la recuperación.

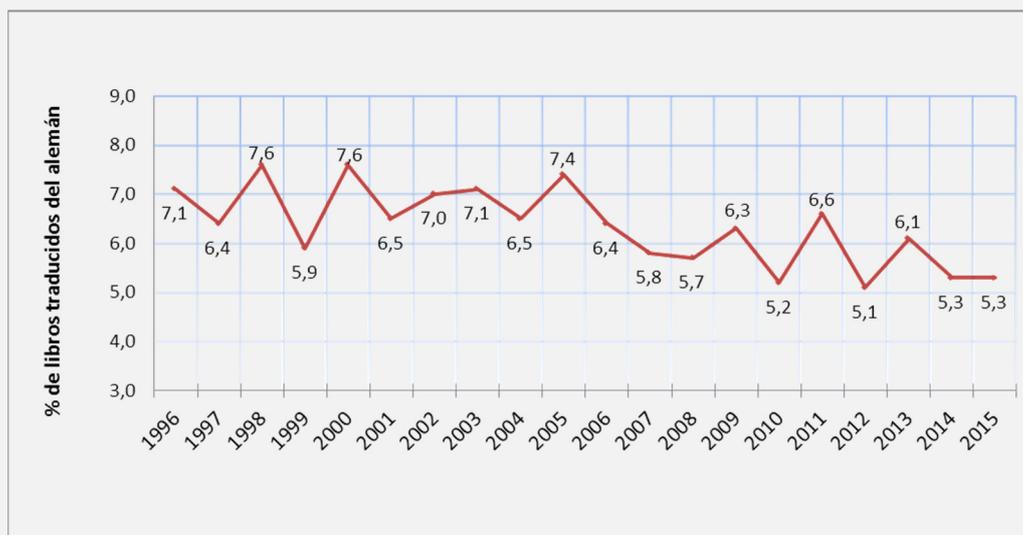


Gráfico 1. Porcentaje de traducciones del alemán sobre el total de libros traducidos (1996-2015)

A la vista de ese resultado, se contactó con el personal técnico de la BNE para que elaborara una estrategia de búsqueda más afinada, que no dependiera de las variables terminológicas dentro de las opciones que posibilitaba la búsqueda avanzada del OPAC web. En concreto se propuso la búsqueda desde un campo del formato MARC que permite señalar aspectos sobre el idioma de la obra, esto es, el campo 041 – Código de lengua, cuyo primer indicador, si es 1, significa que la obra “es o incluye una traducción”. En el subcampo \$h se consigna el idioma original de la obra. Así, se pidió a la BNE que lanzaran dicha búsqueda con el término “ger”, a partir de la cual obtuvimos un documento Excel con los registros solicitados.

RESULTADOS

En total se obtuvieron 22.071 documentos, lo que suponía más del doble de la compilación anterior; de hecho, en porcentaje sobre obras traducidas se ajustaba directamente a lo previsible (5,5%), cumpliendo el primer requisito planteado sobre exhaustividad de la recuperación. En este sentido, cabe señalar que, por la propia estructura del registro, la búsqueda realizada por un campo referido a la obra (041), y por tanto a la traducción como producto, era más certera que la búsqueda por el traductor como persona o agente. En cualquier caso, en este momento ya se contaba con una lista de documentos de la que habíamos eliminado el sesgo de búsquedas inexactas, garantizando que todos los registros eran obras traducidas del alemán.

Por tanto, para el objetivo principal de esta investigación –analizar la visibilidad del traductor según los registros bibliográficos–, este documento constituía finalmente un punto de partida fiable. Desde el punto de vista cualitativo, la observación de esta colección demostró que la información sobre el traductor no era consistente en todos los casos; al contrario, existía una notable disparidad en la consideración del mismo y en la terminología empleada para designar su labor (la traducción figuraba en ocasiones como *versión*, *edición*, o *adaptación*). Por último, en la descripción de cada registro el traductor podía estar consignado –por orden de menor a mayor reconocimiento como participante en la creación de la obra– en los campos 245, 500 y 700. De hecho, en algunos registros no se mencionaba siquiera esa labor.

A continuación, se decidió abordar el objeto de estudio desde un enfoque cuantitativo, aprovechando la circunstancia de contar con un conjunto valioso y muy afinado de registros bibliográficos. Se partió de la premisa de que si el nombre del traductor figuraba en el campo 700 ello denotaba la máxima expresión de visibilidad, pues implicaba tener un punto de acceso personal propio a la colección y, además, llevaba aparejado el cuidado en la elección de la forma del nombre. Y se constató que de los 22.071 registros, algo más de dos tercios (14.730) no tenían cumplimentado ningún campo 700 (lo cual, según los niveles establecidos al comienzo de la investigación, indicaba visibilidad parcial) (Gráfico 2).

El siguiente paso fue examinar el subconjunto restante (7.341 registros) cuyo asiento catalográfico sí contaba con, al menos, un punto de acceso secundario de nombre de persona. Tras el tratamiento y conversión de los datos a distintos formatos que permitieran su análisis, se comprobó que a este grupo de documentos le correspondían 15.413 puntos de acceso secundario personales. El siguiente paso era identificar cuántas de esas entradas se habían creado para traductores y cuántas para otro tipo de agentes implicados (editores, ilustradores, etc.).

Para hacer viable el estudio, se estableció la siguiente hipótesis de trabajo: si a cada documento se le habían conferido, de media aproximada, dos campos 700, se convino en que al menos uno de ellos debía haber sido creado para contener un traductor (dado que ser traducciones era una característica verificada que sí compartían todos). En otras palabras, si el 50% de esos registros estaban referidos a traductores, la visibilidad de los mismos era la deseable.

Se optó consecuentemente por aplicar a esa población (15.413) un procedimiento estadístico de obtención de tamaño muestral con ese porcentaje esperado, y estableciendo un nivel de confianza del 95% y un margen de error del 5%. Esta operación nos devolvía una muestra representativa de 375 registros que debía ser cotejada.

El análisis pormenorizado de la muestra ofreció los siguientes datos: el 44% de los puntos de acceso se referían a traductores (incluyendo un 1% de traductores a lenguas no cooficiales del Estado), mientras que un 55% aludían a otro tipo de agentes participantes en la obra. El 1% restante fue calificado como dudoso por no poder confirmar la labor desempeñada por la persona mencionada.

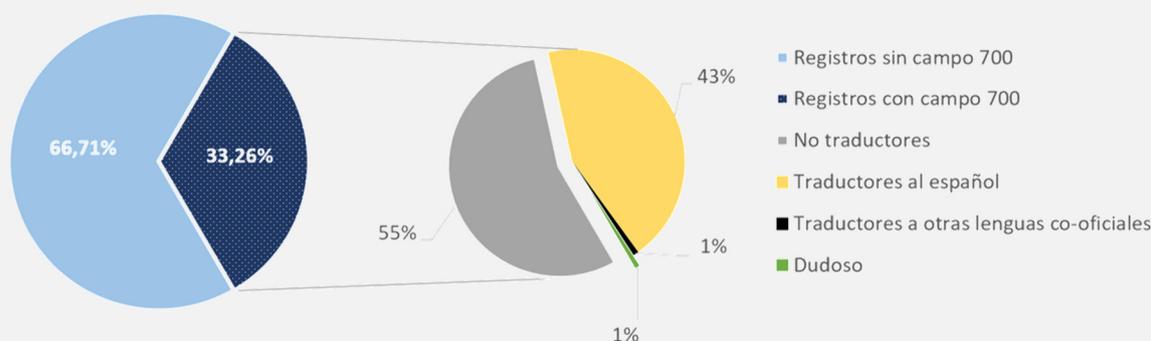


Gráfico 2. Visibilidad del traductor en el campo 700 (Punto de acceso adicional—Nombre de persona)

CONCLUSIONES

En lo referido al cotejo inicial de las fuentes de información empleadas, se pudo extraer una serie de conclusiones:

- Irregular cobertura temporal de los catálogos y bases de datos consultadas.
- Heterogeneidad tanto en los niveles de descripción bibliográfica como en la naturaleza y temática de los documentos incluidos.
- Escasa interoperabilidad real entre las fuentes analizadas.
- Inconsistencia en la consideración de la figura del traductor como responsable de la obra.

Por otro lado, del proceso de tratamiento y examen de la recopilación bibliográfica de obras traducidas del alemán al español se derivan las siguientes reflexiones y líneas futuras de investigación:

- Ninguna fuente bibliográfica es completa en sí misma. La mención al traductor es un dato que se ha ido subsanando de forma retroactiva (a partir de campo 041 del formato MARC, que se refiere a la traducción, no al traductor) y no siempre de manera sistemática.
- En gran medida, los criterios de catalogación dependen de los criterios de edición, que en el caso español también se han visto afectados por cambios jurídicos (por ejemplo, la Ley de Propiedad Intelectual). Este aspecto repercute directamente en la visibilidad del traductor.
- Según los datos obtenidos, la asignación de puntos de acceso secundarios personales en la descripción de una traducción conlleva, en un porcentaje aceptable de los casos, el registro efectivo del nombre del traductor. Por consiguiente, podría afirmarse que la invisibilidad de este está relacionada con la discrecionalidad a la hora de incluir este campo en la descripción, dado que cuando se decide incorporarlo, el traductor alcanza en la práctica cotas cercanas a la visibilidad total.
- Durante mucho tiempo, la visibilidad del traductor no ha sido un objeto de estudio relevante, ni siquiera en trabajos de investigación sobre traducciones. Se considera que esta línea de investigación debe continuarse con el fin de determinar si la representación de la figura del traductor en los catálogos bibliográficos tiene más de causa o de consecuencia respecto a su visibilidad en el ámbito editorial y social.
- Por último, el conjunto documental obtenido permite explotar numerosas vías para comprender la evolución de la traducción, en este caso de obras escritas en lengua alemana, a lo largo de las décadas, así como los condicionantes económicos, políticos y sociales que pueden haber influido en dicha actividad.

REFERENCIAS BIBLIOGRÁFICAS

Calañas Continente, J. A. y Raposo Fernández, B. (2006). *Traducciones españolas de literatura alemana 1972-2006*. CD-ROM. Grupo OSWALD.

Cuéllar Lázaro, M. C. (2000). *Dobletes de traducción y traductología : las traducciones al castellano en España de la literatura contemporánea en lengua alemana (1945 - 1990); estudio lingüístico*. Valladolid : Univ. de Valladolid, Secretariado de Publ. e Intercambio.

Emmerich, K. R. (2013). Visibility (and invisibility). En Gambier, Y. y Van Doorslaer, L. (eds.), *Handbook of Translation Studies* (pp. 200-206). Amsterdam: John Benjamins.

Fortea, C. (2009). *Dos cambios de siglo : ensayos sobre literatura alemana traducida*. Bern [u.a.] : Lang.

Goethe-Institut Madrid (2002). *Bibliografía selectiva de las traducciones de obras literarias del alemán al español desde 1990*. Madrid: Infocentro/Biblioteca Goethe Institut Inter Naciones.

Goethe-Institut Madrid (2004). *Bibliografía selectiva de la traducciones de obras infantiles y juveniles del alemán al español desde 2000*. Madrid: Infocentro/Biblioteca Goethe Institut Inter Naciones.

Goethe-Institut Buenos Aires (2010). *Bibliografía selectiva : traducciones del alemán en la Argentina; 2000-2010*. Buenos Aires: Goethe-Institut Buenos Aires.

Ley de Propiedad Intelectual (texto refundido). (1996). Recuperado de: <https://www.boe.es/buscar/doc.php?id=BOE-A-1996-8930>

Maldonado Alemán, M. (2009). *La narrativa de la unificación alemana : Autores y obras*. - 1. Aufl. Pieterlen : Lang, Peter Bern.

Mönnig, R. (Hrsg.) (1964). *Übersetzungen aus der deutschen Sprache : Spanisch 1948-1963, 2. Aufl.* Göttingen : Vandenhoeck & Ruprecht.

Santana López, B., y Travieso Rodríguez, C. (eds.) (2013). *Puntos de encuentro: los primeros 20 años de la Facultad de Traducción y Documentación de la Universidad de Salamanca*. Salamanca: Ediciones Universidad de Salamanca.

Smith-Yoshimura, Karen. (2014). Challenges posed by translations. En *OCLC Research*. Recuperado de: <http://hangingtogether.org/?p=3878>

Stumpe, Johanna et al. (1978): *Bibliographie deutschsprachiger und deutschlandkundlicher Literatur in spanischer Übersetzung seit 1945*. Bonn-Bad Godesberg : Inter Naciones.

Venuti, Lawrence (1995). *The translator's invisibility: a history of translation*. London: Routledge.



OS VOCABULÁRIOS CONTROLADOS NA ORGANIZAÇÃO E GESTÃO DO PATRIMÓNIO CULTURAL: ORIENTAÇÕES PRÁTICAS

Natália Jorge¹, Filipa Medeiros², Juliana Rodrigues Alves³, Susana Medina⁴

¹CITCEM – Centro de Investigação Transdisciplinar Cultura, Espaço e Memória / Faculdade de Letras da Universidade do Porto/ Doutoranda em Museologia na FLUP, Bolseira de Doutoramento em Empresas (BDE) da FCT/ Sistemas do Futuro, Lda, natalia@sistemasfuturo.pt

²CIDEHUS - Centro Interdisciplinar de História, Culturas e Sociedades da Universidade de Évora, fmcm@uevora.pt

³CITCEM – Centro de Investigação Transdisciplinar Cultura, Espaço e Memória / Faculdade de Letras da Universidade do Porto/ Doutoranda em Museologia na FLUP, julira@yahoo.com

⁴CITCEM – Centro de Investigação Transdisciplinar Cultura, Espaço e Memória / Faculdade de Letras da Universidade do Porto/ Doutoranda em Museologia na FLUP, Museu da Faculdade de Engenharia, smedina@fe.up.pt

RESUMO Pretende-se apresentar o trabalho desenvolvido no âmbito do Grupo de Trabalho Sistemas de Informação em Museus da Associação Portuguesa de Bibliotecários, Arquivistas e Documentalistas (GT-SIM/BAD), denominado *Os vocabulários controlados na organização e gestão do património cultural: orientações práticas* (Jorge, Medeiros, Alves, & Medina, 2017), que tem como objetivo ser uma ferramenta facilitadora em língua portuguesa para a utilização e criação de vocabulários controlados em instituições de memória (arquivos, museus e bibliotecas). Este trabalho utilizou como base a experiência académica e profissional da equipa da sublinha *Terminologias* do GT-SIM e o livro de Patricia Harpring *Introduction to Controlled Vocabularies: Terminology for Art, Architecture, and Other Cultural Works* (2013). Trata-se de um documento em desenvolvimento, passível de vir a ser incorporado com novos elementos resultantes do debate que a sua leitura venha a suscitar, estando disponível online no website da Associação Portuguesa de Bibliotecários, Arquivistas e Documentalistas (BAD). Está ainda prevista a adição futura de novas ferramentas de apoio, como glossários e *toolkits*.

PALAVRAS-CHAVE *Património cultural, Vocabulários controlados, Organização da informação, Gestão da informação, Instituições de memória.*

ABSTRACT The aim of the present paper is to describe the work developed within the Working Group on Information Systems in Museums (WG-ISM) of the Portuguese Association of Librarians, Archivists and Documentation Professionals (GT-SIM / BAD), entitled *Controlled vocabularies in the organization and management of cultural heritage: a practical guide* (Jorge, Medeiros, Alves, & Medina, 2017). This guide aims to be a facilitator tool in portuguese for the use and creation of controlled vocabularies in memory institutions (archives, museums and libraries). It has been based on the academic and professional experience of the WG-ISM-Terminologies' team members and on Patricia Harpring's book *Introduction to Controlled Vocabulary: Terminology for Art, Architecture, and Other Cultural Works* (2013). This guide should be considered as a work-in-progress document which can be incorporated with new elements resulting from the debate that its reading will provoke. It is available online on the website of the Portuguese Association of Librarians, Archivists and Documentation Professionals. The future addition of new support tools, such as glossaries and *toolkits*, is also foreseen.

KEYWORDS *Cultural heritage, Controlled Vocabularies, Information organization, Information management, Memory institutions*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Esta comunicação tem como objetivo apresentar o guia *Os vocabulários controlados na organização e gestão do património cultural: orientações práticas* (2017), criado pela sublinha *Terminologias* do Grupo de Trabalho Sistemas de Informação em Museus da Associação Portuguesa de Bibliotecários, Arquivistas e Documentalistas (GT-SIM/BAD).

Importa aqui fazer uma breve apresentação do GT-SIM, de molde a compreender-se melhor o contexto e os objetivos do trabalho que se expõem nesta comunicação. Este grupo de trabalho foi aprovado e criado pela Associação Portuguesa de Bibliotecários, Arquivistas e Documentalistas (BAD) em 2012, ano durante o qual ficaram também definidos a missão, os objetivos estratégicos, os conceitos operatórios e a organização do trabalho a realizar pelos respetivos membros.

O GT-SIM procura pensar o Museu como um centro de produção de conhecimento ao assumir o objeto de museu como documento e o acervo da instituição museológica, existente nas Reservas, Arquivo, Biblioteca ou Centro de Documentação como um todo unitário nas suas inter-relações informacionais. A visão integradora do acervo do Museu implica um maior enfoque nas potencialidades informativas do acervo, contribuindo assim para uma mais eficiente gestão de toda a informação sobre património produzida em contexto museológico (BAD, n.d.).

A partir de 2012, o GT-SIM definiu linhas de ação diferenciadas com a finalidade de concretizar os objetivos estratégicos que havia definido e que são:

- 1) Diagnóstico aos sistemas de informação nos museus portugueses - esta linha tem como objetivo promover o levantamento e a caracterização dos museus portugueses relativamente às áreas da gestão da informação atinentes aos respetivos bens patrimoniais. Foi já efetuado o *Diagnóstico aos Sistemas de Informação nos Museus Portugueses: Relatório final* (que retomaremos mais adiante em detalhe), apresentado em abril de 2017 numa sessão pública que contou com a participação de vários profissionais ligados aos sistemas de informação nos museus portugueses.
- 2) Metodologias e procedimentos a utilizar pelos profissionais nos museus - esta linha subdivide-se em três sublinhas.
 - 2.1) Vocabulários controlados na organização e gestão da informação sobre património cultural (é o trabalho desta sublinha que se dá a conhecer em pormenor nesta comunicação).
 - 2.2) Tradução do documento *Cataloguing Cultural Objects* (CCO), norma desenvolvida pela VRA – *Visual Resources Association*, em 2006.
 - 2.3) Tradução de guias técnicos, que constituem um conjunto de orientações e aconselhamento para a implementação da norma SPECTRUM na gestão das coleções dos museus.

- 3) Constituição de um Centro de Documentação Virtual (CDV) em livre acesso – esta linha tem como finalidade proporcionar recursos de informação relacionados com as áreas de investigação do GT. Para tal recorreu-se à plataforma Zotero (gestor de referências bibliográficas) para criar uma biblioteca de grupo, que conta já com mais de 300 recursos organizados por coleções e devidamente indexados a partir dos descritores de vários tesouros existentes. Os recursos incidem sobre a gestão da informação nos museus, bibliotecas e arquivos e, ainda, sobre a normalização documental dos acervos.
- 4) Seminários, conferências e encontros - estas iniciativas procuram valorizar e promover o GT no meio profissional e académico, ao mesmo tempo destacam a importância dos sistemas de informação em museus como pilares fundamentais na gestão dos acervos patrimoniais.
- 5) Divulgação - a promoção e divulgação da atividade do GT-SIM constitui-se como um eixo fundamental do grupo. Os três meios de comunicação mais importantes são: Notícia BAD (jornal em linha da BAD, através da categoria «info em museus», com mais de 90 notícias publicadas); participação nas redes sociais e dinamização de um grupo no Facebook; e, ainda, participação e dinamização da mailing list.

Importa salientar que as linhas e sublinhas do GT-SIM interagem constantemente, embora tenham âmbitos de ação diferenciados, complementam-se no sentido de concretizar os objetivos estratégicos estabelecidos, orientados para o desenvolvimento de sistemas integrados de informação nos museus.

Após a apresentação do contexto de realização passa-se à apresentação do referido guia. Este trabalho pretende ser uma ferramenta facilitadora em língua portuguesa para a utilização de terminologia controlada nas instituições de memória. Utilizou como base a experiência académica e profissional da equipa da sublinha *Terminologias* do GT-SIM e o livro de Patricia Harpring *Introduction to Controlled Vocabularies: Terminology for Art, Architecture, and Other Cultural Works* (2013). O guia apresenta os tipos mais comuns de vocabulários controlados, exemplos de aplicação nacionais e internacionais, e orientações para a criação e uso de vocabulários controlados. É de sublinhar que este documento não constitui um fim em si mesmo, encontrando-se em constante atualização e crescimento, na medida em que se encontra aberto à discussão e a eventuais alterações. A equipa pretende alargar a discussão aos profissionais que trabalham na área, ouvir as suas contribuições, e lançar para o próximo ano (2018) uma nova edição em que serão incluídos um glossário e ferramentas de apoio.

METODOLOGIA

Nos últimos anos têm sido publicadas ferramentas de controlo terminológico em língua portuguesa na área do património cultural, resultantes de projetos de investigação e de parcerias entre instituições nacionais e internacionais (como por exemplo, o projeto *Thesaurus de Acervos Científicos em Língua Portuguesa*, coordenado pelo Museu Nacional de História Natural e da Ciência da Universidade de Lisboa e pelo Museu de Astronomia e Ciências Afins do Rio de Janeiro). Paralelamente, existem iniciativas individuais de instituições que desenvolvem ferramentas específicas, as quais surgiram para fazer face à necessidade de dar respostas aos próprios projetos de documentação. Tendo em vista a partilha entre os profissionais, a divulgação de projetos na área das linguagens documentais, a promoção do debate e da sua fundamentação num corpo teórico ajustado ao desenvolvimento de vocabulários controlados, percebeu-se a necessidade de um documento orientador que auxiliasse a discussão sobre a

utilização das ferramentas atualmente existentes (como os tesouros, lista de termos, etc.) ou que ajudasse à criação de novas ferramentas específicas. Para dar resposta a estas demandas organizamos seminários (*A utilização dos thesauri nos sistemas de gestão do património cultural*, em 2013, e *Sistemas de Organização do Conhecimento em Museus*, em 2015), construímos uma lista de referências sobre vocabulários controlados no Centro de Documentação Virtual CDV (Almeida, Ferreira, Medeiros, Patrão, & Salgado, n.d.) e publicamos o documento *Os vocabulários controlados na organização e gestão do património cultural: orientações práticas* (Jorge, Medeiros, Alves, & Medina, 2017).

O público-alvo das diferentes iniciativas é a comunidade de profissionais ligados às instituições de memória, especialmente aqueles que veem a adoção de vocabulários controlados como uma ferramenta indispensável na organização e na gestão do património cultural, bem como os demais interessados, quer da comunidade académica e científica, quer da sociedade civil.

A ideia e construção de um documento orientador surgiu no âmbito do debate empreendido no seio do grupo de trabalho GT-SIM e na sequência da realização de seminários, estando também relacionado com as atividades do Centro de Documentação Virtual (CDV), disponível para consulta online na plataforma Zotero. A partir do trabalho de recolha de referências e organização da informação do CDV percebeu-se a mais-valia de disponibilizar ao público-alvo um guia facilitador da utilização das referências disponíveis no Zotero (e de outras) e que servisse também como auxiliar do trabalho quotidiano de organização de informação em instituições de memória.

O CDV reúne na sua biblioteca virtual, e apenas na área de vocabulários controlados, mais de 100 referências online. A seleção de recolha destas referências foi realizada primeiramente com a contribuição dos membros do GT-SIM, que compreende profissionais e académicos na área da organização da informação em instituições de memória. Como metodologia de trabalho foi realizada uma pesquisa exploratória e a descrição em listagem de diversas fontes que têm disponíveis online vocabulários controlados. Posteriormente foi realizada a conferência dos dados recolhidos, nomeadamente as informações primárias: link do website disponível, título, autor(es), data, idiomas, tipo de vocabulário, inclusão de resumo (abstract), e foram criadas tags para facilitar a posterior recuperação da informação (aqui utilizamos como referência o tesouro da Unesco). Por fim, as referências têm sido constantemente atualizadas no CDV.

O guia é um complemento ao referido trabalho, e surge para responder à seguinte questão: de que forma os utilizadores, ou seja, os profissionais das instituições de memória que trabalham com a organização da informação, podem utilizar os vocabulários controlados referenciados no CDV, entre outros?

Para dar resposta a esta questão, e para a construção de um guia de apoio, foi feito um levantamento bibliográfico sobre o uso e a criação de vocabulários controlados. Para além das normas produzidas em língua portuguesa, usamos como base de referência a publicação da autoria de Patricia Harpring, *Introduction to Controlled Vocabularies: Terminology for Art, Architecture, and Other Cultural Works* (2013). O livro de Patricia Harpring é um guia detalhado para a construção de vocabulários controlados, de índole técnica, que aborda questões conceituais pertinentes para os profissionais da área, reunindo, desenvolvendo e sistematizando reflexões sobre a natureza dos vocabulários controlados, a sua construção, as suas diferentes tipologias, potencialidades, dificuldades e exigências no atual contexto de organização e gestão de informação sobre o património cultural. Por conseguinte, respondia aos objetivos traçados pela equipa da sublinha *Terminologias* do GT-SIM, que o tentou adaptar ao panorama português, usando uma linguagem acessível e referenciando, sempre que possível, exemplos nacionais. Para isso, paralelamente, realizou-se um levantamento de ferramentas terminológicas e

exemplos de aplicação de terminologia, nacionais e internacionais, algumas das quais foram referenciadas no guia, de maneira a contextualizar e a localizar as mesmas para o contexto português. A orientação definida partiu de exemplos passíveis de complementar o texto, em primeiro lugar retirados do contexto português, e seguidamente da Europa ou do Brasil.

Para além disso, foram utilizados como apoio outros documentos da mesma autora que abordam a temática dos vocabulários controlados (Harpring, 2015, 2016) e os *Guidelines do Getty Research Institute* (J. Paul Getty Trust, n.d.).

Pretende-se que o guia, disponível no site da BAD (Jorge, Medeiros, Alves, & Medina, 2017), seja discutido entre os pares através de diferentes canais de comunicação, durante o ano de 2017, e que a partir da discussão gerada sejam criadas novas ferramentas complementares de apoio, como *toolkits*, a serem publicadas numa nova edição.

RESULTADOS

Um dos primeiros resultados a apresentar, fruto do trabalho de levantamento de ferramentas terminológicas nacionais e internacionais desenvolvido pela sublinha *Terminologias* do GT-SIM, consiste na disponibilização, no Centro de Documentação Virtual (CDV), de uma área dedicada a agrupar referências, intitulada *Vocabulários Controlados*.

A principal realização da sublinha *Terminologia*, sobre a qual nos debruçamos na presente apresentação, consiste na publicação do guia *Os vocabulários controlados na organização e gestão do património cultural: orientações práticas* (Jorge, Medeiros, Alves, & Medina, 2017).

Os objetivos deste guia são:

- 1) Sensibilizar a comunidade de profissionais das instituições de memória para a importância da normalização terminológica no tratamento da informação.
- 2) Definir e caracterizar um conjunto de conceitos-chave no âmbito da organização e gestão da informação.
- 3) Caracterizar os diferentes tipos de vocabulários controlados e respetivos domínios de aplicação.
- 4) Dar a conhecer recursos e projetos de referência, nacionais e internacionais, que possam servir de base de apoio ao desenvolvimento/aperfeiçoamento de outros já existentes.
- 5) Oferecer aos profissionais um conjunto de orientações gerais para a construção de vocabulários controlados de natureza diversa, bem como de registos de autoridade.

O guia apresenta um conjunto de orientações sobre o uso e a criação de vocabulários controlados, organizando-se em nove capítulos, a saber:

- 1) Apresentação do projeto.
- 2) Referência aos pré-requisitos para o desenvolvimento de projetos de documentação (estrutura de dados, procedimentos de registo, sintaxe de dados e terminologia).
- 3) Apresentação de normas nacionais e internacionais para o desenvolvimento de vocabulários controlados.

- 4) Quadro metodológico e conceptual.
- 5) Tipologias de vocabulários controlados.
- 6) Ferramentas relacionadas com vocabulários controlados.
- 7) Projetos de referência nacionais e internacionais.
- 8) Orientações para a construção de vocabulários controlados.
- 9) Perspetivas futuras de desenvolvimento.

De salientar que este guia detém um carácter pioneiro no panorama da organização e da gestão da informação nas instituições de memória em Portugal, a dois níveis: por um lado, e do ponto de vista teórico, assume-se como a primeira proposta metodológica relativa à catalogação do património cultural, partindo, para tal, de um quadro referencial teórico nacional e internacional; por outro, e do ponto de vista profissional, tendo nascido no seio de um grupo de trabalho, constitui-se como uma ferramenta que poderá ser afinada, adaptada e aprofundada por cada instituição, à medida das suas necessidades e especificidades.

Como já referimos, a produção deste guia nasceu fruto da discussão realizada no GT-SIM sobre necessidade de desenvolver a temática e promover o debate sobre os vocabulários controlados na organização e gestão da informação sobre património cultural, mediante a realização de seminários e produção de documentação orientadora apoiada em bibliografia de referência. Esta necessidade encontra-se devidamente fundamentada na recente publicação *Diagnóstico aos Sistemas de Informação nos Museus Portugueses: Relatório final* (Santos, Serôdio, & Ferreira, 2017). A caracterização dos museus portugueses, patente neste documento, assenta em várias dimensões: categorias e tipo de bens, modos de incorporação, número de bens e modalidades, informatização da coleção, documentos orientadores, quadro normativo, sistema de vocabulário controlado e, por fim, a interligação entre bases de dados (Santos, Serôdio, & Ferreira, 2017, p. 47). Na nossa perspetiva interessa-nos fazer aqui o recorte relativo ao uso de vocabulários controlados, e é com esse foco que desenvolvemos a análise que de seguida se apresenta. Para o efeito, consideramos que deverão ser abordadas outras dimensões do estudo para uma melhor compreensão, como a quantidade de bens inventariados e a informatização das coleções. Isto porque tais vertentes estão intimamente ligadas ao uso dos vocabulários controlados, a primeira porque nos indica o ponto de situação quanto à totalidade de bens inventariados, permitindo assim estabelecer a comparação com o total de registos informatizados, e a segunda por consideramos que a informatização dos acervos requer, para avaliar a consistência de dados e uma pesquisa eficaz, a utilização de vocabulários controlados.

O *Diagnóstico* é uma referência contextual, uma vez que poderá ser o ponto de partida para aprofundar políticas culturais e definir estratégias de atuação. No âmbito do presente trabalho a relevância deste documento prende-se com o facto de representar a quantificação do uso de vocabulários controlados, a qual por sua vez vai ao encontro da realidade por nós constatada empiricamente em virtude dos contactos com diferentes profissionais e museus.

O *Diagnóstico* assenta no método quantitativo e na técnica de inquérito extensivo por questionário (autoadministrado), dirigido aos responsáveis dos museus, utilizando como suporte de recolha de informação a plataforma online (web survey) - Survey Monkey. O universo de análise foi composto por todas as entidades autodesignadas museu com um funcionamento permanente ou sazonal, o que resultou num total de 710 museus. O levantamento da informação foi realizado entre 13 de março e 10 de junho de 2016, sendo os dados recolhidos relativos ao ano de 2015. O número de respostas válidas resultou numa amostra de 222 museus (Santos, Serôdio, & Ferreira, 2017, p. 11). O questionário foi estruturado

em 9 grupos, compostos por um total de 89 questões, subdivididas em 56 perguntas e 33 desdobramentos (Santos, Serôdio, & Ferreira, 2017, p. 18).

De uma forma genérica, no que respeita aos acervos museológicos, bibliográficos e arquivísticos, constata-se que cerca de dois terços dos museus possuem pelo menos dois tipos de acervo, sendo o mais comum o museológico e o bibliográfico. Os museus que possuem apenas acervo museológico constituem apenas 35% dos casos, sendo que 32% referiu possuir bens em todos os acervos (Santos, Serôdio, & Ferreira, 2017, p. 80).

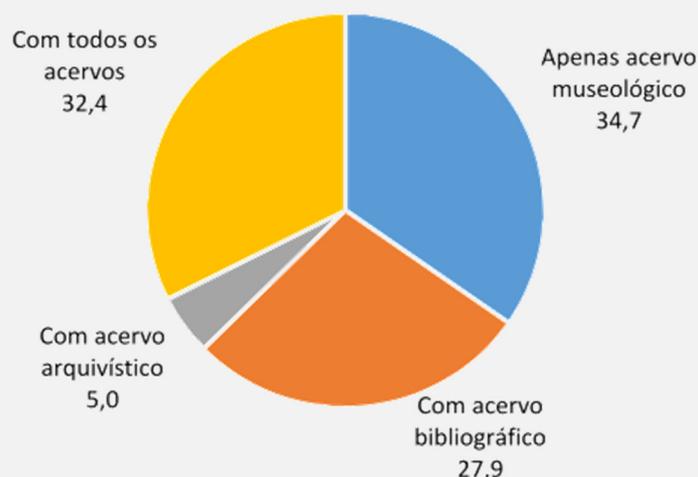


Figura 6 Museus por tipo de acervo (Santos, Serôdio, & Ferreira, 2017, p. 80)

Quanto à informatização do inventário destes acervos, salienta-se que oito em cada dez museus assinalaram ter o inventário em suporte informático. No entanto, ainda havia 17% que não possuía quaisquer bens dos seus acervos em base de dados. Nesse sentido, verifica-se que tanto um terço dos museus possui pelo menos dois dos seus acervos informatizados, destacando-se de novo os acervos museológico e o bibliográfico, enquanto outro terço dos museus apenas tem informatizado o seu acervo museológico. Com os três acervos nessa condição encontram-se 17% dos museus (Santos, Serôdio, & Ferreira, 2017, p. 116).

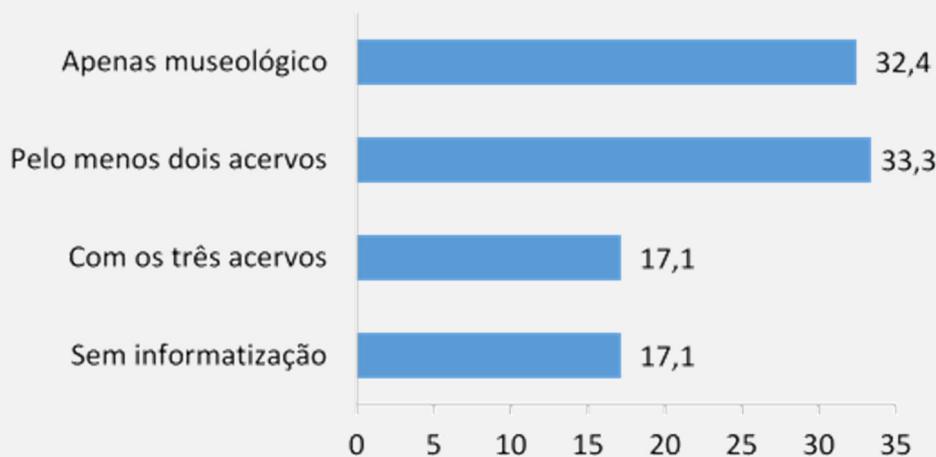


Figura 2 Museus por informatização dos acervos (Santos, Serôdio, & Ferreira, 2017, p. 83)

Partindo para um nível específico, nas respostas à questão sobre o Total de bens museológicos e de bens com cadastro (idêntico ao registo em livro de tombo), com inventário, fotografados, em base de dados (informatizados) e em exposição, os autores constataam, no que diz respeito ao número de museus que responderam, percentagens relativamente elevadas de resposta a qualquer uma das opções da forma de registo. Tendo agora em conta o número de bens, verifica-se que essas percentagens são substancialmente mais baixas, registando-se a mais elevada nos bens com cadastro (36%), seguida dos bens com inventário (34%). É relevante referir que, no que respeita à informatização e aos bens fotografados, a partir das respostas obtidas, as percentagens não ultrapassam os 10%, situando-se os bens informatizados em 9% e os bens em suporte fotográfico em 8% (Santos, Serôdio, & Ferreira, 2017, p. 50).

Forma de registo	Número de casos	Percentagem	Total de bens	Percentagem
Com cadastro	124	59,9	4.789.360	35,8
Com inventário	178	86,0	4.609.212	34,4
Fotografados	153	73,9	997.499	7,5
Em Base de dados	165	79,7	1.157.933	8,7
Em exposição	171	82,6	176.532	1,3
Número de casos	207			
Total de bens			13.384.242	

Figura 3 Total de bens museológicos e de bens com cadastro, com inventário, fotografados, em base de dados (informatizados) e em exposição (Santos, Serôdio, & Ferreira, 2017, p. 49)

Tal como já foi referido, relativamente à informatização do acervo museológico, os autores concluíram que oito em cada dez museus possuíam uma aplicação informática para a gestão do acervo (Santos, Serôdio, & Ferreira, 2017, p. 50).

Quantitativamente, e com base na realidade de 2015, a maioria dos museus referiu ter o processo em curso (parcial), ao passo que apenas 15% dos museus possui o seu acervo todo informatizado. De referir ainda que os autores constaram a inexistência de informatização das coleções em cerca de um quarto dos museus (Santos, Serôdio, & Ferreira, 2017, p. 52).

Chegados ao ponto pretendido, consistente nas respostas válidas à pergunta *O museu utiliza algum sistema de vocabulário controlado/thesauri na indexação, descrição ou classificação do seu acervo museológico? Se utiliza refira qual*, a grande maioria dos museus (70%) que respondeu a esta questão menciona que não usa vocabulários controlados na descrição ou classificação dos acervos. Tendo em conta os que referem utilizar vocabulários controlados (23%, o que corresponde a 52 casos), 21% corresponde a referências cujos autores dizem não ter sido possível identificar e/ou filiar em qualquer das categorias usadas para o tratamento desta resposta aberta, 15% responde com a indicação da designação comercial do software utilizado, podendo denotar, como referem os autores, alguma dificuldade em distinguir entre ferramentas terminológicas e sistemas de gestão de informação, 15% menciona os vocabulários desenvolvidos internamente e 15% refere a utilização de documentos normativos nacionais, com preponderância para as publicações da entidade da tutela dos museus e património cultural, 12% alude a documentos normativos internacionais e 6% a projetos de tesouros temáticos em língua portuguesa (Santos, Serôdio, & Ferreira, 2017, pp. 57-58).



Figura 4 Vocabulário controlado utilizado (Santos, Serôdio, & Ferreira, 2017, p. 58)

O panorama relativo à utilização dos vocabulários controlados traçado com base nas respostas apresentadas no *Diagnóstico* é indicador da relevância da continuidade do debate e discussão sobre esta temática, não só pelo elevado número de respostas que indica não os utilizar, mas também pela aparente confusão concetual entre ferramentas terminológicas e sistemas de informação. Do mesmo se conclui que há necessidade de continuar a produzir documentação de apoio.

Os resultados obtidos com esta investigação traduziram-se na publicação do guia que se apresenta e discute aqui. É de notar que este documento não constitui um fim em si mesmo, encontrando-se em constante atualização e crescimento e estando aberto à discussão e a eventuais alterações, fruto de desenvolvimentos das problemáticas apresentadas que venham a ocorrer. Salienta-se, também, que caberá a cada profissional e instituição utilizadora do guia criar, fomentar e aplicar as suas próprias orientações, atendendo às especificidades inerentes a cada acervo. Por esta razão reitera-se a ideia de que este documento não se trata de um modelo a seguir, funcionando, tão-somente, como um guia de apoio e, em última análise, como uma proposta metodológica suscetível de ser aplicada a diferentes realidades.

CONCLUSÕES

Por se tratar de um documento em desenvolvimento, este guia está aberto à discussão teórica e metodológica. Neste sentido, pretende fornecer à comunidade de profissionais das instituições de memória um conjunto de orientações práticas que permitam uma eficaz e consistente organização e gestão da informação através do uso de vocabulários controlados.

A partir dos contributos que se espera receber na sequência da discussão gerada e da conseqüente revisão deste documento, estima-se a criação de um glossário e de *toolkits*, ferramentas que são auxiliares da gestão da informação nas instituições de memória, ajudando na criação de raiz de vocabulários controlados ou na normalização de terminologias já existentes.

Por fim, importa ainda referir que existe um trabalho de campo que está a ser efetuado junto de alguns museus portugueses, assim como de profissionais da área, no sentido de testarem o guia, esperando-se em breve ter o retorno dessa experiência, a fim de aprimorar o seu conteúdo.

Convidamos todos os interessados a lerem, participarem nas ações do GT-SIM, discutirem e contribuir para esta publicação, orientada para aqueles que desejam estudar e trabalhar na organização e gestão da informação em património cultural.

REFERÊNCIAS BIBLIOGRÁFICAS

Almeida, M. J. de, Ferreira, F., Medeiros, F., Patrão, S., & Salgado, A. (n.d.). Centro de documentação virtual|Zotero |Groups GT_SIM. Retrieved April 28, 2017, from <https://www.zotero.org/groups/81851/gt-sim>

BAD. (n.d.). Secções e Grupos de Trabalho da BAD. Retrieved May 25, 2017, from <http://www.apbad.pt/Seccoes/Seccoes.htm>

Harpring, P. (2013). *Introduction to Controlled Vocabularies Terminology for Art, Architecture, and Other Cultural Works*. Los Angeles, Califórnia: Getty Research Institute.

Harpring, P. (2015). *Introduction to Controlled Vocabularies: Featuring the Getty Vocabularies*. Getty Vocabulary Program. Retrieved from https://www.getty.edu/research/tools/vocabularies/intro_to_vocab.pdf

Harpring, P. (2016). *Introdução aos vocabulários controlados: terminologia para arte, arquitetura e outras obras culturais*. (Secretaria da Cultura do Estado de São Paulo, Pinacoteca de São Paulo, & ACAM Portinari, Eds.). São Paulo.

J. Paul Getty Trust. (n.d.). Getty Vocabularies Editorial Guidelines (Getty Research Institute). J. Paul Getty Trust. Retrieved from <http://www.getty.edu/research/tools/vocabularies/index.html>

Jorge, N., Medeiros, F., Alves, J. R., & Medina, S. (2017). *Os vocabulários controlados na organização e gestão de informação sobre património cultural: orientações práticas*. (Grupo de Trabalho Sistemas de Informação em Museus (GT-SIM) & Associação Portuguesa de Bibliotecários Arquivistas e Documentalistas (BAD), Eds.). Retrieved from https://www.bad.pt/noticia/wp-content/uploads/2017/04/Guia_VocabulariosControlados_final-1.pdf

National Information Standards Organization. (n.d.). ISO 25964 – the international standard for thesauri and interoperability with other vocabularies. Retrieved May 22, 2016, from <http://www.niso.org/schemas/iso25964>

Matos, A., Braga, A., Serafim, C., Cortês, C., Correia, E., Alves, J. R., ... António, R. (2017). Guias técnicos de implementação do SPECTRUM PT – Uma ajuda adicional para a implementação da norma SPECTRUM: Notícia BAD. Grupo de Trabalho Sistemas de Informação em Museus (GT-SIM) Associação Portuguesa de Bibliotecários Arquivistas e Documentalistas (BAD). Retrieved from <https://www.bad.pt/noticia/2017/04/03/guias-tecnicos-de-implementacao-do-spectrum-pt-uma-ajuda-adicional-para-a-implementacao-da-norma-spectrum/>

Museu Nacional de História Natural e da Ciência da Universidade de Lisboa (Coord.), & Museu de Astronomia e Ciências Afins do Rio de Janeiro (Coord.). (2013). *Thesaurus de Acervos Científicos em Língua Portuguesa*. Retrieved from <http://thesaurusonline.museus.ul.pt>

Organização das Nações Unidas para a Educação, a Ciência e a Cultura - (UNESCO). (n.d.). *UNESCO Thesaurus*. Retrieved from <http://vocabularies.unesco.org/browser/thesaurus/en/>

Portugal Biblioteca Nacional CT7. (1992). Norma portuguesa NP 4036 1992. Documentação: Tesouros monolíngues: directivas para a sua construção e desenvolvimento. Lisboa: IPQ - Instituto Português da Qualidade.

Rodríguez Bravo, B. (Dir), Alvite Díez, M. L., Díez Díez, Á., Gallego Lorenzo, J., López García, A., Morán Suárez, M. A., ... Santos De Paz, L. (2004). *Tesouro de Ciencias de la Documentación*. Universidad de León. Retrieved from <https://buleria.unileon.es/handle/10612/3003>

Santos, J., Seródio, C., & Ferreira, F. (2017). *Diagnóstico aos Sistemas de Informação nos Museus Portugueses: Relatório final*. Grupo de Trabalho Sistemas de Informação em Museus (GT-SIM) da Associação Portuguesa de Bibliotecários, Arquivistas e Documentalistas (BAD). Retrieved from <https://www.bad.pt/noticia/2017/05/18/publicacao-dos-resultados-do-diagnostico-aos-sistemas-de-informacao-nos-museus-portugueses/>

VOCABULÁRIOS CONTROLADOS DE REFERÊNCIA

Alfaro Asins, C., Marcos Alonso, C., Otero Morán, P., & Grañeda Miñón, P. (n.d.). *Diccionario de numismática*. Secretaría General Técnica. Centro de Publicaciones. Ministerio de Cultura. Retrieved from <https://sede.educacion.gob.es/publiventa/diccionario-de-numismatica/filatelia-numismatica/13594C>

Archive of Digital Art (ADA), Grau, O., Coones, W., Rühse, V., Hoth, J., & Schiller, D. (1999–2004). *Media Art Research Thesaurus* [Meta-Thesaurus]. Retrieved from <http://mediaartresearch.org/search/thesaurus-hierarchical.html>

British Museum, & Department of Scientific Research. (n.d.). *British Museum Materials Thesaurus*. Retrieved from <http://terminology.collectionstrust.org.uk/British-Museum-materials/>

British Museum. (n.d.). *British Museum Object Names Thesaurus*. Retrieved from <http://terminology.collectionstrust.org.uk/British-Museum-objects/>

CABI - Centre for Agriculture and Biosciences International. (1983–2016). *CAB Thesaurus*. Retrieved from <http://www.cabi.org/cabthesaurus/mtwdk.exe?yi=home>

Campos, T., & Ventura, J. M. (2003). *Terminologia controlada para a indexação de documentos na área da arquitectura*. Lisboa: Biblioteca Nacional.

Centro Nacional de Folclore e Cultura Popular do Instituto do Património Histórico e Artístico Nacional/MinC. (n.d.). *Tesouro de Folclore e Cultura Popular*. Retrieved from <http://www.cnfcp.gov.br/tesouro>

Council of Europe, & European Heritage Network. (n.d.). *Cultural Heritage Thesaurus (HEREIN)*. Retrieved from <http://www.herein-system.eu/thesaurus-presentation>

Dag Hammarskjöld Library, & United Nations Department of Public Information. (n.d.). *UNBIS Thesaurus*. Retrieved from <http://lib-thesaurus.un.org/LIB/DHLUNBISThesaurus.nsf>

DIBAM - Dirección de Bibliotecas, Archivos y Museos - Gobierno de Chile. (n.d.). *Tesouro Regional Patrimonial*. Retrieved from <http://www.tesouroregional.cl/trp/publico/buscar.htm>

European Education Thesaurus. (2009). Retrieved from <http://vocabularyserver.com/tee/en/>

Ferrez, H. D., Castro, A. H. F. de, Almeida, Á., Santos, H. M. dos, Ayres, J. M. R. A., Peregrino, J., ... Peixoto, M. E. S. (n.d.). *Tesouro de Objetos do Patrimônio Cultural nos Museus Brasileiros*. Retrieved from <http://www.tesouromuseus.com.br>

Government of Canada, & Library and Archives Canada. (2007). *dc.subject Vocabularies - Controlled Vocabularies - Information Management - Library and Archives Canada*. Retrieved from <http://www.collectionscanada.gc.ca/government/controlled-vocabularies/007004-110.01-e.php?q1=subject>

Historic England. (2015). *Archaeological Sciences Thesaurus*. Retrieved from http://thesaurus.historicengland.org.uk/thesaurus.asp?thes_no=560

Historic England. (n.d.-a). *Defence of Britain Thesaurus*. Retrieved from http://thesaurus.historicengland.org.uk/thesaurus.asp?thes_no=365&thes_name=Defence%20of%20Britain%20Thesaurus

Historic England. (n.d.-b). *EH and National Trust Resource Description Thesaurus*. Retrieved from http://thesaurus.historicengland.org.uk/thesaurus.asp?thes_no=547&thes_name=EH%20And%20National%20Trust%20Resource%20Description%20Thesaurus

Historic England. (n.d.-c). *FISH - Maritime Cargo Types Thesaurus*. Retrieved from http://thesaurus.historicengland.org.uk/thesaurus.asp?thes_no=77&thes_name=FISH%20Maritime%20Cargo%20Types%20Thesaurus

Historic England. (n.d.-d). *FISH Building Materials Thesaurus*. Retrieved from http://thesaurus.historicengland.org.uk/thesaurus.asp?thes_no=129&thes_name=FISH%20Building%20Materials%20Thesaurus

Historic England. (n.d.-e). *FISH Components Thesaurus*. Retrieved from http://thesaurus.historicengland.org.uk/thesaurus.asp?thes_no=546&thes_name=FISH%20Components%20Thesaurus

Historic England. (n.d.-f). *FISH Evidence Thesaurus*. Retrieved from http://thesaurus.historicengland.org.uk/thesaurus.asp?thes_no=92&thes_name=FISH%20Evidence%20Thesaurus

Historic England. (n.d.-g). *FISH Historic Aircraft Type Thesaurus*. Retrieved from http://thesaurus.historicengland.org.uk/thesaurus.asp?thes_no=225&thes_name=FISH%20Historic%20Aircraft%20Type%20Thesaurus

Historic England. (n.d.-h). *FISH Maritime Craft Types Thesaurus*. Retrieved from http://thesaurus.historicengland.org.uk/thesaurus.asp?thes_no=143&thes_name=FISH%20Maritime%20Craft%20Types%20Thesaurus

Historic England. (n.d.-i). *FISH Maritime Place Name Thesaurus*. Retrieved from http://thesaurus.historicengland.org.uk/thesaurus.asp?thes_no=145&thes_name=FISH%20Maritime%20Place%20Name%20Thesaurus

Historic England. (n.d.-j). *FISH Thesaurus of Monument Types*. Retrieved from http://thesaurus.historicengland.org.uk/thesaurus.asp?thes_no=1&thes_name=FISH%20Thesaurus%20of%20Monument%20Types

ICOM International Committee for the Museums and Collections of Costume. (n.d.). *Vocabulary of Basic Terms for Cataloguing Costume*. Retrieved from http://old.collectionstrust.org.uk/assets/thesaurus_icombts/vbt00e.htm

Istituto Centrale per il Catalogo e la Documentazione (ICCD). (n.d.). *Strumenti terminologici*. Retrieved from <http://www.iccd.beniculturali.it/index.php?it/473/standard-catalogfici/Standard/95>

J. Paul Getty Trust, & Getty Research Institute. (n.d.-a). *Art & Architecture Thesaurus (AAT)*. Retrieved from <http://www.getty.edu/research/tools/vocabularies/aat/>

J. Paul Getty Trust, & Getty Research Institute. (n.d.-b). *Cultural Objects Name Authority (CONA)*. Retrieved from <http://www.getty.edu/research/tools/vocabularies/cona/index.html>

J. Paul Getty Trust, & Getty Research Institute. (n.d.-c). *Thesaurus of Geographic Names (TGN)*. Retrieved from <http://www.getty.edu/research/tools/vocabularies/tgn/index.html>

J. Paul Getty Trust, & Getty Research Institute. (n.d.-d). *Union List of Artists Names Online (ULAN)*. Retrieved from <http://www.getty.edu/research/tools/vocabularies/ulan/index.html>

Junta de Andalucía: Consejería de Educación, Cultura y Deporte. (n.d.). *Tesouro de Património Histórico Andaluz*. Retrieved from <http://www.iaph.es/tesouro/init.htm>

Kroustallis, S. K. (2008). *Diccionario de materias y técnicas. (I. Materias)*. Secretaría General Técnica. Centro de Publicaciones. Ministerio de Cultura. Retrieved from <https://sede.educacion.gob.es/publiventa/diccionario-de-materias-y-tecnicas-i-materias/museos/13342C>

Kroustallis, S. K. (2015). *Diccionario de materias y técnicas (II. Técnicas)*. Secretaría General Técnica. Centro de Publicaciones. Ministerio de Educación, Cultura y Deporte. Retrieved from <https://sede.educacion.gob.es/publiventa/diccionario-de-materias-y-tecnicas-ii-tecnicas/diccionarios-enciclopedias-y-tesauros-museos/20692C>

Library of Congress. (n.d.). *Thesaurus for Graphic Materials II: Genre and Physical Characteristic Terms (TGM II)*. Retrieved from <http://www.loc.gov/rr/print/tgm2/>

Ministère de la culture et de la communication (France). (n.d.). *Vocabulaires scientifiques - Joconde*. Retrieved from <http://www.culture.gouv.fr/documentation/joconde/fr/partenaires/AIDEMUSEES/vocabulaires.htm>

Ministère de la Culture et de la Communication (France). (2000). *Thésaurus de la désignation*. Retrieved from <http://www.culture.gouv.fr/culture/inventai/patrimoine/index.htm>

Ministerio de Educación, Cultura y Deporte. (n.d.). *Tesauros - Diccionarios del patrimonio cultural de España - Portada*. Retrieved from <http://tesauros.mecd.es/tesauros>

Bourcier, P., Dunn, H., Chenhall, R. G., & Nomenclature Taskforce. (2015). *Nomenclature 4.0 for Museum Cataloging: Robert G. Chenhall's System for Classifying Cultural Objects* (4^a). American Association for State and Local History. Retrieved from <http://resource.aaslh.org/view/nomenclature-4-0-for-museum-cataloging/>

Padilla, C., Maicas, R., & Cabrera, P. (2002). *Diccionario de materiales cerámicos*. Secretaría General Técnica. Centro de Publicaciones. Ministerio de Educación, Cultura y Deporte. Retrieved from <https://sede.educacion.gob.es/publiventa/diccionario-de-materiales-ceramicos/artesania-artes-populares/10968C>

Projecto CLIP - Compatibilização de Linguagens de Indexação em Português. Grupo de Belas Artes, Neves, A. L. P., Sarmiento, A. de M., & Santos, M. L. (1996). *Terminologia controlada para indexação de documentos na área da arquitectura militar*. Lisboa: Inst. da Biblioteca Nacional e do Livro. Retrieved from <http://purl.pt/250>

Projecto CLIP - Compatibilização de Linguagens de Indexação em Português. Grupo de Belas Artes, Resende, J. M., Ventura, J., & Duarte, E. (1996). *Terminologia controlada para a indexação de documentos na área do design*. Lisboa: Inst. da Biblioteca Nacional e do Livro. Retrieved from <http://purl.pt/251>

Projecto CLIP - Compatibilização de Linguagens de Indexação em Português. Grupo de Belas Artes, Sarmiento, A. de M., & Santos, M. L. (2012). *Terminologia Controlada para Indexação de Obras da Área da Fotografia* (3a ed.). Lisboa: Instituto Português de Fotografia. Retrieved from <http://www.ipf.pt/files/terminologia+controlada+setembro+2012.pdf>

Projecto CLIP - Compatibilização de Linguagens de Indexação em Português. Grupo de Belas Artes, & Ventura, J. (1998). *Terminologia controlada para a indexação de documentos na área da arquitectura religiosa*. Lisboa: Biblioteca Nacional. Retrieved from <http://purl.pt/259>

Rocca, S. V. (Dir.), Guedes, N. C. (Coord.), Roque, M. I., & Guerreiro, D. (2004). *Thesaurus - Vocabulário de Objetos do Culto Católico*. Lisboa: Fundação da Casa de Bragança; Universidade Católica Portuguesa.

Rodríguez Bernis, S. (2005). *Diccionario de mobiliario*. Secretaría General Técnica. Centro de Publicaciones. Ministerio de Cultura. Retrieved from <https://sede.educacion.gob.es/publiventa/diccionario-de-mobiliario/artes-decorativas-conservacion-restauracion/12574C>

Rubio Liniers, M. C. (1999). *Tesouro de Historia Contemporânea de Espanha*. Madrid: CSIC - Centro de Información y Documentación Científica (CINDOC). Retrieved from <http://digital.csic.es/handle/10261/35800>

Trinidad Lafuente, I. (2011). *Tesouro y diccionario de objetos asociados a ritos, cultos y creencias*. Secretaría General Técnica. Centro de Publicaciones. Ministerio de Cultura. Retrieved from <https://sede.educacion.gob.es/publiventa/tesouro-y-diccionario-de-objetos-asociados-a-ritos-cultos-y-creencias/antropologia-museos/13865C>

Trinidad Lafuente, I. (2012). *Tesouro y diccionario de objetos asociados a la expresión artística*. Secretaría General Técnica. Centro de Publicaciones. Ministerio de Educación, Cultura y Deporte. Retrieved from <https://sede.educacion.gob.es/publiventa/tesauro-y-diccionario-de-objetos-asociados-a-la-expresion-artistica/diccionarios-enciclopedias-y-tesauros/14268C>

The Library of Congress. (2016). *Library of Congress Authorities*. Retrieved from <http://authorities.loc.gov/>

Van de Waal, H., Royal Netherlands Academy of Arts and Sciences (KNAW), & University of Utrecht. (n.d.). *Iconclass*. Retrieved from <http://www.iconclass.nl/home>



A MEDIAÇÃO DA INFORMAÇÃO NA REDE DE ARQUIVOS DIPLOMÁTICOS IBERO-AMERICANOS – RADI

Ana Cláudia Medeiros¹, Bruna Lessa², Carolina Magalhães³, Daniel Marins⁴, Hildenise Novo⁵,
Ivana Lins⁶, Luis Carlos Batista⁷, Maíra Salles⁸

¹UFBA ana.medeiros@ufba,

²UFBA, 0000-0003-4485-203X, brunalessa@ufba.ba

³UFBA, 0000-0002-2391-3393, carolina.magalhaes@ufba.br

⁴UFBA, 0000-0002-2253-7410, danielmarinss@hotmail.com

⁵UFBA, 0000-0003-0986-7164, hново@ufba.br

⁶UFBA, /0000-0003-0422-4135, Ivana.lins@gmail.com

⁷UFBA, 0000-0002-8071-3494, luisjesusb@bol.com.br

⁸UFBA, 0000-0003-4151-7690, maira.salles@ufba.br

RESUMO Este artigo se propõe a identificar a mediação da informação na Rede de Arquivos Diplomáticos Ibero-Americanos – RADI, considerando o acesso à informação e a visibilidade aos arquivos diplomáticos. Ao disponibilizar os recursos digitais é preciso garantir a sua recuperação, acesso e uso a partir de plataformas que promovam uma mediação tecnológica no ciberespaço. Os conceitos de usabilidade são aplicáveis para analisar a eficácia dessa mediação. Propõe-se, portanto, uma triangulação de conceitos, problematizando as questões que envolvem a mediação, rede de informação e tecnologia para um potencial acesso e uso no ambiente virtual. Essas redes apresentam-se como novo modelo de produção, centradas na necessidade do usuário e surgem nos campos da memória e do patrimônio. Realizou-se uma observação direta dos principais recursos disponíveis no *website* da RADI a fim de coletar informações importantes sobre as políticas de cooperação e mediação e identificar como se dá o acesso ao acervo a partir de uma abordagem centrada na usabilidade, tendo a mediação, como ação de interferência do profissional no processo de organização, busca e apropriação da informação. A RADI contribui para fortalecer o papel sociocultural dos arquivos dos ministérios das Relações Exteriores dos países ibero-americanos, porque a documentação diplomática é de suma importância para a história da independência da América Latina.

PALAVRAS-CHAVE Rede de informação, mediação da informação, acesso à informação, usabilidade, RADI

ABSTRACT This article aims to evaluate information mediation mechanisms in the Ibero-American Diplomatic Archives Network – RADI, taking into account ease of access to the information and visibility of the diplomatic archives. When making digital resources available, it is necessary to safeguard their recovery, access and use based on platforms that promote technological mediation in cyberspace. The concepts of user friendliness are applicable to analyze the efficacy of this mediation. Within this context, a triangulation of concepts is proposed, critically evaluating issues that involve mediation, information networks and technology for potential access and use in virtual environments. These networks are presented as a new production model, focused on user needs and are applicable to the fields of collective memory and heritage. Direct observations were carried out of the main resources available on the RADI website to collect important information on cooperation and mediation policies and identify how the archives are accessed based on a user friendliness-focused approach, with mediation as the professionals' tool in the process of information organization, search and appropriation. RADI contributes to strengthening the socio-cultural role of the ministries of foreign relations' archives in Ibero-American countries, as diplomatic documentation is of paramount importance to the history of independence in Latin America.

KEY WORDS *Information network, information mediation, information access, user friendliness, RADI.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

1 INTRODUÇÃO

O processo de busca e apropriação da informação se intensifica com a utilização da *internet*, que através das ferramentas e dispositivos tecnológicos redefinem o arquivo tradicional e virtual. Assim, o sentido de hibridação pode ser observado nas redes de informação entre instituições arquivísticas como tentativas mediadoras e surgem nos campos da memória e do patrimônio possibilitando a divulgação dos itens documentais arquivísticos, além de favorecer a sua preservação.

As tecnologias informacionais na contemporaneidade atuam nas diversas práticas sociais, entretanto, pesquisas relacionadas às temáticas da Tecnologia da Mediação e Tecnologias da Informação e Comunicação (TIC), têm sido de grande valia, pois os mesmos têm a preocupação em abordar a organização, armazenamento, arquitetura, recuperação da informação, satisfação do usuário, usabilidade de programas entre outros aspectos.

Diante disso, este artigo se propõe a identificar a mediação da informação na Rede de Arquivos Diplomáticos Ibero-Americanos, considerando o acesso à informação e a visibilidade aos arquivos diplomáticos. Preocupou-se com a mediação humana inserida nas redes virtuais (de informação, cooperação, colaboração), por possibilitar articulações na comunicação.

Os arquivos dos ministérios das Relações Exteriores dos países ibero-americanos salvaguardam uma documentação de suma importância para a história da independência da América Latina. Através da *internet*, o acervo disponibilizado e os instrumentos de pesquisa atingem um grande número de usuários / pesquisadores que estão interessados nas informações arquivísticas, seja para estudos científicos ou atender a uma mera curiosidade. Assim, as redes de arquivo promovem as instituições, garantem a autenticidade e integridade da informação e possibilitam o acesso à informação de forma rápida e precisa.

2 TECNOLOGIA E REDE DE INFORMAÇÃO: USABILIDADE

A discussão sobre as unidades de informação (arquivo, bibliotecas, museus e centros de documentação) e a tecnologia da informação e comunicação – TIC tem o objetivo de evidenciar os meios de produção e difusão da informação no mundo contemporâneo, em que a *internet* direciona o rumo das novas formas de armazenamento e acesso aos documentos.

A melhoria da transmissão de dados e a criação de protocolos de comunicação possibilitaram a conexão das redes locais e regionais, definida por Castells (2002) como sociedade em rede, baseada no paradigma econômico e tecnológico da informação, que aborda as práticas sociais, espaços de fluxos, mobilidades e desterritorializados.

A *web 2.0*, apresentada por O'Reilly (2005), é uma ferramenta que proporciona a comunicação de forma participativa e dinâmica, a fim de promover a troca de informação nas comunidades virtuais. Os

aplicativos devem ser elaborados a partir dos modelos de negócio abertos e flexíveis, com padrões, protocolos e sistemas operacionais *open source*, que favorecem a democratização da informação. Como exemplo, destaca-se o *facebook*, *twitter*, *youtube* (com postagem de vídeos), *instagram* e *flickr* (para fotografias e imagens), *wikis* (páginas *web* cujos conteúdos podem ser editados por múltiplos utilizadores) e *blog* (espaços para discussões e diários de campos).

A virtualização das comunidades possibilita que a informação e os serviços sejam adaptados às necessidades específicas do usuário, considerando as funções cognitivas na busca e uso da informação. Para Varela, Barbosa e Guimarães (2009), as soluções que facilitem o acesso ao conhecimento, por meio das próprias ferramentas, são identificadas como processos mediados por computador.

Nesse contexto, as redes de informação apresentam-se como novo modelo de acesso à informação, cuja estrutura conecta (hierarquicamente ou não) diversas unidades de informação sobre um mesmo objetivo, indicando caminhos para a efetivação do direito à informação arquivística.

As redes são formas de organização flexível e adaptável nos seus objetivos, normas e atores, cuja finalidade é a difusão da informação e a articulação de saberes. Apresenta-se como um sistema de nós interligado e compartilha os mesmos códigos de comunicação, formada por protagonistas (indivíduos, grupos e/ou organizações). É por isso, a comunicação e a facilidade de acesso à informação estão atreladas a inclusão digital. Para Varela *at al.* (2009), as redes favorecem a democratização do conhecimento, permitindo o diálogo entre os atores sociais, a comunidade científica e a leiga.

Do ponto de vista de Tomáel (2005, p. 3), “as redes de informação reúnem pessoas e organizações para o intercâmbio de informações, ao mesmo tempo em que contribuem para a organização de produtos e a operacionalização de serviços que sem a participação mútua, não seriam possíveis.” Portanto, neste trabalho, considera-se rede de informação entre instituições arquivísticas como a redefinição dos arquivos virtuais, pois permitem a organização, preservação e propagação dos documentos históricos.

As redes de instituições arquivísticas emergem dessa necessidade de acesso à informação, autenticidade e integridade, em que a rápida obsolescência tecnológica coloca em risco a construção do conhecimento, o patrimônio cultural e, conseqüentemente, a memória. Os documentos natos-digitais, digitalizações, áudios e vídeos, ou seja, todos os registros independentes da forma e suporte em ambiente digital precisam ser acessíveis a longo tempo.

Nessa perspectiva, o objeto digital apresenta uma dimensão física, que é o conjunto de símbolos registrados em suporte físico e interpretável pelo hardware; uma dimensão lógica, em que o código/estrutura de dados é compreensível pelos computadores; dimensão conceitual, no qual o código adquire significado para o ser humano; e dimensão essencial, com a autenticidade e ciclo de vida da informação (Pinto, 2014).

A cooperação entre instituições arquivísticas estabelece o compartilhamento dos registros fidedignos que refletem as ações institucionais, servindo como fonte para a construção do conhecimento e a memória. Nesse contexto, a mediação visa ampliar as relações entre os objetos digitais e usuários, contribuindo com a formação individual e coletiva na área cultural, histórica e científica.

A mediação da informação mostra-se como temática relevante nesse contexto, pois é a mesma que orienta a implantação das redes, bem como determina as formas de disponibilizar descrições arquivísticas e instrumentos de pesquisas (guias, inventários), resumos e imagem do documento. Assim, a dimensão social da rede é de relevância e associa-se com o contexto e valor da informação.

3 ESPAÇOS DE PRESERVAÇÃO DIGITAL: MEDIAÇÃO PARA ACESSO E USO DA INFORMAÇÃO

A criação de espaços de preservação digital, tais como bibliotecas e repositórios digitais, tem se tornado cada vez mais comum entre instituições públicas e privadas. Segundo Santaella (2010) as tecnologias do acesso se relacionam ao advento da *internet*, que possibilitou a convergência dos computadores com as telecomunicações, permitindo o acesso a qualquer hora e em qualquer lugar conferindo uma autonomia aos usuários que utilizam a rede.

Por esse aspecto, a mediação para o uso e acesso da informação nos espaços de preservação digital pode se caracterizar como intrínseca, pois está relacionada ao processo que envolve a organização e indexação de conteúdos, a escolha dos formatos para disseminação dos documentos, o planejamento para o uso e preservação dos dados, a estrutura do ambiente virtual para armazenamento, interface, entre outros, ou seja, ações relacionadas às atividades técnicas do profissional da informação.

Para Almeida Júnior (2015) a mediação acontece em ambientes informacionais de forma direta ou indireta, o que nos faz inferir que tais ambientes também podem ser virtuais. Nesse sentido, embora não haja uma mediação direta, onde há o contato direto entre o usuário e o profissional da informação, no acesso às coleções nos espaços de preservação digital, é a sua interface que atua como mediador na busca por informação.

Nesse sentido, é possível falar em mediação tecnológica quando o uso de meios artificiais (Vygotsky, 1993), as interações no ciberespaço (Pierre Levy, 1998), nos espaços de fluxos (Manuel Castells, 1999), na infoesfera (Luciano Floridi, 2009), é feita pelos *designers* de sistemas interativos (desenvolvedores de páginas da *web*), que potencializam o acesso a fontes documentais de diferentes culturas de forma global. Diante disso, a interface em espaços de preservação digital e transmissão do conhecimento no espaço virtual caracterizam-se como instrumentos de mediação entre os homens e os objetos (reais ou virtuais), pois assumem uma função sociocultural no processo de compartilhamento de saberes e deve ser projetada de modo a permitir a interatividade entre a informação e o usuário, para que assim, haja satisfação da sua necessidade informacional.

Partindo dessa lógica, tais espaços precisam ter usabilidade, termo definido por Nielsen e Loranger (2007, p. xvi) como:

[...] um atributo de qualidade relacionado à **facilidade do uso de algo**. Mais especificamente, refere-se à **rapidez com que os usuários podem aprender a usar alguma coisa, a eficiência deles ao usá-la**, o quanto lembram daquilo, seu grau de propensão a erros e o quanto gostam de utilizá-la. Se as pessoas não puderem ou não utilizarem um recurso, ele pode muito bem não existir. (destaque nosso).

Conforme definição dos autores entende-se que os espaços de preservação digital devem possibilitar a autonomia dos usuários na busca da informação. Por esse aspecto, arrisca-se construir uma aproximação entre a ação de mediar e os conceitos de usabilidade descritos acima – ao utilizar um espaço de preservação digital (a exemplo de um repositório ou biblioteca digital), o usuário procura a partir da interação com a interface do *website* determinado conteúdo, o qual está organizado sob uma demanda previamente identificada por um profissional (mediação indireta), estabelecendo-se uma intermediação na interatividade do sistema e o usuário. Caso a necessidade informacional do usuário não seja satisfeita

nessa primeira busca, autônoma, a figura de um agente mediador, humano, mesmo que utilizando ferramentas tecnológicas é imprescindível, realizando com isso uma mediação direta.

4 REDE DE ARQUIVOS DIPLOMÁTICOS IBERO-AMERICANOS - RADI: ACESSO AO PATRIMÔNIO CULTURAL

A Rede de Arquivos Diplomáticos Ibero-Americanos, proposta pelo Chile, México, Uruguai e Venezuela, em 1997, foi aprovada como programa na VIII Cimeira Ibero-Americana de Chefes de Estado e de Governo, em Porto - Portugal, em 1998, com o objetivo de promover a cooperação na organização, gestão, conservação e utilização de sistemas dos arquivos dos Ministérios das Relações Exteriores da América Latina e estimular a investigação sobre as relações diplomáticas dos países membros.

O trabalho colaborativo, de intercâmbio de informações e experiências, contribui para a implementação de projetos que valorizam as instituições arquivísticas e a formação especializada em documentos diplomáticos. Dentre as linhas de ação, destaca-se a realização de reuniões anuais, para informar sobre os progressos do programa e partilhar experiências, diagnóstico de avaliação das condições e necessidades específicas de cada membro, incentivo aos países participantes a fazer contribuições financeiras para a manutenção da Rede, financiamento de projetos de preservação e digitalização de documentos e formação de arquivistas.

A Rede é coordenada por uma administração central (Secretaria Executiva), eleita por Assembléia Geral. Para o triênio 2015-2018, a Sede de la Secretaria General Iberoamerica (Madrid, Espanha) assume o gerenciamento, em conjunto com a Oficina Subregional de México e Montevideo (Uruguai).

Segundo o *Informe Anual* de 2013 da RADI, a página da *web* atingiu o total de 3.960 visitantes e, na Aula Virtual tem 595 utilizadores cadastrados, que fazem os cursos destinados a descrição, tipologia documental e documentos eletrônicos. Desta forma, a divulgação da Rede acontece principalmente nos eventos científicos da área de Arquivologia (congressos, encontros e simpósios), motivando o seu uso para lazer, pesquisa individual – curiosidade, pesquisa escolar e/ou acadêmica, objetivos pessoais (turismo, visitas guiadas, atividades culturais).

5 METODOLOGIA

Por se tratar de um estudo descritivo, utiliza-se a observação direta da Rede de Arquivos Diplomáticos Ibero-Americanos, que foi realizada pelos autores desta pesquisa com base em uma das abordagens apresentadas por Saracevic (2005) sobre avaliação de unidades de informação digitais. Tal abordagem é centrada na usabilidade, a qual envolve avaliação de diferentes características, particularmente em relação a portais, pelos usuários. Nesse sentido, as características escolhidas nesta pesquisa estão relacionadas à interface do portal, a partir de atributos de usabilidade definidos por Nielsen (2012):

- a) aprendizado: a interface possibilita ao usuário realizar tarefas básicas de forma autônoma desde o primeiro uso;
- b) eficiência: o portal é agradável de utilizar, facilitando a velocidade nas ações dos usuários.

- c) memorização: a estrutura para o acesso às informações do portal são fáceis de lembrar;
- d) satisfação: o design possibilita maior interação com o usuário.

5.1 OBSERVAÇÃO DIRETA DA RADI: INTERFACE E USO

A análise de dados consistiu na descrição da rede de arquivo histórico selecionada, observando estrutura, conteúdo, tipos de mediação da informação que envolve o acesso à informação e a usabilidade. Esses elementos contribuem na visibilidade da RADI e na preocupação em atender a necessidade informacional do usuário.

Desta forma, o processo considerou as técnicas e princípios da arquivística adotadas na rede, a inserção da rede dentro das redes sociais, tipos de busca (simples e avançada), interatividade da interface para o usuário e formato dos documentos digitais.

A Rede de Arquivos Diplomáticos Ibero-Americanos é composta por 22 (vinte e dois) países membros: Andorra, Argentina, Brasil, Bolívia, Costa Rica, Colômbia, Cuba, Chile, Equador, El Salvador, Espanha, Guatemala, Honduras, México, Nicarágua, Panamá, Paraguai, Peru, Portugal, República Dominicana, Uruguai e Venezuela. Todos contam com a participação exclusiva de arquivos públicos, ou seja, importantes para a memória nacional.

A *Página principal da Rede* é clara, com um layout limpo, disponível no idioma espanhol, português e inglês, com notícias sobre os eventos e atividades realizadas pela Rede. Expõem vídeos dos eventos publicados no *YouTube*, *link* para o *Facebook*, e ainda *hiperlinks* de apresentação e dos arquivos diplomáticos.

Com referência aos *Elementos sobre busca e recuperação da informação* observa-se que as buscas podem ser por tipo de publicações, assuntos, países, e ano de publicações. Os países participantes colaboram com bancos de dados dos arquivos nacionais, documentos digitalizados (correspondências, fotografias, relatórios) e instrumentos de pesquisas (catálogos, inventários, guias). Ao selecionar o “Brasil”, por exemplo, é direcionada para áreas de interesse, que remetem as páginas *web* da Presidência do Brasil e Ministério das Relações Exteriores, permitindo o acesso aos catálogos dos Documentos sobre o Brasil Colônia nos Arquivos de Goa (Índia) e aos documentos digitalizados sobre Relatórios Ministeriais do Império, de 1830 a 1960, entre outros.

Portanto, a Rede apresenta descrições dos acervos, documentos digitalizados e instrumentos de pesquisa. Os arquivos diplomáticos representam a memória das relações internacionais, fonte de conhecimento de políticas externas, das lutas pela independência e processos de criação dos Estados-Nação da América Latina.

Com base na estrutura, estão disponíveis o *Documento de Formulacion e Bases de Operación, Proyectos de Financiamientos* (2010-2012), Boletins das Reuniões anuais, Relatórios das atividades da RADI (2004-2014), Programa da RADI e Ficha descritiva do Projeto. Documentos que comprovem o processo de implantação da Rede, destacando a cooperação e compartilhamento de informações e experiências.

A *Comunicação com o usuário* ocorre por endereço, telefone, *e-mail*, formulário de contato, *blog*, *Facebook*, *YouTube*, e cadastramento no *site*. Interessante comentar que a visibilidade da Rede no *Facebook* está associada à capacitação de qualidade para os profissionais de arquivos diplomáticos.

A apresentação visual do *site* atrai o usuário, bem como as informações sobre a rede, como tipos de membros e objetivos. Os itens como logomarca, apresentação / histórico da rede e buscador interno mostram-se elementos essenciais, referindo-se ao ‘que’, ‘para que’ e ‘como’ da proposta de implantação das redes.

É interessante comentar que a Rede selecionada, apesar de ser descentralizada na sua atuação, em que cada membro tem autonomia para gerir conteúdo, possui uma administração central, que visa controlar a participação dos atores (membros), gerenciando os recursos financeiros, humanos e materiais. Existe a orientação sobre as normas e políticas da Rede, padronização dos produtos e os aspectos de estruturação, como protocolos de compatibilidade, formatos de interoperabilidade e tipos de *softwares*. Itens de preservação que permitem o acesso ao objeto digital.

Isso confirma que a centralidade é medida pela indicação dos laços e tipos de relacionamentos do ator central com o grupo. Por isso, as informações sobre histórico são divulgadas nos próprios sítios e podem ser localizadas instituições administradoras, que permitem o acesso aos Estatutos, Regimentos, decretos, portarias, Atas, relatórios de atividades, planejamentos, balancetes, entre outros.

O usuário precisa de auxílio para atender a sua necessidade informacional e a RADI não tem manual de orientação para quem “se sente perdido”, dificultando o processo de busca e recuperação da informação.

Considerando as características descritas por Nielsen (2012), quanto a usabilidade de portais, a RADI viabiliza o aprendizado, eficiência, memorização e satisfação, uma vez que, apesar de não ter o manual de orientação, sua interface possibilita ao usuário realizar tarefas básicas, além de promover a difusão do conhecimento e preservação da memória, através da salvaguarda da informação arquivística, e ainda propiciar a qualificação de profissionais arquivistas.

6 CONSIDERAÇÕES FINAIS

A Rede de Arquivos Diplomáticos Ibero-americanos representa a mediação através da máquina e da mediação humana, tendo o arquivista, como responsável pelo fazer profissional. A cooperação e/ou colaboração entre as instituições arquivísticas em um espaço desterritorializado, dispondo da descrição do acervo arquivístico e acesso *on-line* as fontes primárias, apresenta-se como novo modelo de visibilidade institucional e promoção ao acesso à informação.

A RADI é estruturada formalmente, consolidada por documentos legais, administração centralizada e compartilhamento de informação e experiências. Por garantir um nível suficiente de qualidade para o usuário final, possui políticas de adesão de membros, de disponibilização de conteúdo e de acesso e uso da informação.

Os elementos sobre o *layout* (desenho *web*) são essenciais para conquistar o utilizador. Nas observações foram indicados como fácil navegação, arquitetura fixa (elementos não mudam com frequência), seções visíveis, *links* secundários funcionando, URL identificado, títulos descritivos na barra do navegador, carregamento da página rápido.

A questão da usabilidade na rede de informação tornar o usuário em um sujeito proativo do sistema de acesso e uso da informação, tendo a autonomia e a liberdade de acesso que favorecem os *feedbacks* e formas de interação presentes da *web* social.

É importante considerar que as redes de arquivos têm contribuído não só com a preservação do documento, mas também com a questão do acesso físico, que minimiza as barreiras geográficas, atua em consonância com a velocidade da produção da informação, além de garantir a autenticidade e integridade dos documentos. Com isso, a consolidação da RADI faz com que as tecnologias estejam inseridas no cotidiano dos arquivos físicos e digitais. Portanto, a temática Tecnologia da Informação, Mediação e Estudo de usuário, estão bem próximas, pois se preocupam com o acesso, uso, necessidade e satisfação dos usuários.

REFERÊNCIAS BIBLIOGRÁFICAS

Almeida Júnior, Oswaldo Francisco de (2015). Mediação da informação: um conceito atualizado. In: Bortolin, Sueli, Santos Neto, João Arlindo dos, Silva & Rovilson José da (Org.). *Mediação oral da informação e da leitura*. Londrina: ABECIN, 2015. p. 9-32.

Associação Brasileira de Normas Técnicas (2002). NBR 9241-11: requisitos ergonômicos para trabalho de escritório com computador – Parte 11 – orientações sobre usabilidade. Rio de Janeiro: ABNT, 2002.

Castells, Manuel (2002) *A Sociedade em rede*. (Majer, R. V. Trad.). 6. ed. São Paulo: Paz e Terra.

Floridi, Luciano (2009). *Infosfera: ética e filosofia nell'età dell'informazione*. Editora: Giappichelli

Levy, Pierre. *Cibercultura*. (2008). 7ª reimpressão. São Paulo: Editora 34

Nielsen, Jakob (2016). *Usability 101: introduction to usability*. 2012. Recuperado em 10 de abril de 2017 de <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>

Nielsen, Jakob; Loranger, Hoa. (2007). *Usabilidade na Web: projetando websites com qualidade*. Rio de Janeiro: Elsevier.

O'Reilly, Tim (2015). *Web 2.0 Compact Definition?* 2005. Recuperado em 10 de março de 2015 de <http://radar.oreilly.com/2005/10/web-20-compact-definition.html>

Pinto, Maria Manuela (2014). Da Preservação de Documento à preservação da Informação. In: DUARTE, Zeny (Org.). *A conservação e a restauração de documentos na era pós-custodial*. Salvador: EDUFBA, p.127-196.

Purday, Jon (2016). Think culture: Europeana.eu from concept to construction. *The Electronic Library*, v. 27, n. 6, p.919-937, 2009. Recuperado em 12 de abril de 2017 de <http://www.emeraldinsight.com/doi/abs/10.1108/02640470911004039>

RADI (2014). Informe Anual de 2013 [online]. 2014. Recuperado em 03 de abril de 2015 de http://segib.org/sites/default/files/INFORME_RADI_2013.pdf

Santaella, Lúcia (2010). A aprendizagem ubíqua substitui a educação formal? *Revista de Computação e Tecnologia*, São Paulo, v. 2, n.1, p. 17-22, 2010. Recuperado em 25 de março de 2016 de <http://revistas.pucsp.br/index.php/ReCET/article/view/3852>

Saracevic, Tefko (2005). How were digital libraries evaluated? In: LIDA CONFERENCE LIBRARIES IN THE DIGITAL AGE, 2005, Dubrovnik, Croatia. *Anais..* Dubrovnik, Croatia: LIDA, 2005. Recuperado em 22 de março de 2016 de https://comminfo.rutgers.edu/~tefko/DL_evaluation_LIDA.pdf

Tomaél, Maria Inês (2005). Redes de informação: o ponto de contato dos serviços e unidades de informação no Brasil. *Informação & Informação*, Londrina, v.10, n.1/2, p.1-26, jan. / dez.. Recuperado em 13 de dezembro de 2014 de <http://www.uel.br/revistas/uel/index.php/informacao/article/view/1611>

Varela, Aida Varela, Barbosa, Marilene Lobo Abreu, Guimarães & Igor Baraúna (2009). Dos processos analógicos às tecnologias digitais contemporâneas de recuperação da informação: caminhos cognitivos na mediação para o acesso ao conhecimento. In: Sayão, Luis et al. (Org.). *Implantação e gestão de repositórios institucionais: políticas, memória, livre acesso e preservação*. Salvador: EDUFBA. p.123-161.

Vygotsky, L. S. (1993) *Pensamento e linguagem*. São Paulo: M. Fontes.



METODOLOGIAS DE SALVAGUARDA DOS BENS CULTURAIS: CONTRIBUIÇÕES DA CIÊNCIA DA INFORMAÇÃO NO CAMPO PATRIMONIAL

Vitória Gomes Almeida¹, Izabel França de Lima², Henry Poncio Cruz de Oliveira³

¹Universidade Federal da Paraíba, 0000-0003-2663-4936, vitoriagomesalmeida@yahoo.com.br

²Universidade Federal da Paraíba, 0000-0003-2701-5432, belbib@gmail.com

³Universidade Federal da Paraíba, 0000-0003-2330-2442, henry.poncio@gmail.com

RESUMO Reflete criticamente sobre a metodologia e instrumento de salvaguarda para bens culturais intangíveis, desenvolvida pelo Instituto do Patrimônio Histórico e Artístico Nacional (IPHAN): o Inventário Nacional de Referências Culturais (INRC) e o Registro de Bens Culturais de Natureza Imaterial. Objetiva compreender de que maneira se estruturam e funcionam essas metodologias, e verificar as correlações e mútuas contribuições existentes entre patrimônio e Ciência da Informação (CI). Por meio de análise documental, dos manuais e documentos técnicos do IPHAN, bem como da literatura especializada da Ciência da Informação, observa tensões, limitações e dimensões sociais da representação da informação para identificação e registro da informação patrimonial, verificando a contribuição da CI para preservação da memória cultural.

PALAVRAS-CHAVE *Informação patrimonial, Memória, Patrimônio cultural, Metodologias de Salvaguarda.*

ABSTRACT It reflects critically about the methodology and instrument for safeguarding intangible cultural assets, developed by the National Historical and Artistic Heritage Institute (IPHAN): the National Inventory of Cultural References (INRC) and the Register of Cultural Goods of Intangible Nature. It aims to understand how these methodologies are structured and how they function, and to verify the correlations and mutual contributions between patrimony and Information Science (CI). Through documentary analysis of the manuals and technical documents of IPHAN, as well as the specialized literature of Information Science, it observes tensions, limitations and social dimensions of information representation for identification and registration of patrimonial information, verifying the contribution of CI to preserve cultural memory.

KEY WORDS *Heritage information, Memory, Cultural heritage, Safeguarding methodologies.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

São consenso entre os pesquisadores, gestores culturais, mestres da cultura e demais envolvidos no campo patrimonial as percepções acerca dos avanços existentes após a criação das políticas públicas de proteção ao patrimônio cultural brasileiro, sobretudo, quando referentes ao patrimônio imaterial.

Constituído, de acordo com o Instituto do Patrimônio Histórico e Artístico Nacional (IPHAN), por bens culturais “que se manifestam em saberes, ofícios e modos de fazer; celebrações; formas de expressão

cênicas, plásticas, musicais ou lúdicas; e nos lugares (como mercados, feiras e santuários que abrigam práticas culturais coletivas)”, teve seu reconhecimento institucional através do Decreto 3.551/2000, fazendo do Brasil pioneiro em políticas patrimoniais.

Hoje, a compreensão de patrimônio cultural é vinculada à noção de referência cultural, reconhecendo a pluralidade e diversidade da produção humana, manifesta de forma tangível e intangível (FONSECA, 2000), produzida e mantida pelos mais heterogêneos grupos que compõem a sociedade brasileira, e que representam diferentes sentidos e valores sociais e culturais.

Em meio à diversidade cultural existente no contexto brasileiro, a criação de metodologias e estratégias capazes de identificar a convergência identitária e memorialística dos grupos, expressas em bens culturais de natureza intangível se fez urgente e culminou em duas ferramentas metodológicas: o Registro de Bens Culturais de Natureza Imaterial e o Inventário Nacional de Referências Culturais (INRC).

Diante desses processos que envolvem a discussão de memória, identidade, preservação e registro de informação, propomos a discussão de patrimônio cultural por meio da Ciência da Informação (CI), por meio da reflexão sobre ferramentas metodológicas para salvaguarda dos bens culturais do IPHAN.

Buscamos compreender de que maneira se estruturam e funcionam essas metodologias, verificando as correlações e as possibilidades de contribuições existentes entre o campo dos estudos informacionais e o dos estudos patrimoniais, questionando de que maneira os pressupostos teórico-metodológicos da Ciência da Informação podem contribuir no campo patrimonial, uma vez que essas metodologias são pautadas, sobretudo, na produção de documentação sobre os bens culturais alvos de salvaguarda.

As informações e reflexões obtidas seguem estruturadas em seções. Inicialmente, faz-se levantamento historiográfico da ação institucional brasileira para com o patrimônio cultural do país, em que se aplica análise documental e se discute acerca da estrutura das metodologias do IPHAN. Por fim, buscam-se na literatura especializada da Ciência da Informação (CI) perspectivas teóricas dos processos de representação do patrimônio, do acesso/uso de documentos e sua possível contribuição para a preservação da memória cultural.

PATRIMÔNIO CULTURAL E AÇÃO INSTITUCIONAL NO BRASIL

No ano de 1937 acontece a criação do Serviço do Patrimônio Histórico e Artístico Nacional (SPHAN)¹, primeira instituição do Brasil para proteção do patrimônio cultural do país. Enquanto entidade oficial mais antiga de proteção de bens culturais na América Latina, é responsável pela promoção permanente do tombamento, da conservação, do enriquecimento e conhecimento do patrimônio artístico e histórico nacional.

Contudo, é preciso ressaltar que essa atuação voltada quase exclusivamente para os bens móveis, priorizando o chamado patrimônio de pedra e cal, difere da proposição feita por Mário de Andrade,

¹ Hoje renomeado para Instituto do Patrimônio Histórico e Artístico Nacional (IPHAN).

autor do anteprojeto do SPHAN, que, através de suas viagens, pesquisas e olhar etnográfico, já demonstrava no início do século XX atenção com a cultura imaterial ou intangível.

A grande atuação da instituição em prol da preservação e conservação voltou-se, durante muito tempo, exclusivamente para os bens móveis, expresso por meio do seu instrumento de preservação o *tombamento* e a divisão dos bens a serem preservados (Livro do Tombo Arqueológico, Etnográfico e Paisagístico; Livro do Tombo Histórico; Livro do Tombo de Belas-Artes; Livro do Tombo das Artes Aplicadas).

São consequências do contexto político em que é criada e implantada: Movimento Modernista (traz em pauta perspectiva etnográfica, que pode ser percebida nas proposições de Mário de Andrade), Revolução de 1930 (levou à constituição do Estado brasileiro enquanto Estado nacional, capitalista e burguês); Estado Novo (regime ditatorial que utilizou cultura e políticas culturais para fortalecer o Brasil enquanto Estado Nacional). Castriota (2009, p. 74) lembra que "as primeiras ações do SPHAN (criado em 1937) derivam da 'narrativa ortodoxa': estabelece-se afinidade eletiva entre nosso passado barroco e colonial e a arquitetura moderna que então se fazia".

Ao ser preterido pelos intelectuais que se articularam em torno do projeto que criou o Sphan, o caráter etnográfico de sua orientação teórica se revelou abrangente demais para a representação, a partir do patrimônio histórico e artístico nacional, da nação em processo de construção. A proposta totalizante de cultura de Mário de Andrade encontrou resistência entre os vários grupos que lidavam pragmaticamente com o chamado patrimônio cultural dedicando-se a práticas de colecionar, restaurar e preservar objetos com o propósito de colocá-los à mostra segundo as funções didáticas ou políticas que lhes eram atribuídas (NOGUEIRA, 2005, p. 220).

O intuito da política da época era então desenvolver “alta cultura” no país, em que os valores da nação que se construía deviam estar em primeiro lugar. Manifestações e expressões populares como as oriundas do folclore, além de serem tidas como cultura “menor”, viviam sob rígida política de vigilância e controle.

As ações institucionais desenvolvidas para cultura imaterial ou intangível eram concebidas dentro da noção de cultura popular ou folclore, onde estavam implícitas sua subalternidade (diante da chamada cultura erudita, por exemplo) e valorização da tradição como vestígio do passado, considerando todo progresso ou mudança como processo de dessacralização da sabedoria popular (ORTIZ, 2006).

O resultado é a difusão de noção enrijecida no campo do patrimônio e conseqüentemente da memória, em que tradição e cultura (sobretudo a denominada popular) deveriam se manter puras e autênticas, sendo rejeitadas quando essas apresentavam novas dinâmicas e configurações.

Somente com a promulgação da Constituição de 1988 uma nova e mais abrangente compreensão de patrimônio cultural passa a ser adotada, na qual se considera que sejam constituídos por bens tangíveis e intangíveis, e que incluem as formas de expressão: modos de criar, fazer e viver; criações científicas, artísticas e tecnológicas; obras, objetos, documentos, edificações e demais espaços destinados às manifestações artístico-culturais; conjuntos urbanos e sítios de valor histórico, paisagístico, artístico, arqueológico, paleontológico, ecológico e científico.

Nove anos mais tarde, os debates acerca desses bens intangíveis, já entendidos como integrantes do patrimônio cultural, culminam no Seminário Patrimônio Imaterial: estratégias e formas de proteção,

realizado em Fortaleza, com o intuito de discutir os instrumentos legais e administrativos de preservação dos bens culturais intangíveis, resultando na Carta de Fortaleza.

Como decorrência dessas ações surge o Decreto 3551/2000, que reconhece institucionalmente o patrimônio cultural imaterial brasileiro, constituído, de acordo com o IPHAN, por bens culturais “que se manifestam em saberes, ofícios e modos de fazer; celebrações; formas de expressão cênicas, plásticas, musicais ou lúdicas; e nos lugares (como mercados, feiras e santuários que abrigam práticas culturais coletivas)”. Com isso, prioriza-se a construção e legitimação de um patrimônio que constate sua singularidade e relevância para a memória e identidade dos brasileiros.

Esse reconhecimento institucional em âmbito nacional surge no período em que a crítica ao patrimônio tangível, elitizado e eurocêntrico também era desenvolvida internacionalmente. Organizações como a Organização das Nações Unidas para a Educação, Ciência e Cultura (UNESCO) lançam, ainda em 1989², a noção de patrimônio cultural imaterial como fonte de diversidade cultural e garantia de desenvolvimento sustentável, reiterada em alguns documentos internacionais, como a Declaração Universal da UNESCO sobre a Diversidade Cultural, de 2001, e na Declaração de Istambul, de 2002.

Somente na Convenção de 2003 cria-se, efetivamente, instrumento multilateral com propósito de salvaguardar patrimônio cultural imaterial, criado a partir da perspectiva da alteridade, que considera constantes “recriações” aos quais esses bens são alvos, tendo em vista a dinâmica que se insere em suas comunidades e grupos (PELEGRINI; FUNARI, 2008).

O reconhecimento das expressões da cultura como algo dinâmico, móvel e passível de novas reconfigurações exige adoção de novos instrumentos e metodologias que consigam atender especificidades que bens culturais intangíveis demandam em sua preservação e salvaguarda.

Nesse sentido, enquanto diretrizes básicas para salvaguarda dos bens culturais intangíveis, elenca-se produção de inventários e registros, cuja base está em produzir documentação e conhecimento acerca do bem cultural em processo de reconhecimento, enquanto patrimônio cultural brasileiro, abordando seu contexto (questões sociais, territoriais, materiais e ambientais), tornando possíveis sua manutenção e reprodução.

Segundo o IPHAN (2017), os instrumentos utilizados para a salvaguarda dos bens culturais intangíveis foram construídos e aplicados levando em consideração sua natureza processual e dinâmica.

Com relação à INRC, essa se constitui enquanto metodologia de pesquisa ou procedimento de investigação em que o objetivo principal é produzir conhecimento sobre os marcos e referências de identidade de determinado grupo social.

tal metodologia possibilita: 1) auxílio, quando for o caso, à instrução dos processos de registro; 2) promoção do patrimônio cultural imaterial junto à sociedade; 3) orientação para ações de apoio e fomento a bens culturais em situação de risco ou de atendimento a demandas advindas do processo de inventário; 4) tratamento e acesso público às informações produzidas sobre esse universo (CAVALCANTI; FONSECA, 2008, p. 20).

² Recomendação para a Salvaguarda da Cultura Tradicional e Popular.

Dividida em três etapas (levantamento preliminar, identificação e documentação), é composta por atividades de planejamento da pesquisa, elaboração de plano de trabalho, formação das equipes e capacitação, preparação de metodologia e instrumentos de pesquisa, pesquisa em campo com aplicação dos instrumentos de pesquisa (realização de entrevistas, reuniões, questionários), além da documentação audiovisual. Na etapa final, é realizada sistematização das informações, elaboração das publicações, promoção e divulgação dos resultados.

Com relação ao registro, define-se este como instrumento legal de preservação, reconhecimento e valorização do patrimônio intangível, que requer pesquisa documental e de campo, mobilização e consenso social sobre motivações e propósitos; argumentação sobre o valor patrimonial da expressão cultural; diagnóstico sobre vulnerabilidade e recomendações para salvaguarda do bem cultural, finalizando com inscrição do bem em um dos livros do Patrimônio Imaterial (Celebrações, Lugares, Saberes ou Formas de Expressão).

As propostas de registros definem-se no movimento coletivo da própria sociedade. Recebidas pelo IPHAN e avaliadas em caráter preliminar, se julgadas procedentes, são encaminhadas para instrução. A instrução dos processos de registro – a elaboração dos dossiês de registros – é sempre supervisionada pelo IPHAN. Consta de descrição pormenorizada do bem a ser registrado, acompanhada da documentação correspondente. [...] Realizada a instrução do processo, o IPHAN emite parecer publicado no Diário Oficial da União. Após 30 dias, que acolhem eventuais manifestações da sociedade sobre o registro, o processo é encaminhado ao Conselho Consultivo do Patrimônio Cultural para deliberação. [...] Um critério-chave para a legitimidade de qualquer pleito ao registro é a sua relevância para a memória, a identidade e a formação da sociedade brasileira (CAVALCANTI; FONSECA, 2008, p. 20).

O foco do registro está em documentar, com técnicas específicas, a continuidade histórica da manifestação, buscando captar diversas características que o bem cultural apresenta ao longo do tempo, possibilitando avaliação de suas modificações e tornando acessível ao público as informações coletadas e registradas.

Conhecendo procedimentos adotados no registro e no INRC, fazemos algumas observações:

- Em ambos prevalece noção de referência cultural, concepção em que sentidos e valores atribuídos aos bens culturais são determinados pelos grupos e comunidades;
- O INRC tem, por finalidade, identificar e mapear bens tangíveis e intangíveis constituídos enquanto referência cultural, uma vez que o Registro proporciona o reconhecimento de determinado bem cultural como patrimônio cultural imaterial brasileiro;
- Essas ações têm, como principal produto, a criação e sistematização de grande quantidade de informação patrimonial, que contribui diretamente para o conhecimento das especificidades dos bens para o planejamento de ações de preservação e salvaguarda;
- Apesar da grande quantidade de informação patrimonial produzida e do objetivo de disponibilizá-las para acesso ao público, através do Banco de Dados de Bens Culturais Imateriais Registrados, criado por exigência do Decreto 3551/2000, o número de acessos registrados até a última vez que visitamos o site era de quatro mil, como pode ser observado na Figura 1 abaixo:



Figura 1. Banco de Dados de Bens Culturais Imateriais Registrados

A crítica de nossa análise no banco de dados recai sobre dois aspectos: o número de acessos e sua arquitetura da informação. No primeiro, não foi possível localizar data de criação do banco de dados no próprio site do IPHAN, porém, por meio do site MatrizPCI³ verificou-se que foi lançado em 10 de agosto de 2010, somando, em sete anos de atuação, apenas 4 mil acessos. Considerando os dados supracitados, o site possui uma média 1,58 acessos diários. Com relação à estrutura do site, foram identificadas dificuldades relacionadas a organização das informações no banco de dados e localização dos itens, também foram percebidas dificuldades relacionadas a estrutura de navegação em camadas de páginas. Também identificamos problemas para recuperação de informações por meio do sistema de busca. O que ora expomos extrapola os objetivos do presente texto, mas indica a necessidade de estudos posteriores com foco na Arquitetura da Informação do referido ambiente.

Diante dessas percepções, permanece o questionamento: de que maneira os pressupostos teórico-metodológicos da Ciência da Informação podem contribuir no campo patrimonial, para coleta, registro, disseminação e acesso da informação patrimonial? Para tanto, trataremos na próxima seção algumas reflexões.

INFORMAÇÃO E PATRIMÔNIO: NOTAS NO ÂMBITO DA CI

Os debates acerca da informação, memória, cultura e patrimônio encontram-se consolidados no campo de CI, através da sua institucionalização nos grupos de trabalho da associação de pesquisa e nos programas de Pós-Graduação da área, também como reflexo pelo grande número de produções teóricas sobre suas interações, discursos, representações e fluxos.

Em suas reflexões sobre esses temas, Azevedo Netto (2008) afirma que, nos processos de preservação cultural, é demandada uma série de questionamentos (quais, quem e para quem os patrimônios são preservados?), na qual o principal instrumento de preservação que se pode dispor é a informação, uma

³ Site sob a tutela do Governo de Portugal, com recursos em língua portuguesa para difusão de boas práticas e valorização do Patrimônio Cultural Imaterial e sistema de informação de suporte ao *Inventário Nacional do Patrimônio Cultural Imaterial*.

vez que experiências na atualidade têm mostrado que os patrimônios selecionados a serem preservados foram fruto de estudos, tendo suas informações repassadas e incorporadas pelas populações de entorno.

Considerando o objeto de estudo da CI e sua responsabilidade social oriunda de processo cultural (desenvolvimento científico da sociedade) e estabelecida através de representações que abarcam sua assimilação e reprodução (MARTINS e MORAES, 2012; DODEBEI, 2002), e pensando nas atuais metodologias do IPHAN para salvaguarda do patrimônio intangível, que têm como foco documentar informações sobre os bens culturais estudados, ao trabalharem com a coleta e registro de informação patrimonial, verificamos uma das correlações e possibilidades de contribuição da área para a preservação da memória cultural, através da organização e tratamento da informação.

Para isso, evocamos as discussões dos processos de representação realizadas no campo da CI que, segundo Alvarenga (2002), estão divididos entre representação primária, feita pelo autor, e representação secundária, realizada em unidades de informação para organização e recuperação em sistemas documentais. Entre esses processos de representação, abarcariam três momentos distintos que, segundo ela, compreenderiam a produção dos registros do conhecimento, organização dos sistemas de informação documentais e acesso às informações pelos usuários.

Ao lidar diretamente com processos de representação dos bens culturais e ter como produtos documentos (dossiês) sobre os bens culturais, essas ações institucionais de identificação, salvaguarda e socialização do patrimônio cultural colocam a informação como elemento central, tanto no que se refere aos atributos dos bens culturais, quanto sua representação no âmbito da documentação produzida sobre eles. Nesse sentido, o campo de estudos da organização e tratamento da informação, sobretudo os referentes à representação, poderia contribuir para o aprimoramento da representação da informação no âmbito do INRC e do registro, colaborando, por conseguinte, na recuperação de informações e na disseminação e apropriação da informação patrimonial.

Além disso, as reflexões no âmbito dos sistemas documentários também poderiam contribuir para a salvaguarda, acesso e uso, uma vez que se referem aos sistemas resultantes das atividades documentárias, ou seja, atribuição de significados com o fim de orientar usuários em seus processos de busca e uso de informação (ORTEGA, 2013).

METODOLOGIA

De acordo com Aquino (2013), pensar-conhecer-agir sobre a realidade social é tarefa do/a pesquisador/a. No âmbito da CI essa tarefa ganha especificidades, pois o fazer da pesquisa significa idealizar prática sociocultural de interação do pesquisador com seu objeto (informação), contextualizando-a e mobilizando seu potencial cognitivo para organizá-la, selecioná-la, interpretá-la, ressignificá-la e atribuir sentidos, tendo em vista atender demandas sociais, culturais (diversidade cultural e institucional) e pessoais (reconhecimento, aprendizado e construção de significados dos sujeitos/pesquisadores sobre a ciência).

Considerando essa tarefa, na clareza de suscitar os debates e refletir acerca do nosso questionamento, que visa compreender de que maneira os pressupostos teórico-metodológicos da Ciência da Informação podem contribuir no campo patrimonial, temos como objeto de estudo a metodologia do Inventário Nacional de Referências Culturais (INRC) e o instrumento de salvaguarda, o Registro de Bens Culturais de Natureza Imaterial.

Para alcançarmos os objetivos propostos, faremos pesquisa bibliográfica ao utilizarmos o aporte teórico da CI. Buscando maior número possível de informações e detalhamento sobre o tema, recorreremos à análise documental como técnica para identificação, reconhecimento e análise da metodologia de inventário de referências culturais, perfilhando seus sentidos sociais e institucionais, uma vez que, segundo Souza (2011), mostrou-se como técnica eficaz para subsidiar estudos que utilizam abordagem qualitativa a fim de compreender cientificamente os fenômenos sociais da atualidade.

REFLEXÕES FINAIS

A produção e sistematização de informação patrimonial são essenciais para garantia e salvaguarda de bens culturais, sobretudo quando nos referimos aos bens culturais intangíveis. Dentro do campo da CI, as discussões que remetem a essas temáticas ganham força em consonância com a preocupação em âmbito institucional (sob a forma de políticas públicas), o que nos oferece terreno propício para trazer reflexões de cunho teórico-metodológico, considerando aproximação e múltiplas intersecções existentes.

Ortega (2013), ao tratar o objeto de estudo da CI quanto à sua finalidade, coloca que essa seria contemplar as necessidades humanas de informação⁴ e defende que o objeto da Ciência da Informação relaciona-se, então, à intervenção realizada entre a produção e o uso de informação, por meio da elaboração de registros ou inscrições, e das demais atividades que possibilitam a permanência destes registros para acesso e usos posteriores.

Se pensarmos nesses termos, teremos a necessidade de informação acerca dos bens culturais, presentes no INRC e no registro do IPHAN. O objeto da CI seria a informação patrimonial, permeada pelas atividades que compõem área de organização e tratamento, uma vez que, por meio delas, seria possível ter contribuições visando o aprimoramento das representações dos bens culturais, repercutindo, por conseguinte, na recuperação, acesso, uso e apropriação das informações relativas às referências e patrimônios culturais alvos de salvaguarda.

Com isso, acreditamos que alguns dos desafios no acesso ao patrimônio cultural, como o apresentado sobre a base de dados do IPHAN sobre os bens culturais registrados, possam ser superados, garantindo que a preservação da memória e a salvaguarda do patrimônio cultural sejam cada vez mais efetivas.

REFERÊNCIAS BIBLIOGRÁFICAS

Alvarenga, L.. (2001, dezembro). A Teoria do Conceito revisitada em conexão com Ontologias e Metadados no contexto das bibliotecas tradicionais e digitais. DataGramaZero, 6. http://www.brapci.ufpr.br/brapci/_repositorio/2010/01/pdf_adf84bf8a1_0007457.pdf

⁴ “inerentes a todo ser humano, e relativas a aspectos de ordem científica, educacional, utilitária, de entretenimento, estética, profissional, outras” (ORTEGA, 2013, p. 152).

Aquino, M. A. (2013). Ciência e método: elementos para reflexão nas pesquisas em Ciência da Informação. In: Aquino, M. A., Oliveira, H. P. C., Lima, I. F. (Org). Experiências metodológicas em Ciência da Informação (pp. 19-47). João Pessoa: Editora da UFPB.

Netto, C. X. A. (2008). A abordagem do conceito como uma estrutura semiótica. *Transinformação*, 20(1), 47-58. <https://dx.doi.org/10.1590/S0103-37862008000100004>

Fonseca, M. C. L. (2001). Referências culturais: base para novas políticas de patrimônio. In: IPHAN. *Inventário Nacional de Referências Culturais: manual de aplicação* (pp. 11-21). Brasília: IPHAN.

Castriota, L. B. (2009). *Patrimônio cultural: conceitos, políticas, instrumentos*. São Paulo: Annablume.

Fonseca, M. C., & Cavalcanti, M. L. D. C. (2008). *Patrimônio Imaterial no Brasil. Legislação e Políticas Estaduais*. Brasília: Instituto Brasileiro de Educação e Cultura.

Dodebei, V. (2002). *Tesouro: linguagem de representação da memória documentária*. Intertexto.

Autor, A. A., Autor, B. B., & Autor, C. C. (Ano). *Título do trabalho*. In *Título do evento* (páginas). Cidade, Estado/País: Editora.

Martins, G. K., & Netto, C. X. A. (2011). Representação da informação e preservação da memória: mapeamento conceitual do patrimônio imaterial brasileiro. In: ISKO Brasil - Congresso Brasileiro de Representação e Organização do Conhecimento (pp. 122-136). Brasília, DF. ISKO Brasil.

Ortega, C. D. (2013). Ciência da Informação: do objetivo ao objeto. In: Rojas, M. Á. R. (Org.). *El objeto de estudio de la bibliotecología/documentación/ciencia de la información: propuestas, discusión, análisis y elementos comunes* (pp. 151-177). México: Instituto de Investigaciones Bibliotecológicas y de la Información/UNAM.

Pelegrini, S. D. C. A., & Funari, P. P. A. (2008). *O que é patrimônio cultural imaterial*. Brasiliense.



DIAGNÓSTICO DA SOBREVIVÊNCIA E DISPONIBILIDADE DOS PERIÓDICOS CIENTÍFICOS ON-LINE DE ACESSO ABERTO: ESTUDO DO ESTADO DE MINAS GERAIS (BRASIL)

Edna da Silva Angelo¹, Marlene Oliveira²

¹Universidade Federal de Minas Gerais, 0000-0002-8739-6436, ednasangelo@yahoo.com.br

²Universidade Federal de Minas Gerais, 0000-0003-2834-1272, marlene@eci.ufmg.br

RESUMO O objetivo é realizar um estudo exploratório da taxa de periódicos científicos on-line de acesso aberto inativos e indisponíveis. Para tal, utilizou-se para a investigação o estado de Minas Gerais (Brasil). Após uma minuciosa revisão bibliográfica, os dados foram levantados a partir da lista de periódicos científicos cadastrados no Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), para a solicitação do ISSN (Número Padrão Internacional para Publicação Seriada). A definição da situação da publicação de cada periódico do universo citado se baseou nas informações e indícios presentes no próprio site do periódico. Os dados coletados e os resultados foram submetidos a uma análise descritiva. Localizou-se 612 periódicos que atendiam aos critérios estabelecidos, a saber, possuir caráter de comunicação científica, ser editado no Estado de Minas Gerais, ter ISSN, ter o meio físico on-line e ser de acesso aberto. No geral, 423 estão ativos (69%), 101 inativos (17%) e 88 indisponíveis (14%). Com base nos registros identificados foi possível verificar e mensurar a quantidade de periódicos científicos on-line de acesso aberto constituídos formalmente por década, a quantidade de que permaneceram em atividade, que encerraram suas atividades e que estão com acesso indisponível. Observa-se que sobre a questão da acessibilidade dos periódicos de suporte digital, o grande problema de não preservar as informações dos trabalhos é que estes se perdem por completo. Conclui com a recomendação da consolidação de políticas públicas de apoio às publicações, sabendo que este suporte informacional é importante para o desenvolvimento científico e econômico da sociedade.

PALAVRAS-CHAVE *Organização do Conhecimento, Divulgação científica, Periódico Científico, Periódico científico on-line, Acesso à informação.*

ABSTRACT The aim of this work is to carry out an exploratory study of the rate of inactives and the availability of open access scientific journals. For this purpose, the State of Minas Gerais (Brazil) was used for the study. Following an in-depth bibliographical review, data was collected from a list of scientific journals which are registered at the Brazilian institute of Science and Technology Information (IBICT) at the request of ISSN (International Standard Number for Serial Publication). The definition of the publication situation of each journal in the cited universe was based on the information and on the evidences present in the journal's website. The data was collected and the results were submitted to a descriptive analysis. 612 journals were found that met the established criteria, namely to have scientific communication, to be edited in the State of Minas Gerais, to have ISSN, to have the physical medium on-line and to be open access. In general, 423 are active (69%), 101 inactive (17%) and 88 unavailable (14%). Based on the records identified, it was possible to verify and measure the number of open access on-line scientific journals formally constituted per decade, the number of those who remained in activity, who closed their activities and are unavailable. On the issue of accessibility of digital media, the biggest problem of not preserving the information of studies is

that they are lost altogether. The study concludes with the recommendation to consolidate public policies to support publications, knowing that this informational support is important for the scientific and economic development of society.

KEYWORDS *Knowledge organization, Scientific dissemination, Scientific journal, On-line scientific journal, Access to information.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Desde a invenção da imprensa por Johann Gutenberg, no século XV, a informação registrada cresce de forma exponencial e com o desenvolvimento das tecnologias computacionais, em especial após a Segunda Guerra Mundial no bojo da revolução científica e técnica, esse aumento ocasionou embaraços para acessar e recuperar a informação com eficácia. Bush, em 1945, descreve esse fenômeno como explosão informacional (Saracevic, 1996).

Além disso, após a era industrial - século XVIII e XIX - em que a riqueza provia pelo bem tangível, a informação tornou-se a nova fonte de valor, caracterizando a sociedade de tal forma que passou a ser denominada como da Informação e do Conhecimento. Nessa conjuntura, o desenvolvimento econômico e o bem-estar social tornaram-se estreitamente ligados à existência de um intelectual bem informado (Mayer-Schönberger & Cukier, 2013; Pisoschi & Pisoschi, 2016). O sucesso não é mais determinado pelo o que se possui e sim pelo saber (Prusak & Mcgee, 1994). Conforme explica Drucker (1993), se nas sociedades anteriores a pessoa instruída era um ornamento, na Sociedade do Conhecimento é o emblema, o símbolo, o porta bandeira... ela define a capacidade de desempenho da sociedade, mas também incorpora seus valores, crenças e compromissos.

Cientistas e inovadores nos domínios das diversas ciências e tecnologias só podem ampliar as fronteiras do conhecimento se forem asseguradas boas e rápidas informações. O sucesso do mercado dos pesquisadores e o desenvolvimento econômico implícito estão intimamente ligados ao acesso à informação, especialmente a altamente especializada e de qualidade (Mayer-Schönberger & Cukier, 2013).

O periódico científico – objeto de interesse da presente investigação –, criado em 1665 na Europa, promove normas de qualidade na condução da ciência e na sua comunicação por meio de critérios para avaliação dos artigos que integram seu conteúdo. Suas publicações são metodologicamente verificadas por pares. É o meio formal e historicamente utilizado para divulgar e debater acerca das pesquisas finalizadas ou em andamento (Severino, 2000; Mabe, 2001). É um meio de divulgação crível, de acesso rápido e prático no que diz respeito a recuperação da informação científica, cujo principal fim é disponibilizar para sociedade a produção científica das diversas áreas do conhecimento (Mueller, 2000; Curty & Boccato, 2005).

Os periódicos científicos on-line de acesso aberto, diferente do livro impresso, transpassam as barreiras físicas e geográficas, tornando o conhecimento acessível a qualquer lugar do mundo. Destacam-se como principal meio de visibilidade científica. Seguindo o traço do avanço da internet e a popularização dos portais virtuais que indexam e disponibilizam trabalhos científicos, na atualidade, assiste-se uma rápida expansão de títulos nesse formato, no entanto não há registro se continuam ativos.

Os periódicos encerram as publicações sem qualquer meio que possa monitorar a sua trajetória. Alguns diretórios, a exemplo do DOAJ (Directory of Open Access Journals. <https://doaj.org>) e do LATINDEX (Sistema Regional de Información en Línea para Revistas Científicas de América Latina, el Caribe, España y Portugal. <http://latindex.org>), apesar de terem a intenção de agrupar os registros dessa fonte de informação, não conseguem na integralidade e não possuem políticas de monitoramentos da atividade dos que estão inseridos. A Rede ISSN (ISSN Network. <http://www.issn.org>) também não tem essa atribuição como rotina de trabalho, além de não disponibilizar para consulta pública os títulos pertencentes ao seu catálogo.

O objetivo desta investigação é realizar um estudo exploratório da taxa de inativos e disponibilidade dos periódicos científicos on-line de acesso aberto. Pergunta-se: qual a taxa de sobrevivência dos periódicos científicos on-line de acesso aberto? Esses mantêm os trabalhos publicados disponíveis para acesso?

A importância deste trabalho se fundamenta em levantar a questão da descontinuidade dos periódicos e a preservação do acesso ao seu conteúdo - rico teor informacional que ajuda a compreender o desenvolvimento da ciência. A falta desse canal de informação interrompe as principais funções dos periódicos científicos: i) ser uma espécie de arquivo da ciência; ii) ser os veículos de divulgação e comunicação do saber e, iii) ser os meios de conferir prestígio e reconhecimento aos autores (Stumpf, 1998). Justifica-se pela ausência de pesquisa na temática.

Utilizou-se para o estudo o estado de Minas Gerais (Brasil) por trata-se de uma região brasileira polo de inovação e empreendedorismo. Com uma população estimada em 20,7 milhões, concentra 339 instituições de ensino superior, 10,2% de matrículas em cursos presenciais (631 mil) do país (Sindicato dos Mantenedores de Ensino Superior, 2015).

COMUNICAÇÃO CIENTÍFICA

A importância de se estudar a produção científica repousa na tese desta manter uma relação direta com o desenvolvimento científico, tecnológico e econômico de um país. Conforme Meadows (1999) quem mais produz Ciência e Tecnologia é quem mais avança no processo de desenvolvimento global. O filósofo Francis Bacon, já no século XVII, expressava que o conhecimento é poder, ou seja, o resultado da falta do conhecimento levaria também a falta de poder (Targino, 2001; Lehfeld, 2007).

A comunicação do conhecimento constitui parte essencial da ciência. Meadows (1999, p.vii) refere-se como tão vital quanto a própria ciência, pois a esta não cabe reivindicar com legitimidade este nome enquanto não houver sido analisada e aceita pelos pares. Isso exige, necessariamente, que seja comunicada.

Apoia-se na intenção de divulgar os resultados de pesquisa e também de possibilitar a colaboração entre os pares. Fundamenta-se na informação científica e representa um acréscimo ao entendimento universal até então existente sobre algum fato ou fenômeno. Isso porque a ciência possui caráter evolutivo e mutável, o que faz da pesquisa científica seu instrumento-mor e da comunicação científica seu elemento básico (Targino, 2000).

O advento das novas tecnologias e aprimoramento das formas de comunicação proporcionaram à informação maior velocidade de recuperação, acesso e disseminação. Sobretudo depois da internet,

mudou-se a forma de divulgar o conhecimento, alterou a interface e os serviços possíveis. Possibilitou muitas vantagens na forma de gerar e notificar a informação. Os limites geográficos em escala mundial foram eliminados de forma instantânea. (Rodrigues & Oliveira, 2012).

PERIÓDICO CIENTÍFICO

De acordo com Meadows o termo *periodical* (periódico) entrou em uso na segunda metade do século XVIII e se refere a qualquer publicação que apareça em intervalos determinados e contenha diversos artigos de diferentes autores (Meadows, 1999).

A história dos periódicos científicos teve seu início em janeiro de 1665, na França, quando o parisiense Denis de Sallo deu início à primeira revista, denominada *Journal des Sçavans*, dedicada a publicar todo tipo de notícia de interesse científico e cultural. Embora essa tenha alguma pretensão de ser a primeira revista científica, é na verdade uma mistura eclética de, principalmente, livro e outros comentários em uma ampla gama de assuntos, e até mesmo de romances, por exemplo, *Le Mariage de Belseger* (Meadows, 1999; Singleton, 2014).

Tendo obtido uma publicação do *Journal des Sçavans*, os membros da *Royal Society* discutiram a possibilidade de publicar algo semelhante. Entretanto, concluíram que poderia ser feito de maneira mais científica; excluíram algumas seções e deram ênfase aos relatos de casos de experimentos realizados pelos cientistas. Assim, deram origem ao periódico *Philosophical Transactions* em março de 1665. Também chamado de *Phil Trans*, é considerado o protótipo das revistas científicas. A periodicidade mensal logo alcançou 1200 exemplares, cuja subscrição era taxada em dez libras (Merlo, 2012).

Antes do seu surgimento, as informações sobre ciência eram veiculadas em folhetins, volantes e jornais cotidianos. Desde a invenção da imprensa até o século XVII essa prática era comum, porém, quando se fazia necessário a comunicação de um conhecimento mais especializado, esse era realizado por meio de correspondência entre os cientistas e/ou por atas ou memórias das reuniões científicas. As atas ou memórias transcrições das descobertas que eram relatadas durante as reuniões de uma sociedade e depois impressas de forma resumida para servirem de fonte de consulta (Freitas, 2006).

Desde sua gênese o periódico científico é um instrumento de referência valioso para acompanhar os avanços provenientes da Ciência e Tecnologia. Possibilita o acesso ao estado da arte dos campos de conhecimento, na medida em que difunde a produção consolidada pela *démarche* das pesquisas. Além disso, realiza com sucesso a disseminação e recuperação eficaz de informações científicas por dois fatores: o tempo (determinado pela velocidade de distribuição) e o espaço (sob a forma da área de distribuição) (Pisoschi & Pisoschi, 2016).

Por seu formato ágil, é o suporte ideal para obras que unam reflexão crítica com o desejo de intervenção na ciência e sociedade. Possui diversas funções, dentre elas a de informar, a de obter reações, a de registrar autoria e a de indicar e localizar documentos (Mueller, 2000).

Os artigos publicados em periódicos científicos são de suma importância para a divulgação formal mais ampla. O papel do periódico científico, e conseqüentemente do artigo científico, é fundamental para favorecer a geração e a comunicação de conhecimentos e de atividades de pesquisas (Mueller, 2000; Curty & Boccato, 2005).

ACESSO ABERTO

Com o surgimento da Web, na atualidade, estima-se que possa haver cerca de milhares de periódicos científicos no mundo (Spinak & Packer, 2015). As novas tecnologias melhoram a produtividade, permitindo o rápido crescimento de informações disponíveis (Meadows, 1999).

Atenta-se que o modelo tradicional de publicação tem sido, para muitos, aberto ao debate à medida que os preços dos periódicos chegam ao ponto em que as bibliotecas não podem continuar a renovar todas as suas assinaturas e a Internet oferece soluções alternativas cada vez mais atraentes (Poynder, 2001).

Com o objetivo declarado de eliminar os obstáculos que impedem o acesso à informação científica, decidindo por tornar a literatura acessível on-line e sem restrições, foi definida estratégias básicas fundamentadas no uso do protocolo *Open Archives Initiative – Protocol for Metadata Harvesting* (OAI-PMH). Em 2002 esse Movimento se consolidou com a publicação da *Budapest Open Access Initiative* (BOAI) que deliberou duas estratégias básicas que tem norteado as discussões sobre o tema desde então (Jange & Kademani, 1999; Packer, 2011).

A primeira, o auto-arquivamento, é determinada como a Via Verde (*Green Road*) e compreende o arquivamento, por parte dos autores, de artigos científicos já publicados ou aceitos para publicação em um periódico referendado cujo acesso se dá por assinatura. É garantida quando os repositórios os tornam os trabalhos disponíveis de forma ampla e livre. A segunda estratégia definida em Budapeste constitui a Via Dourada (*Golden Road*) e envolve os periódicos científicos eletrônicos cujo acesso aberto ao seu conteúdo é garantido pelos próprios editores. Nesse caso, a publicação se dá, primariamente, no próprio periódico (Costa, 2006).

ASPECTOS METODOLÓGICOS

Esta investigação realizou um estudo exploratório da taxa de inativos e disponibilidade dos periódicos científicos on-line de acesso aberto. Como aborda Gil (1999) as pesquisas exploratórias têm como principal finalidade desenvolver, esclarecer e modificar conceitos e ideias, com vistas na formulação de problemas mais precisos. Visa oferecer informações sobre o objeto desta e orientar a formulação de hipóteses (Cervo & Silva 2006).

Iniciou-se com uma revisão minuciosa da literatura. Por meio de pesquisa bibliográfica, trabalhos dos cientistas Fachin & Hillesheim (2006); Meadows (1999); Mueller (2000); Packer (2005); Rodrigues & Fachin (2008); Stumpf (1998) e Targino (2000 & 2001); dentre outros, auxiliaram a compreender o conceito e a formar o escopo teórico da investigação.

Após, realizou-se um levantamento de dados a partir da lista de periódicos científicos cadastrados no Centro Brasileiro do ISSN coordenado pela Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT. <http://cbissn.ibict.br>), para a solicitação do ISSN (*International Standard Serial Number*).

A escolha pelo catálogo do ISSN deu-se por possuir o cadastro mais abrangente dos periódicos. Apesar de não ser obrigatória a adesão, seu uso é definido pela norma técnica *Internacional Standards Organization ISO 3297* (Bomfá, 2003). É um parâmetro para o controle de qualidade de revistas científicas e também um critério de indexação em base de dados nacionais e internacionais (<http://cbissn.ibict.br/>).

Para o desenvolvimento da pesquisa, optou-se como amostra o estado brasileiro de Minas Gerais. Assim, o universo de investigação foi restringido para periódicos que atendem os seguintes critérios: possuir caráter de comunicação científica; ser editado no Estado de Minas Gerais; ter ISSN; ter o meio físico on-line e ser de acesso aberto.

Seguiu-se o seguinte percurso metodológico:

- Identificação dos periódicos científicos on-line de acesso aberto criados no estado de Minas Gerais, Brasil;
- Identificação da continuidade ou da interrupção dos periódicos científicos on-line de acesso aberto criados no estado de Minas Gerais, Brasil;
- Identificação da taxa de inativos dos periódicos científicos on-line de acesso aberto criados no estado de Minas Gerais, Brasil;
- Identificação da taxa de periódicos que não mantem a disponibilidade das edições publicadas.

As seguintes premissas foram utilizadas para a formulação do método de cálculo das taxas de inativos e de disponibilidade:

- Os periódicos que solicitam formalmente o número ISSN começam a operar (entram em atividade) a partir da data de constituição;
- Um indício de que os periódicos científicos estão ATIVOS é quando cumprem a periodicidade indicada;
- Um indício de que os periódicos estão INATIVOS, isto é, encerraram suas atividades, é quando se omitem por mais de dois anos com a periodicidade indicada ou informam que estão inativos; e
- Um indício de que os periódicos estão INDISPONÍVEIS é quando não é possível localizar a página de acesso, mesmo após três tentativas.

SITUAÇÃO DO PERIÓDICO CIENTÍFICO

A definição da situação da publicação de cada periódico do universo citado se baseou nas informações e indícios presentes nas variáveis da base de dados do próprio site do periódico.

Situação “ATIVO”:

Nesse estudo, para que um periódico constituído em meio eletrônico e com acesso aberto pudesse ser considerado “ATIVO” no ano de 2017, seria preciso que o mesmo:

- Apresente no ano 2017 a variável PERIODICIDADE diferente de “ENCERRADA” e/ou “INATIVA”;
- Tenha publicado no ano de 2017, 2016 e/ou 2015; e
- Possua indícios de estar “CORRENTE” no ano de 2017, por exemplo, com chamada pública para receber artigos.

Situação “INATIVO”:

Para que um periódico constituído em meio eletrônico e com acesso aberto seja considerado “INATIVO” no ano de 2017, é preciso que esse:

- Esteja com a variável PERIODICIDADE igual a “ENCERRADA” e/ou “INATIVA” e
- Não tenha publicado no ano de 2017, 2016 e 2015; e
- Não possua indícios de estar “CORRENTE” no ano de 2017, por exemplo, com chamada pública para receber artigos.

Situação “INDISPONÍVEL”:

- Quando não for possível localizar o site do periódico; ou
- Quando não for possível acessar o site do periódico, mesmo após três tentativas em diferente horário e dia.

APRESENTAÇÃO E ANÁLISE DOS DADOS

Os dados de cada periódico foram coletados e os resultados foram submetidos a uma análise descritiva. Atentou-se para a reanálise da listagem disponibilizada pelo IBICT para não incorrer em omissão de qualquer periódico científico on-line de acesso aberto do estado de Minas Gerais. Os títulos foram ordenados para evitar a duplicação.

A pesquisa localizou 612 periódicos que atendiam aos critérios estabelecidos. No geral, 423 estão ativos (69%), 101 inativos (17%) e 88 indisponíveis (14%), como é possível constatar na tabela 1 e 2.

Tabela 1. Quantidade de periódicos on-line de acesso aberto por década de criação

	Antes de 1960	Década de 1960	Década de 1970	Década de 1980	Década de 1990	Década de 2000	2010-2016	Total por situação
Ativo	9	4	15	23	55	149	168	423
Inativos	0	1	0	1	9	51	39	101
Indisponíveis	0	0	0	0	1	51	36	88
Total por década	9	5	15	24	65	251	243	612

Referência: Dados da pesquisa (junho 2017)

Tabela 2. Porcentagem de periódicos on-line de acesso aberto por década de criação

	Antes de 1960	Década de 1960	Década de 1970	Década de 1980	Década de 1990	Década de 2000	2010-2016	Total por situação
Ativo	100%	80%	100%	96%	85%	60%	69%	69%
Inativos	-	20%	-	4%	14%	20%	16%	17%
Indisponíveis	-	-	-	-	1%	20%	15%	14%

Referência: Dados da pesquisa (junho 2017)

Observa-se um crescimento exponencial quanto ao ano de criação desses registros. 81% dos periódicos foram lançados a partir do ano de 2000, 494 do total de 612 registros. Fato este compreensível pelas facilidades proporcionadas pelas novas tecnologias que agilizaram a criação, organização e manutenção dos instrumentos de informação.

Analisando os periódicos inativos (Figura 1), todos os criados antes de 1960 (9 periódicos) e da década de 1970 (15 periódicos) que alteraram o meio físico de impresso para digital continuam vigente. O primeiro periódico localizado data de 1894. Apenas dois (um da década de 60 e outro da década de 80) não são mais editados.

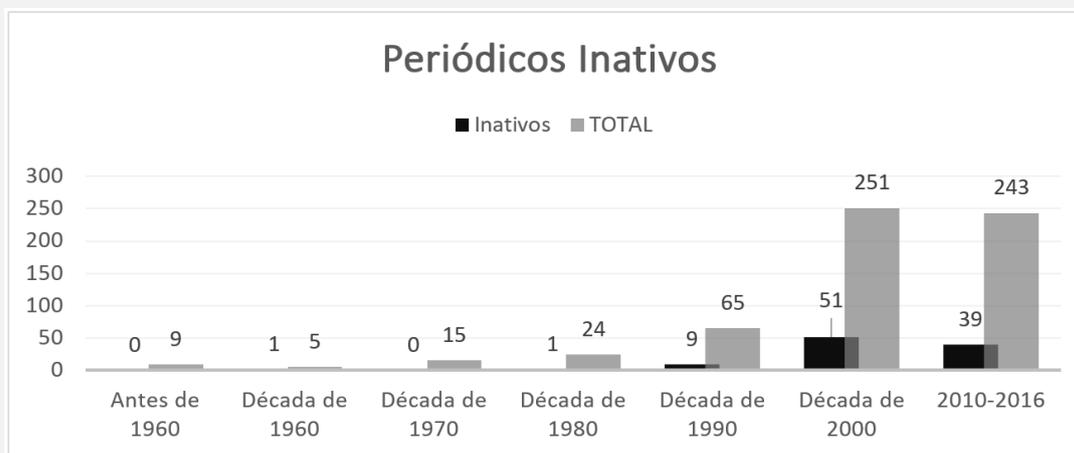


Figura 1. Quantidade de periódicos inativos on-line de acesso aberto por década de criação

Referência: Dados da pesquisa (junho 2017)

Nenhum periódico criado antes de 1980, que somam 53 periódicos, está indisponível (Figura 2), isto é, todos mantêm informações dos trabalhos publicados.

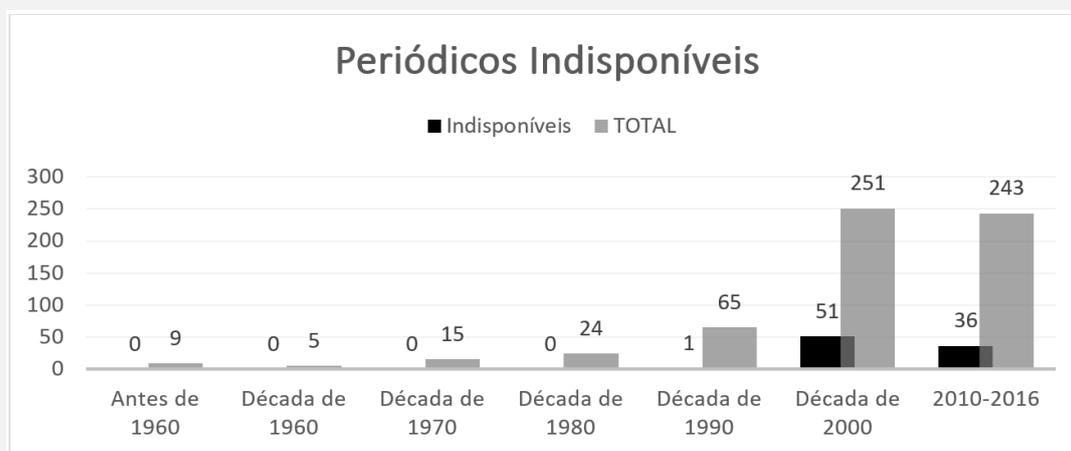


Figura 2. Quantidade de periódicos indisponíveis on-line de acesso aberto por década de criação

Referência: Dados da pesquisa (2017)

ANÁLISE POR DÉCADA

ANTES DE 1960

Foram localizados nove periódicos on-line de comunicação científica de acesso aberto do estado de Minas Gerais (Brasil) criados antes de 1960. Esses se adaptaram ao meio digital e publicam até os dias atuais.

Os periódicos com data de criação de antes de 1960 correspondem:

- 1% do universo.

O período possui:

- 100% periódicos on-line de acesso aberto ativos.

DÉCADA DE 1960

Cinco periódicos da década de 1960 são publicados até o presente e um vigorou até o ano de 2002, no entanto mantém disponível na internet todos os trabalhos publicados.

Representa:

- 1% do universo.

Possui:

- 80% periódicos on-line de acesso aberto ativos;
- 20% periódicos on-line de acesso aberto inativos.

DÉCADA DE 1970

No total de 15 registros localizados, todos estão ativos.

Representa:

- 2% do universo.

Possui:

- 100% periódicos on-line de acesso aberto ativos.

DÉCADA DE 1980

Foram 24 registros identificados. A grande maioria com a periodicidade sem atraso. Em comparação à década anterior, o número de registro praticamente duplica.

Representa:

- 4% do universo.

Possui:

- 96% periódicos on-line de acesso aberto ativos;
- 4% periódicos on-line de acesso aberto inativos.

DÉCADA DE 1990

É a década que se identifica pela primeira vez um registro que não é possível localizar o endereço eletrônico, apesar da tentativa com inúmeras estratégias.

Representa:

- 11% do universo.

Possui:

- 85% periódicos on-line de acesso aberto ativos;
- 14% periódicos on-line de acesso aberto inativos;
- 1% periódicos on-line de acesso aberto indisponíveis.

DÉCADA DE 2000

Período em que a internet se popularizou. O número de periódicos científicos dessa década em comparação a década passada quase que quadruplicou, em contrapartida, o número de periódicos inativos e inacessíveis aumentaram em proporções também elevadas.

Representa:

- 41% do universo.

Possui:

- 60% periódicos on-line de acesso aberto ativos;
- 20% periódicos on-line de acesso aberto inativos;
- 20% periódicos on-line de acesso aberto indisponíveis.

2010-2016

Em plena expansão, com apenas seis anos praticamente já possui a mesma quantidade numérica de periódicos científicos on-line de acesso aberto da década anterior (apenas oito a menos).

Essa tendência vem ao encontro da explosão informacional vivenciada no presente. Hoje, a sociedade está emergida com um grande volume informacional.

Destaque para o número de periódicos que não é mais possível localizar na internet, em números absolutos já somam 36. Em estatística ponderada, é maior que a década passada.

Esta década representa:

- 40% do universo.

Possui:

- 69% periódicos on-line de acesso aberto ativos;

- 16% periódicos on-line de acesso aberto inativos;
- 15% periódicos on-line de acesso aberto indisponíveis.

CONSIDERAÇÕES FINAIS

Existe uma íntima relação entre crescimento científico-tecnológico e crescimento econômico de um país. Como abordado por diversos pesquisadores, na contemporaneidade, o poderio político e econômico das nações está vinculado mais à sua capacidade científica e tecnológica do que à riqueza dos seus recursos naturais. A comunicação científica e suas fontes, em especial os periódicos científicos, são importantes elementos integrantes desse sistema.

Com base nos registros identificados foi possível verificar e mensurar a quantidade de periódicos científicos on-line de acesso aberto constituídos formalmente por década no estado brasileiro de Minas Gerais, a quantidade que permanecem em atividade, que encerraram suas atividades e que estão com acesso indisponível.

Dos periódicos da década de 1990 e anteriores, apenas 106 se adaptaram as novas tecnologias e passaram do meio físico para o on-line, e destes 90% continuam ativos. Após o ano 2000, quando a internet passou a alcançar a população em massa, houve uma alteração significativa do cenário. A partir dessa década, além do crescimento exponencial dos periódicos, esses, apesar das facilidades da tecnologia passaram a tornar-se inativos e inacessíveis em maior porcentagem.

Sobre a questão da acessibilidade dos periódicos de suporte digital, o grande problema de não preservar as informações dos trabalhos é que estes se perdem por completo. Caso não houver algum repositório de garanta o acesso, não é mais possível recuperar seus dados. Este fato gera um transtorno para os pesquisadores e uma perda imensurável para a memória da ciência.

Muitas publicações, por mais importantes que sejam, não estão disponíveis e mesmo cientistas experientes encontram dificuldade em ter acesso a esse material. A tendência de formato apenas por meio digital eleva a questão uma dimensão maior. Quando não existe a preocupação de conservar a disponibilidade do endereço eletrônico, ou direcionamento para outro, a informação simplesmente não existe e não é possível ser recuperada por qualquer forma.

É necessário um estudo mais aprofundado a respeito das causas de interrupção dos periódicos científicos. Recomenda-se a consolidação de políticas públicas de apoio às publicações, sabendo que este suporte informacional é importante para o desenvolvimento científico e econômico da sociedade. Já é possível constatar que 14% dos periódicos de acesso aberto, do estado estudado, não é possível localizar.

REFERÊNCIA

Bomfá, C.R.Z. (2003). *Revistas Científicas em Mídias Digit@l: critérios e procedimentos para publicação*. Santa Catarina: VisualBook.

Budapest Open Access Initiative - BOAI. (2012). *Ten years on from the Budapest Open Access Initiative: Setting the default to open*. Recuperado em 25 junho, 2017, de <http://www.opensocietyfoundations.org/openaccess/boai-10-recommendations>

Costa, S. M. S. (2006). Filosofia aberta, modelos de negócios e agências de fomento: elementos essenciais a uma discussão sobre o acesso aberto à informação científica. *Ciência da Informação*, 35(2), 39-50. Recuperado em 25 junho, 2017, de <https://dx.doi.org/10.1590/S0100-19652006000200005>

Cervo, A.L. & Silva, R. (2006). *Metodologia científica*. São Paulo: Prentice Hall Brasil.

Curty, M.G. & Boccato, V.R.C. (2005). O artigo científico como forma de comunicação do conhecimento na área de Ciência da Informação. *Perspect. ciênc. inf.*, 10(1), 94-107. Recuperado em 25 junho, 2017, de <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/305/108>.

Drucker, F.P. (1993). *Sociedade Pós-Capitalista*. São Paulo: Pioneira.

Fachin, G.R.B. & Hillesheim, A.I.A. (2006). *Periódico científico: padronização e organização*. Florianópolis: Ed. da UFSC.

Freitas, M.H. (2006). Considerações acerca dos primeiros periódicos científicos brasileiros. *Ci. Inf.*, 35(3), 54-66. Recuperado em 25 junho, 2017, de http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652006000300006

Gil, A.C. (1999). *Métodos e técnicas de pesquisa social*. São Paulo: Atlas.

Jange, S. & Kademani, B.S. (1999) Metamorphosis of the scientific journal: past, present and future. *Malaysian Journal of Library & Information Science*, 4(1), 61-69.

Lehfeld, N. (2007). *Metodologia e conhecimento científico: horizontes virtuais*. Petrópolis: Vozes.

Mabe, M.A. (2001). Digital dilemmas: Electronic challenges for the scientific journal publisher. *Aslib Proceedings*; Bradford, 53(3), 85-92.

Mayer-Schönberger, V. & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.

Meadows, A.J. (1999). *A comunicação científica*. Brasília: Briquet Lemos.

Merlo, I. (2012). Das origens das revistas científicas ao Jornal Vascular Brasileiro. *J. vasc. bras.*, 11(2), 93-94. Recuperado em 25 junho, 2017, de http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1677-54492012000200002&lng=en&nrm=iso

Mueller, S.P.M. (2000). O periódico científico. In: Campello, B. S.; Cendón, B. V. & Kremer, J. M. (Org.). *Fontes de informação para pesquisadores e profissionais*. Belo Horizonte: UFMG.

Packer, A.L. (2005). A construção coletiva da Biblioteca Virtual em Saúde. *Interface – Comunicação Saúde, Educação*, 9(17), 249-72. Recuperado em 25 junho, 2017, de <http://www.scielo.br/pdf/icse/v9n17/v9n17a04.pdf>.

Packer, A. L. (2011). Os periódicos brasileiros e a comunicação da pesquisa nacional. *Rev. USP, São Paulo*, 89. Recuperado em 25 junho, 2017, de http://rusp.scielo.br/scielo.php?script=sci_arttext&pid=S0103-99892011000200004&lng=pt&nrm=iso

Pisoschi, A.M. & Pisoschi, C.G. (2016) Is open access the solution to increase the impact of scientific journals? *Scientometrics*, 109, 1075–1095.

Prusak, L. & Mcgee, J. (1994). *Gerenciamento Estratégico da Informação*. Rio de Janeiro: Ed. Campus.

Rodrigues, R. & Fachin, G.R.B. (2008). A comunicação científica e o uso de portais: estudo. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 9., São Paulo. Anais eletrônicos, ECA-USP; ANCIB. Recuperado em 25 junho, 2017, de <http://enancib.ibict.br/index.php/enancib/ixenancib/paper/view/3113/2239>

Rodrigues, R.S. & Oliveira, A.B. (2012). Periódicos científicos na America Latina: títulos em acesso aberto indexados no ISI e SCOPUS. *Perspect. ciênc. inf.*, 17(4), 77-99. Recuperado em 25 junho, 2017, de <http://www.scielo.br/pdf/pci/v17n4/06.pdf>

Saracevic, T. (1996). Information science: origin, evolution and relations. *Perspect. ciênc. inf.*, 1(1). Recuperado em 25 junho, 2017, de <http://basessibi.c3sl.ufpr.br/brapci/v/a/16970>.

Severino, A.J. (2000). *Metodologia do trabalho científico*. São Paulo: Cortez.

Sindicato dos Mantenedores de Ensino Superior. (2015). *Mapa de ensino superior*. Brasília. Recuperado em 25 junho, 2017, de <http://convergenciacom.net/pdf/mapa-ensino-superior-brasil-2015.pdf>.

Singleton, A. (2014). The first scientific journal. *Learned Publishing*, 27(1).

Spinak, E. & Packer, A. (2015). 350 anos de publicação científica: desde o “Journal des Sçavans” e “Philosophical Transactions” até o SciELO [on-line]. *SciELO em Perspectiva*. Recuperado em 25 junho, 2017, de <http://blog.scielo.org/blog/2015/03/05/350-anos-de-publicacao-cientifica-desde-o-journal-des-scavans-e-philosophical-transactions-ate-o-scielo>

Stumpf, I.R. (1998). *Periódicos científicos: Documentos ABBD*, 8. Porto Alegre, Associação Brasileira de Ensino em Biblioteconomia e Documentação.

Targino, M.G. (2000). Comunicação científica: uma revisão de seus fundamentos básicos. *Informação e Sociedade: estudos*, 10(2), 1-27. Recuperado em 20 março, 2017, de <http://www.ies.ufpb.br/ojs/index.php/ies/article/view/326/248>.

Targino, M.G. (2001). Comunicação científica na sociedade tecnológica: Periódicos eletrônicos em discussão. *Comunicação e Sociedade*, 3(1-2), 93-112.



DESAFIOS À ORGANIZAÇÃO E ACESSO AO PATRIMÓNIO CULTURAL

USO DE UM PERFIL DE APLICAÇÃO PARA A GESTÃO E AGREGAÇÃO DA INFORMAÇÃO AUDIOVISUAL DA UNIVERSIDADE DO PORTO: O CASO TVU

Sara Oliveira

*Mestre em Ciência da Informação da Faculdade de Engenharia da Universidade do Porto,
saraoliveira878@gmail.com*

RESUMO A presente comunicação resulta do projeto de dissertação desenvolvido no Mestrado em Ciência da Informação e tem como objetivo demonstrar como gerir o património cultural e científico, em formato audiovisual, e como aumentar a sua visibilidade criando uma janela de oportunidade para a produção e disseminação do conhecimento produzido na Universidade do Porto. Como abordagem metodológica, considerou-se o método de Investigação-ação, desenvolvendo para o efeito um único ciclo, com as inerentes cinco fases. A principal contribuição, num contexto de experimentação, consiste na construção de um perfil de aplicação Dublin Core (DCAP) – o Audiovisual Content Application Profile (ACAP) –, que contribua para uma melhor gestão dos dados gerados no âmbito da produção audiovisual da TVU, e que possibilite a sua integração no agregador especializado EUscreen e posterior acesso através do portal Europeana.

PALAVRAS-CHAVE *Informação audiovisual, Património cultural e científico, Perfil de aplicação, EUscreen, Universidade do Porto.*

ABSTRACT The present communication results from a dissertation project developed in the Masters in Information Science and has as goal show how manage cultural and scientific heritage, in audiovisual format, and how to increase your visibility by creating a window of opportunity for the production and dissemination of knowledge produced in University of Porto. As a methodological approach, it was considered the Action-Research method, developing for this purpose a single circle, with the inherent five phases. The main contribution, in an experimental context, is the construction of an application profile Dublin Core (DCAP) - the Audiovisual Content Application Profile (ACAP) -, which contributes to a better management of the data generated in scope of TVU's audiovisual production and to enable their integration into the specialized aggregator named EUscreen and later access through the Europeana portal.

KEYWORDS *Audiovisual Information, Cultural and scientific heritage, Application profile, EUscreen, University of Porto.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Com o desenvolvimento das TIC, emergiu e afirmou-se a prática da criação e divulgação de conteúdos audiovisuais em plataformas *online*. “A produção audiovisual deixou de ser pensada exclusivamente para *broadcasting*, e passou a ser estudada e planificada para a *Web*, de forma a se adequar às

características do novo meio e a tirar partido das diversas potencialidades que este coloca ao seu dispor” (Oliveira, 2010).

As funcionalidades de transmissão, visualização e partilha na Internet fazem com que os conteúdos audiovisuais cheguem rapidamente aos consumidores, em tempo real e à simples distância de um clique. Dada a facilidade de disseminação desta produção digital, são muitas as entidades que atualmente optam por publicitar os seus produtos e serviços através deste meio importando compreender os inerentes processos e impactos.

A presente comunicação resulta da investigação realizada neste domínio, no âmbito do Mestrado em Ciência da Informação¹. Coloca-se o foco na produção informacional audiovisual e remete-se, especificamente, para a TVU., um serviço da Universidade do Porto onde são produzidos e distribuídos conteúdos audiovisuais com uma particular incidência na instituição, na ciência e na cultura. Este canal universitário junta as pessoas através da partilha de conhecimento e interesses e permite que o público fique não só a conhecer os acontecimentos e marcos importantes de uma instituição dedicada à formação superior e à criação de conhecimento científico, cultural e artístico, como também a contribuir de diferentes formas para este acervo.

A promoção e disseminação dos conteúdos audiovisuais da TVU. são atualmente realizadas através de um portal colaborativo, no qual a informação associada aos recursos sob a forma de metadados desempenha um papel preponderante. No entanto, uma análise ao mesmo permitiu identificar dois problemas: um número de metadados insuficiente para descrever todas as características e propriedades dos conteúdos; e um baixo nível de disseminação e de visibilidade dos conteúdos audiovisuais na *Web*. Deste modo, foram fixados como principais objetivos: 1) melhorar a gestão de dados da TVU., através da inclusão de metadados que proporcionem uma descrição mais completa e representativa dos conteúdos, nas duas faces do portal (*front office* e *back office*); 2) integrar a TVU. na rede de instituições parceiras do EUscreen, a fim de possibilitar a agregação dos conteúdos audiovisuais na Europeia, tornando-os visíveis e disponíveis para uma audiência à escala global.

A informação audiovisual tem ganho um papel de cada vez maior destaque na preservação da memória cultural, nomeadamente ao nível das universidades, instituições cuja memória evidencia os seus traços identitários, a importância da sua ação e o impacto produzido nas comunidades com que interage.

Para que a promoção e disseminação dos conteúdos audiovisuais na *Web* garantam a comunicação da ciência e o acesso aberto ao conhecimento e à cultura, torna-se necessário recorrer a tecnologias da chamada *Web Semântica* que expande a *Web* existente (Harth, Janik, & Staab, 2010), adicionando uma multidão de padrões de linguagem e componentes de *software* que proporcionam aos seres humanos e às máquinas um acesso direto aos dados.

Entre as tecnologias da *Web Semântica* encontram-se os esquemas de metadados. Estes são definidos por Ikematu (Ikematu, 2009) como dados associados a objetos e que ajudam os potenciais utilizadores a ter um conhecimento completo das suas características. Também as ontologias estão presentes no conjunto de tecnologias da *Web Semântica*, sendo definidas por Almeida e Bax (Almeida & Bax, 2003)

¹ Oliveira, S. (2017). *Disseminação de conteúdos audiovisuais na web: uso de um perfil de aplicação para a gestão e agregação dos recursos da TVU.* (Dissertação de Mestrado), Universidade do Porto. Desenvolvida sob a orientação de Cristina Ribeiro e coorientação de Maria Manuela Pinto.

como estruturas que se organizam a partir de conceitos e dos seus relacionamentos, num domínio do conhecimento.

De entre um conjunto de vocabulários abertos analisados, dá-se no estudo realizado um particular destaque ao padrão de metadados Dublin Core, um vocabulário formado por propriedades amplas e genéricas, comumente utilizado na descrição de todo o tipo de recursos digitais na *Web* (Sayão, 2010), e ao padrão de metadados *VideoMD*, um esquema XML de metadados para descrever as características técnicas de objetos de vídeo digital, originalmente criado pela Biblioteca do Congresso (Library of Congress, 2015). Entre as ontologias analisadas no estudo destacam-se a *DBpedia Ontology*, um projeto que resulta de um esforço comunitário para extrair informação estruturada do Wikipédia e torná-la acessível na *Web* (DBpedia, 2017), a *EBUcore Ontology*, uma ontologia propositadamente concebida com uma lista mínima e flexível de atributos para descrever conteúdos áudio e vídeo (EBU, 2017), e a *Programmes Ontology*, uma ontologia criada pela British Broadcasting Corporation (BBC), com o objetivo de modelar os programas, séries e episódios transmitidos pela emissora pública de televisão.

A existência de vocabulários específicos, que descrevem e ligam dados na *Web*, é importante para que a interoperabilidade seja assegurada, no entanto, os metadados pertencentes a um único vocabulário podem não ser suficientes para descrever ao pormenor um determinado conjunto de dados. A necessidade de descrever conjuntos de dados de diferentes domínios tem levado, assim, à criação de perfis de aplicação (Castro, Ribeiro, & Silva, 2013), do inglês *application profiles*. Estes são esquemas de metadados que incorporam elementos de dados retirados de um ou mais esquemas existentes, combinados entre si e otimizados para uma aplicação local em particular (Heery & Patel, 2000).

A comunidade Dublin Core recomenda a construção de perfis de aplicação de acordo com as regras definidas pelo *Singapore Framework*, um modelo de referência para a descrição de recursos, que tem como objetivo implementar a interoperabilidade semântica entre diferentes comunidades de prática (Malta & Baptista, 2014). A um perfil de aplicação construído segundo as regras deste modelo é dado o nome de *Dublin Core Application Profile* (DCAP). O desenvolvimento de um DCAP deverá, assim, ser considerado e conter requisitos funcionais, modelo de domínio, *description set profile*, guias de utilização e guias de sintaxe.

A esta opção há que acrescer o papel dos agregadores de conteúdos audiovisuais. Este torna-se essencial, não só para permitir a disponibilização dos conteúdos no portal Europeana, mas também no mapeamento e uniformização dos modelos de metadados provenientes de diferentes origens para um esquema em comum, o *Europeana Data Model* (EDM). Entre os projetos agregadores destes conteúdos, identificou-se e selecionou-se o EUscreen, que permite a uniformização de diferentes esquemas de metadados através de uma ferramenta de nome *Metadata Ingestion Tool* (MINT).

O dispositivo metodológico para a realização deste estudo incluiu a Investigação-ação, desenvolvida, nesta fase, ao longo de um ciclo com as respetivas cinco fases.

Este projeto apresenta dois resultados principais. O primeiro consiste na concretização de um DCAP, que visa não só responder às necessidades da TVU., melhorando a sua gestão de dados, mas, também, viabilizar a agregação dos recursos na Europeana. O segundo resultado é a avaliação dos níveis de usabilidade e agregação do DCAP, permitindo perceber, respetivamente, se os metadados que fazem parte do perfil de aplicação são facilmente preenchidos pelos colaboradores da TVU., e se cumprem os requisitos definidos pelo portal europeu EUROPEANA. Ao DCAP foi atribuído o nome de *Audiovisual Content Application Profile* (ACAP).

METODOLOGIA

A Investigação-ação foi a metodologia adotada para guiar o trabalho desenvolvido. Esta distingue-se de outras, pois, tendo como foco principal a resolução de problemas reais que ocorrem em situações do mundo real, considera a colaboração ativa entre o investigador e os membros do serviço/ organização, transformando-os, também, em investigadores e agentes de mudança (O'Brien, 1998). Gerald Susman (1983), mencionado por O'Brien (1998), elaborou um modelo que integra um ciclo de cinco fases, mostrado na Figura 1 e reproduzível tantas vezes quanto necessário, e que foi considerado no estudo aqui exposto.

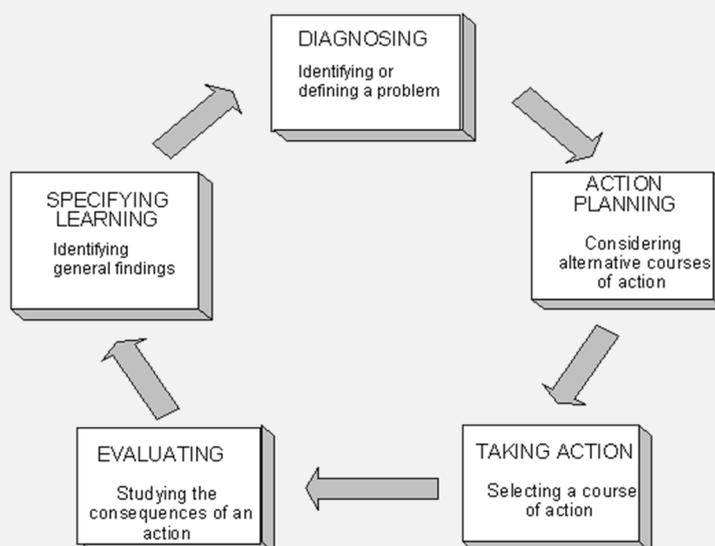


Figura 1. Modelo Investigação Ação (Susman, 1983)

Na primeira fase do ciclo, *Diagnosing*, foram identificados os problemas com que este serviço da Universidade do Porto atualmente se depara. Este diagnóstico foi realizado com base nas técnicas da observação direta, análise de conteúdo e entrevistas semiestruturadas e abertas, semelhantes a conversas normais. A observação direta ao *front office* e ao *back office* do portal, e a análise de conteúdo dos vídeos, permitiram detetar o primeiro problema: os metadados presentes no portal não fornecem uma descrição completa e representativa dos conteúdos audiovisuais. As entrevistas semiestruturadas à coordenadora da TVU. e restantes colaboradores, levaram à identificação do segundo problema: os conteúdos audiovisuais, quando publicados no portal, ganham um crescente número de visualizações, mas depois ficam estagnados, demonstrando que o património universitário não é suficientemente acedido e disseminado. Esta fase do diagnóstico envolveu ainda a recolha de documentação na TVU., com o intuito de compreender o seu funcionamento interno e a evolução que este serviço de produção audiovisual sofreu com o passar dos anos.

Realizado o diagnóstico, passou-se à segunda fase do ciclo, *Action Planning*, onde foram analisadas soluções para a resolução dos dois problemas identificados. A solução para o primeiro problema passaria por fornecer uma melhoria à gestão de dados da TVU., através da inclusão de um maior número de metadados nas duas faces do portal, e a solução para o segundo problema seria possibilitar o acesso aos conteúdos audiovisuais através de portais reconhecidos por diferentes audiências, nomeadamente a Europeia.

Colocadas estas duas possibilidades, foi definida uma solução comum para os dois problemas, que corresponde à fase *Taking Action*. A solução comum encontrada foi a elaboração de um perfil de aplicação, que respeitasse as regras definidas pelo modelo *Singapore Framework*, ao qual foi dado o nome de ACAP. O ACAP deveria incluir metadados que respondessem às necessidades da TVU. e que ao mesmo tempo viabilizassem a disseminação dos conteúdos audiovisuais ao nível da Europeia, através do agregador de audiovisuais EUscreen. Para a concretização desta ação foi necessário recorrer à revisão da literatura, que se focou, primeiramente, em construir um enquadramento teórico sobre os conceitos de “conteúdo audiovisual”, “arquivo audiovisual”, e “património audiovisual”. Seguidamente, foram analisados vocabulários comumente utilizados na descrição de recursos audiovisuais e estudadas as regras definidas pelo *Singapore Framework* para a criação de um perfil de aplicação. De modo a assegurar uma futura agregação dos conteúdos produzidos pela TVU. na Europeia, foram estabelecidos contactos via *email* com os responsáveis do agregador de audiovisuais EUscreen, com o objetivo de integrar a TVU. na rede de instituições parceiras.

Elaborado o ACAP, prosseguiu-se com a avaliação dos resultados da ação – o perfil de aplicação -, passando, assim, para a fase *Evaluating*, junto dos colaboradores da TVU. e do agregador de conteúdos audiovisuais EUscreen.

Estas duas avaliações permitiram tirar conclusões e interpretá-las à luz de quão bem-sucedida foi a ação, chegando-se, assim, à última fase do ciclo, *Specifying Learning*. A avaliação do ACAP junto dos colaboradores da TVU. permitiu perceber se os metadados escolhidos para o perfil de aplicação correspondem às necessidades do serviço de produção audiovisual, e se a curadora e a restante equipa possuem conhecimentos para o preenchimento de todas as propriedades. A avaliação do ACAP junto do agregador de audiovisuais EUscreen permitiu obter percentagens representativas do nível de agregação que uma futura exportação do perfil de aplicação poderá oferecer.

RESULTADOS

Os resultados deste projeto foram obtidos através dos testes que permitiram avaliar os níveis de usabilidade e de agregação do ACAP.

Para a concretização da avaliação do nível de usabilidade do ACAP, foi necessário proceder à sua formalização em ontologia, através de um *software* gratuito e de código aberto de nome *Protégé*. A ontologia foi ingerida no Dendro, uma plataforma colaborativa e multidisciplinar de gestão de dados de investigação, concebida por uma equipa de investigadores da Faculdade de Engenharia da Universidade do Porto². Após a ingestão da ontologia na plataforma Dendro, foi pedido aos colaboradores da TVU. que atribuíssem valores às propriedades do ACAP, com o objetivo de perceber se os futuros utilizadores do perfil de aplicação entendiam o significado dos metadados selecionados, provenientes de diferentes vocabulários. A curadora, colaboradora responsável pela preservação, manutenção e arquivo dos dados digitais, afirmou sentir dificuldades no preenchimento dos metadados que visam o registo das propriedades técnicas dos vídeos. Por outro lado, os colaboradores ligados à produção e edição de vídeo mencionaram não sentir qualquer dificuldade no preenchimento dessas propriedades, pois tratam-se de conceitos com os quais contactam diariamente. De um modo geral, todos consideraram que o perfil de

² Equipa que acolheu a fase de experimentação do projeto que aqui se apresenta.

aplicação estava bem definido e que correspondia às necessidades da TVU. A dificuldade demonstrada pela curadora em preencher as propriedades técnicas não foi considerado um problema, pois também os colaboradores ligados à produção e edição de vídeo têm acesso ao *back office*, podendo eles próprios preencher essas propriedades. Outra solução foi levantada, como é o caso da elaboração de uma tabela pré-formatada com os valores a atribuir a cada uma das propriedades técnicas, isto porque há propriedades cujos valores são comuns a grande número de vídeos.

Para a concretização da avaliação do nível de agregação do ACAP, foi necessário criar um ficheiro em formato XML (formato requerido para a agregação), com o objetivo de simular uma exportação do ACAP do portal. Nesse ficheiro foram incluídos os metadados e respetivos valores, relativos a três conteúdos audiovisuais representativos da história da Universidade do Porto, e dois valores essenciais para a agregação (o URL dos vídeos e o URL do *thumbnail* dos vídeos). O ficheiro foi inserido na ferramenta MINT e foi realizado o processo de agregação à Europeana com o objetivo de perceber se os metadados selecionados para o ACAP correspondiam aos requisitos do esquema EDM. Depois de realizadas as etapas do processo de agregação, foi possível retirar resultados sob a forma de percentagens. O ACAP consegue preencher todos os metadados obrigatórios do esquema EDM (100%), mais de metade dos metadados opcionais recomendados (57%), e um baixo número dos metadados opcionais não prioritários (13%). Foi também possível deduzir o número de metadados agregados pertencentes a cada vocabulário, sendo que o vocabulário com o maior número de propriedades agregadas foi o Dublin Core (8), e o vocabulário com o menor número de propriedades agregadas foi o DBpedia Ontology (1). A avaliação do nível de agregação permitiu ainda pré-visualizar como ficaria a agregação dos conteúdos audiovisuais da TVU. na Europeana, usando o ACAP, como mostra a Figura 2.

Coleção Egípcia da Universidade do Porto - integral

Description: A TVU. realizou em parceria com o investigador Rogério Sousa do CITCEM uma produção em vídeo e fotografia em que cerca de 50 dos objectos egiptológicos do património do Museu de História Natural da U.Porto aparecem no seu contexto histórico e cultural e nos contam pequenas narrativas associados a cada um deles

Subject: CITCEM, Reitoria, História da Arte

Creator: TVU.

Contributor: CITCEM, U.Porto, Museus da U.Porto

Place: Reitoria da U.Porto

Date: 2014-12-29

Publication date: 2014-12-30

Type: Documentário

Format: 4:3Coresmono 00:02:35

Relation: <http://tv.up.pt/videos/hfn2vybo> <https://repositorio-tematico.up.pt/handle/10405/22824>

Is part of: <http://tv.up.pt/premiums/47>

Language: Portuguese

Rights: Atribuição – Uso Não-Comercial – Partilha nos Termos da Mesma Licença (by-nc-sa). Esta licença permite a redistribuição, comercial ou não-comercial, desde que a sua obra seja utilizada sem alterações e na íntegra. É também essencial que seja dado o devido crédito ao autor da obra original.

Data provider: TVU.

Provider: EUscreenXL

Search also for:

Title
Coleção Egípcia da Universidade do Porto - integral

Who
TVU.
CITCEM, U.Porto, Museus da U.Porto

What
Documentário
CITCEM, Reitoria, História da Arte
4:3Coresmono

Provider
TVU.
EUscreenXL

Figura 2. Pré-visualização da agregação na Europeana

Através da observação da figura anterior, realça-se a presença do *Thumbnail* que funciona como um ponto de atração para o utilizador que navega na Europeana, e a presença de uma hiperligação que encaminha o utilizador para o portal da TVU. contribuindo para um aumento do número de visitas e para um aumento do número de visualizações dos conteúdos.

Os metadados “dc: relation” e “dcterms: isPartOf” são considerados metadados muito importantes, pois permitem a ligação para outros conteúdos e fontes. O primeiro permite a ligação para outro conteúdo no portal da TVU. relacionado com o tema Coleção Egípcia, e para um artigo de jornal presente no Repositório Temático da U. Porto, cujo assunto é Coleção de Objetos Assírios. O segundo permite a ligação para um dos canais premium da TVU., levando o utilizador a explorar o portal e potenciais conteúdos de interesse. Estas ligações para conteúdos e fontes externas são consideradas enriquecedoras, não só para o utilizador da Europeana, mas também para os portais/ repositórios para onde as ligações são remetidas.

No canto superior direito, em Search also for, é possível recuperar conteúdos relacionados dentro do portal. É permitido recuperar, por exemplo, conteúdos com o mesmo título ou idêntico, com as mesmas palavras-chave, e conteúdos classificados como sendo do mesmo tipo, neste caso, documentários.

CONCLUSÕES

Dos resultados apresentados é possível inferir que o nível de usabilidade do ACAP é favorável, uma vez que apenas a curadora apresentou dificuldades no preenchimento de um determinado tipo de metadados, o que pode ser resolvido através da intervenção dos colaboradores ligados à produção e edição de vídeo. No que respeita ao nível de agregação do ACAP, este é igualmente positivo, pois um grande número dos metadados requeridos no esquema EDM foi preenchido, revelando um elevado nível de agregação à Europeana.

Relativamente aos problemas e limitações do trabalho, importa realçar a dificuldade na seleção de padrões de metadados e ontologias que se aplicassem às necessidades deste projeto, uma vez que foi necessária uma seleção rigorosa por entre a vasta lista de vocabulários existentes.

É, também, necessário mencionar o modelo *Singapore Framework* seguido para a realização do perfil de aplicação. Embora se trate de um modelo de referência recomendado pela comunidade Dublin Core, lamenta-se a pouca informação existente sobre o assunto, o que dificultou o processo de elaboração do ACAP. As únicas fontes encontradas foram as páginas *web* mantidas pela comunidade Dublin Core, sendo que a página que fornece os guias para a realização de um DCAP data de 2009, revelando que não foram feitas muitas atualizações a este modelo.

Através de um pedido de exportação do atual portal da TVU., com o intuito de perceber qual o nível de agregação do portal, foi possível concluir que o mesmo não está preparado para a agregação dos conteúdos na Europeana, e que deve sofrer significativas melhorias, que poderão ser conseguidas através da implementação do ACAP.

Apesar de o contacto com o agregador de audiovisuais e o manuseamento da ferramenta MINT terem sido enriquecedores, importa mencionar que os metadados do esquema EDM encontra-se dividido em diferentes categorias (obrigatórios, opcionais dentro do grupo de obrigatórios, opcionais recomendados e opcionais não prioritários), o que dificultou bastante a fase de mapeamento/ enriquecimento no

processo de agregação. Esta opinião pessoal já foi transmitida aos colaboradores do EUscreen, que prometeram ter em consideração este aspeto em futuras reformulações, uma vez que já mais provedores de dados que contactaram com a ferramenta MINT referiram este aspeto. Outro aspeto negativo encontrado está relacionado com o facto de algumas propriedades estarem integradas no grupo dos metadados opcionais não prioritários. Estas propriedades permitem a inclusão de hiperligações para conteúdos relacionados, externos à Europeia, e, portanto, um maior nível de descoberta. No teste realizado, que visou a simulação da agregação à Europeia, foram atribuídos a esses valores hiperligações para outros conteúdos da TVU., e para conteúdos presentes nos repositórios aberto e temático da U. Porto., contribuindo assim para uma maior visibilidade do património da instituição de ensino. Dada a importância destes metadados, não se justifica estarem incluídos num grupo de metadados considerados de tão pouca importância pela Europeia. Os provedores têm tendência em utilizar primeiramente os metadados obrigatórios e opcionais recomendados, ficando muitas vezes os metadados opcionais não prioritários esquecidos.

Este trabalho implicou que a TVU. ficasse incluída na rede de instituições parceiras do EUscreen, fornecendo-lhe oportunidades de disseminação e visibilidade na *Web*. Como recomendações destaca-se a assinatura do *Data Exchange Agreement*, que permitirá que a TVU. se torne numa parceira oficial do EUscreen, a seleção rigorosa de conteúdos audiovisuais representativos do património da Universidade do Porto para uma futura agregação na Europeia, e a implementação do ACAP, com vista à facilitação do processo de agregação dos conteúdos e consequente disseminação à escala global via Europeia.

REFERÊNCIAS BIBLIOGRÁFICAS

- Almeida, M. B., & Bax, M. P. (2003). Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. *Ciência da Informação*, 32(3), 7-20. doi:<http://dx.doi.org/10.1590/S0100-19652003000300002>
- Castro, J. A., Ribeiro, C., & Silva, J. R. d. (2013). *Designing an Application Profile Using Qualified Dublin Core: A Case Study with Fracture Mechanics Datasets* Paper presented at the DCMI International Conference on Dublin Core and Metadata Applications, Lisboa. <http://dcpapers.dublincore.org/pubs/issue/view/165>
- DBpedia. (2017). DBpedia. Retrieved from <http://wiki.dbpedia.org/about>
- EBU. (2017). EBU - Operating Eurovision and Euroradio. Retrieved from <https://www.ebu.ch/about>
- Harth, A., Janik, M., & Staab, S. (2010). *Semantic Web Architecture*. Retrieved from <https://userpages.uni-koblenz.de/~staab/Research/Publications/2011/final-semantic-web-architecture.pdf>
- Heery, R., & Patel, M. (2000). Application profiles: mixing and matching metadata schemas. *Ariadne*(25).
- Ikematu, R. S. (2009). Gestão de metadados: sua evolução na tecnologia da informação. *DataGramZero - Revista de Ciência da Informação*, 2(6), 1-5.

Library of Congress. (2015, 13 de julho de 2015). Standards at the Library of Congress. Retrieved from <https://www.loc.gov/standards/>

Malta, M. C., & Baptista, A. A. (2014). A panoramic view on metadata application profiles of the last decade *Int. J. Metadata, Semantics and Ontologies*, 9(1), 58-73.

O'Brien, R. (1998). An Overview of the Methodological Approach of Action Research. Retrieved from <http://www.web.ca/~robrien/papers/arfinal.html>

Oliveira, T. C. M. (2010). *User generated content audiovisual para instituições de ensino superior: proposta de boas práticas*. (Dissertação de Mestrado), Universidade de Aveiro. Retrieved from <https://ria.ua.pt/handle/10773/3805>

Sayão, L. F. (2010). Uma outra face dos metadados: informação para a gestão da preservação digital. *Encontros Bibli*, 15(30), 1-31. doi:<http://dx.doi.org/10.5007/1518-2924.2010v15n30p1>



PATRIMÔNIO: IDENTIDADES CULTURAIS NO SANTUÁRIO BOM JESUS DO MATOSINHOS EM CONGONHAS-MG

Cátia Rodrigues Barbosa¹, Luana Rodrigues Godinho Silveira²

¹Escola de Ciência da Informação da UFMG, 0000-0002-2684-504X, catiarbp1@hotmail.com

²Instituto dos Arquitetos do Brasil/MG, 0000-0002-9757-1808, luanargodinho@gmail.com

RESUMO Trata-se da segunda fase de uma pesquisa que visa criar as ontologias do patrimônio, mapeando as práticas culturais tradicionais e os termos usados sobre o patrimônio, pela comunidade da cidade de Congonhas. O objetivo dessa fase é analisar as diversas formas de apropriação do patrimônio da humanidade - O Santuário de Bom Jesus de Matosinhos. Compreender o surgimento da memória coletiva com base na interpretação de narrativas de identidade que se referem especialmente a relação entre o espaço urbano e as experiências das comunidades, no que tange ao espaço imaginário é de fundamental importância, uma vez que, encenado em museus e exacerbado no contexto da política de desenvolvimento do turismo e comunicações promocionais. A metodologia é qualitativa, um estudo de caso que constou de aplicação de questionário semi-aberto, após a visualização de uma vinheta, elaborada na primeira fase da pesquisa sobre O Santuário de Matosinhos. O questionário foi aplicado para os alunos do ensino fundamental e médio, da cidade de Congonhas. Essa fase da pesquisa teve início, no ano de 2016 e será finalizada em 2017, com uma perspectiva de uma amostragem significativa para elaboração das ontologias e proposta de modelo de gestão do conhecimento patrimonial. Os resultados alcançados, considerando a amostra da análise, 92% visitaram o Santuário de Matosinhos; 27% confirmaram que participaram de algum evento, a maioria citou a Semana Santa, o Jubileu e Celebrações de Missas como alguns dos eventos culturais que ocorreram no Santuário. 86% participaram de atividades lúdicas no local. Entre as atividades citadas, estão: rolar na grama, pique-pega, pique-esconde e correr. Deste modo, pretende-se analisar as diversas formas de apropriação do patrimônio, tendo em conta o contexto ecológico, político, econômico e social.

PALAVRAS-CHAVE *Patrimônio, espaço urbano, gestão do conhecimento, museu.*

ABSTRACT This is the second phase of a research that aims to create heritage ontologies, mapping the traditional cultural practices and terms used on heritage, by the community of the city of Congonhas. The purpose of this phase is to analyze the various forms of appropriation of the patrimony of humanity - The Sanctuary of Bom Jesus de Matosinhos. Understanding the emergence of collective memory based on the interpretation of identity narratives that refer especially to the relationship between urban space and the experiences of communities, with regard to imaginary space is of fundamental importance, since it is staged in museums and Exacerbated in the context of tourism development policy and promotional communications. The methodology is qualitative, a case study that consisted of application of a semi-open questionnaire, after the visualization of a vignette, elaborated in the first phase of the research on The Sanctuary of Matosinhos. The questionnaire was applied to elementary and middle school students in the city of Congonhas. This phase of the research began in 2016 and will be finalized in 2017, with a perspective of a significant sampling for elaboration of the ontologies and proposal of patrimonial knowledge management model. The results obtained, considering the sample of the analysis, 92% visited the Sanctuary of Matosinhos; 27% confirmed that they attended some event, most cited Holy Week, Jubilee and Mass Celebrations as some of the cultural events that took place in the Shrine. 86% participated in on-site recreational activities. Among the activities mentioned are: rolling in the grass, picking up, picketing and running. In this way, we intend to

analyze the different forms of appropriation of the patrimony, taking into account the ecological, political, economic and social context.

KEYWORDS *Heritage, urban space, knowledge management, museum.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

O presente trabalho apresenta os resultados da segunda fase de uma pesquisa que visa criar ontologias do patrimônio sobre o município de Congonhas/MG. Conforme define Moraes (2007), a ontologia pode ser interpretada como um conjunto de especificações formais e explícitas de uma conceitualização compartilhada. Nessa fase da pesquisa são avaliadas as forma de apropriação pela comunidade local do Santuário do Bom Jesus do Matosinhos.

O município de Congonhas está localizado na Região Central de planejamento do Estado de Minas Gerais, conforme definido pelo Governo Estadual. Segundo dados do Instituto Brasileiro de Geografia e Estatística (IBGE), o município possui cerca de 53.348 pessoas em 2016. O salário médio mensal para os trabalhadores formais é de 3,1 salários mínimos, sendo que cerca de 32,8% da população possuía ocupação em 2014.

A origem do município de Congonhas remonta à inauguração do Santuário do Bom Jesus do Matosinhos. (CONGONHAS, 2011). O santuário foi construído na segunda metade do Século XVIII, é uma obra barroca com relevância mundial, tendo sido tombado em 1939 pelo Instituto do Patrimônio Histórico e Artístico Nacional (IPHAN) e reconhecido como Patrimônio Cultural Mundial pela Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) em 1985. O santuário é um conjunto edificado composto por uma igreja, adro com muro e uma escadaria externa decorada com 12 estátuas de profetas em pedra sabão. Há ainda, após a escadaria, na parte frontal do templo, em terreno com aclive, seis capelas, chamadas de Passos, dispostas lado a lado, contendo esculturas de madeira em tamanho natural, ilustrando a via crucis de Jesus Cristo. (IPHAN, 2017).



Figura 1. Santuário de Bom Jesus de Matosinhos- Cidade de Congonhas

Fonte: https://pt.wikipedia.org/wiki/Doze_profetas_de_Aleijadinho. Consulta 27/03/2017.

As estátuas dos doze profetas (1800-1805) são impressionantes e notáveis; consideradas no seu conjunto como uma das mais relevantes obras de toda a produção de Antônio Francisco Lisboa (1738-1814), apelidado por Aleijadinho. Todas as estátuas em pedra-sabão de tamanho próximo do natural. A entrada superior do adro é composta pelas esculturas de Daniel e Oséias, no topo da escadaria.

A fundação da igreja do Bom Jesus de Matosinhos foi inspirada nos santuários de Bom Jesus de Matosinhos, norte de Portugal, na cidade do Porto e o Bom Jesus de Braga. Dispostas em duas alas que dá acesso ao Santuário Bom Jesus do Matosinhos, em Congonhas, há um conjunto arquitetônico de capelas dos Passos da Paixão.

O nome Congonhas remete a um certo tipo de arbusto que cobria a região, com a exploração do ouro, o povoado expandiu-se, com aventureiros procedentes de outras regiões da colônia e mesmo de Portugal. Em 1734 foi criada a freguesia de Nossa Senhora de Conceição das Congonhas.

A decoração do Santuário de Congonhas é uma relíquia do Rococó mineiro. O Rococó é um estilo artístico, que surgiu na Europa no século XVIII. Trata-se de um estilo que busca a sutileza dos excessos e suntuosidades do Barroco. O Rococó ao chegar no Brasil passou a ter influências de temas religiosos, podendo ser vistos nas cidades históricas de Minas Gerais e demais estados do Brasil. Antônio Francisco Lisboa foi um dos principais representantes do Rococó no Brasil.

O adro dos profetas, local que abriga a série de profetas de Congonhas, é de uma beleza harmoniosa, onde conteúdo e forma integram o espaço urbano. Os rolos de pergaminhos, em pedra sabão, contendo as citações bíblicas, em latim; desvela uma cumplicidade de temporalidade com a arte cristã.

O fato de Congonhas ser reconhecida como Patrimônio da Humanidade, implica em avanços culturais, educacionais, sociais e políticos para a cidade. Envolve questões de sustentabilidade econômica para a preservação dos bens patrimoniais e a atração de doadores privados, aumento do turismo e potencial intercâmbio de ideias sobre educação patrimonial.

Questões sobre o vandalismo, mineração, a gestão da cidade pensa na. A opinião pública é dividida, quando se trata de pensar substituir os profetas originais por réplicas possibilidade de retirada dos profetas para ocupar as galerias do museu de Congonhas. A intenção é proteger as obras.



Figura 2. Santuário de Bom Jesus de Matosinhos- Cidade de Congonhas

Fonte: https://pt.wikipedia.org/wiki/Doze_profetas_de_Aleijadinho. Consulta 27/03/2017

O santuário tem papel relevante na conformação urbana da cidade de Congonhas e se coloca na paisagem urbana da cidade como um importante espaço livre de uso público com relevância artística, histórica, cultural, religiosa, dentre outras. Neste sentido, parte-se da compreensão que o santuário deve ser entendido como patrimônio ambiental urbano. Conforme define Castriota (2009):

Assim, quando se pensa em termos de patrimônio ambiental urbano, não se pensa apenas na edificação, no monumento isolado, testemunho de um momento singular do passado, mas torna-se necessário, antes de mais nada, perceber as relações que os bens naturais e culturais apresentam entre si, e como o meio ambiente é fruto dessas relações. (...) Abordar o patrimônio ambiental urbano vai ser assim, como se pode perceber, muito mais do que simplesmente tomar determinadas edificações ou conjuntos: é antes conservar o equilíbrio da paisagem, pensando sempre como inter-relacionados a infra-estrutura, o lote, a edificação, a linguagem urbana, os usos, o perfil hitórico e a própria paisagem natural. (CASTRIOTA, 2009, p. 160).

Compreender o surgimento da memória coletiva com base na interpretação de narrativas de identidade que se referem especialmente à relação entre o espaço urbano e o uso desse espaço pela população, no que tange ao imaginário, é de fundamental importância; uma vez que, é encenado em museus e exacerbado no contexto da política de desenvolvimento do turismo e comunicações promocionais.

Costa (2012) complementa:

A Museologia pode colaborar para que os museus estabeleçam vínculos entre o que o público diversificado sabe e o que os pesquisadores científicos, artistas, literatos, músicos, entre outros, estão produzindo. O Patrimônio Cultural, quando bem preservado, divulgado e tornado efetivamente conhecido e apreendido pela população, garante a adoção de, ao menos, duas atitudes pró-ativas: a) conservação dos bens de valor cultural, em todos os níveis da tipologia patrimonial; b) aquisição de saúde cultural pelos participantes envolvidos na tarefa de salvaguardar e preservar os bens culturais, os quais possibilitam a formação da identidade cultural e a elevação da autoestima e do compromisso engajador e solidário. (COSTA, 2012, p.100).

Nesse sentido, cabe abordar sobre o Museu de Congonhas, situado ao lado do Santuário Bom Jesus de Matosinhos, inaugurado em dezembro de 2015. O prédio foi construído para ser museu, o projeto foi do arquiteto Gustavo Penna. O museu possui um acervo de arte sacra- barroca, que por meio do uso de tecnologias, apresentam exposições às quais mesclam o que há ao ar livre, no espaço urbano do Santuário Bom Jesus do Matosinhos e registros das festividades, depoimentos de moradores da região, jogos interativos em dispositivos tecnológicos, envolvendo os profetas; audios, vídeos sobre o Santuário.

Sendo assim, realizar o diagnóstico das diversas formas de apropriação do Santuário de Bom Jesus de Matosinhos, patrimônio da humanidade, é aplicar ao processo de gestão do patrimônio, uma visão crítica sobre a patrimonialização, perceber em que esses conhecimentos poderiam favorecer na elaboração de um modelo de gestão patrimonial, no sentido de relacionar o espaço ambiental - urbano com o espaço museológico.

Nesse contexto, a percepção do patrimônio pela própria comunidade, vista como valor e como fator que garanta a sustentabilidade da comunidade, implica em uma abordagem de gestão participativa, do patrimônio cultural. Os bens culturais, por meio de seu reconhecimento pela comunidade se tornam, cada vez mais relevantes, no sentido de exercerem suas funções econômicas e sociais.

A pesquisa apresenta uma análise das formas de apropriação do santuário pela comunidade local, elaborada a partir da aplicação de questionários semiaberto. Como resultado, pôde-se perceber se tratar

de um patrimônio ambiental urbano vivo, não apenas no imaginário, mas na vida cotidiana da comunidade local.

METODOLOGIA

Trata-se de um estudo de caso que utiliza método qualitativo. Avalia-se as formas de apropriação, bem como as narrativas acerca do Santuário de Bom Jesus do Matosinhos, em Congonhas, por sua comunidade local.

Para isso fora aplicado um questionário semiaberto para a amostra selecionada, em uma escola estadual, da região metropolitana de Congonhas, que contempla alunos do ensino fundamental e médio da cidade.

A amostra selecionada considerou os dados disponibilizados pelo IBGE, referente ao ano de 2010, importante salientar que a variação dos dados entre 2010 e 2017 não altera o objetivo apresentado nessa pesquisa; que Congonhas apresenta uma taxa de escolarização de pessoas com idade entre 6 e 14 anos de 99,1%. Considerando um total de 48.519 habitantes em 2010, sendo que destes, cerca de 6.899 estão na faixa etária de 6 a 14 anos, é possível dizer que cerca de 14% da população do município frequentam o ensino fundamental.

Em Congonhas há cerca de 21 escolas públicas de ensino fundamental, conforme dados do Índice de Desenvolvimento da Educação Básica (IDEB)

A amostra considerou também, o universo de estudantes do ensino médio das escolas públicas de Congonhas. Das 4 escolas de ensino médio segundo dados do Exame Nacional do Ensino Médio (ENEM), as duas públicas são as de maior porte e quantidade de alunos inscritos, com um total equivalente à 443 pessoas em relação aos 46 alunos da rede particular. As do ensino médio da rede pública de Congonhas.

Antes da aplicação do questionário, para uma melhor inteiração com o tema da pesquisa, fora apresentado aos estudantes uma vinheta elaborada na primeira fase da pesquisa.

A vinheta fez parte de um projeto de iniciação científica intitulado: patrimônio cultural e sua relação com a Ciência da Informação- Imagem em movimento, envolveu estudantes-bolsistas do curso de Museologia da Universidade Federal de Minas Gerais. O vídeo consta de quatro vinhetas, sendo uma delas sobre o Santuário de Matosinhos da cidade de Congonhas.

Elas podem ser acessadas pelo site: <https://www.youtube.com/watch?v=8UR1VsfNDEU>

As perguntas do questionário contemplaram questões sobre o conhecimento do Santuário de Bom Jesus do Matosinhos como Patrimônio da humanidade, a participação em algum evento cultural no Santuário de Bom Jesus do Matosinhos, o termo Santuário de Bom Jesus do Matosinhos, brincadeiras no entorno dos espaços abertos do Santuário de Bom Jesus do Matosinhos.

As respostas foram sistematizadas em tabelas e gráficos, de onde pôde-se extrair as análises.

A pesquisa teve início no ano de 2016 e se encontra em desenvolvimento, devendo ser finalizada em 2017, após a ampliação da amostra de forma a viabilizar a elaboração de ontologias e uma proposta de modelo de gestão do conhecimento patrimonial.

A organização do conhecimento gerada, por meio das respostas do questionário, ou seja, por meio das análises, implicará na terceira fase da pesquisa. Essa fase tratará da caracterização do domínio Patrimônio Santuário Bom Jesus do Matosinhos, que será definido por meio de uma linguagem, proporcionando interpretação.

Nesse sentido, a pesquisa ficará no âmbito de uma ontologia de domínio, à qual tem como objetivo criar um conjunto de classes, que cobre o domínio Patrimônio. Será necessário cobrir o domínio com localização, dados georreferenciados, termos usados pela comunidade local a respeito do Santuário Bom Jesus do Matosinhos.

Trata-se de categoria de instrumentos de representação que possibilita o processamento automático dos recursos informacionais representados, no domínio do patrimônio cultural e a realização de inferências.

RESULTADOS

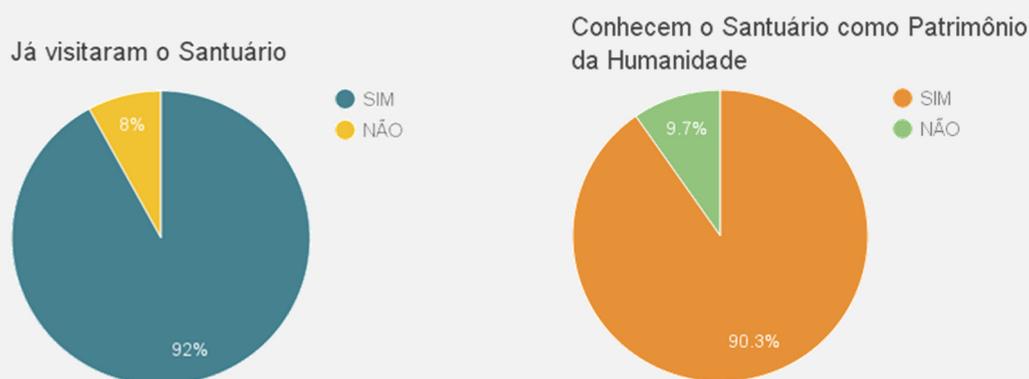
Preliminarmente, considerando a amostra atual da pesquisa, pôde-se verificar que:

92% dos que responderam ao questionário visitaram o Santuário Bom Jesus do Matosinhos;

86% participaram de atividades lúdicas no local.

27% confirmaram que participaram de algum evento, a maioria citou a Semana Santa, o Jubileu e Celebrações de Missas como alguns dos eventos culturais que ocorreram no Santuário.

80% disseram que conheciam o local com o nome de Santuário Bom Jesus do Matosinhos;
20% conheceram com o nome de Igreja.



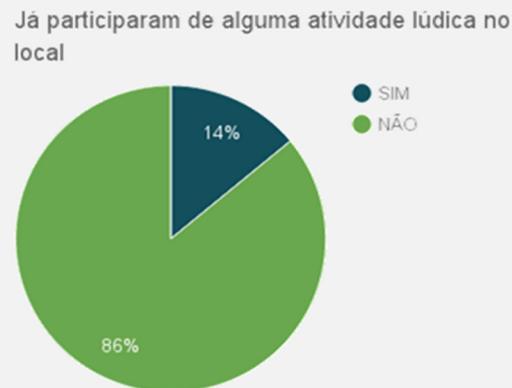


Figura 3. Análise da amostra que compõe de 68 alunos do Ensino fundamental e médio

Entre as atividades lúdicas citadas, estão: rolar na grama, pique-pega, pique-esconde e correr. Pelo teor destas, concluímos que a relação entre essa amostra caracterizada por jovens de 15 aos 18 anos foi em sua minoria estabelecida na infância para 8,68% dos participantes. O que indica que a utilização desse espaço urbano, turístico, Patrimônio da Humanidade é valorado pela escola, para desenvolver atividades de educação patrimonial e, provavelmente, pouco utilizado pelas famílias como um espaço de lazer, de atividades lúdicas.

Dos resultados alcançados foi possível aferir que o Santuário é amplamente conhecido por parte dos alunos que responderam ao questionário, evidenciando algum tipo de relação entre eles e o bem. Essa utilização pode ser considerada como de maior relevância para a vitalidade do espaço urbano e da conservação do bem tombado.

Importante salientar que o fato de 20% conhecerem o Santuário com o nome de Igreja, é significativo para a elaboração das ontologias do Patrimônio de Congonhas.

Conforme consta no gráfico acima, 90,3% responderam que tinham conhecimento do título do Santuário como Patrimônio da Humanidade; o que podemos inferir que a relação do monumento santuário com o turismo e a necessidade de preservação por parte dos alunos, sendo estes, moradores de Congonhas é estreita, no sentido de ter como identidade esse espaço urbano, como algo de um valor reconhecido mundialmente. Não foi questionado sobre os critérios da UNESCO para definir como Patrimônio da Humanidade. Entretanto havia um senso comum, a respeito desse título.

Deve-se ter uma atenção voltada para as respostas referentes à participação em algum evento cultural, no entorno do Santuário Bom Jesus de Matosinhos. Mesmo ciente que apenas 27% confirmaram a participação em algum evento cultural, no entorno do Santuário. A atenção deve estar voltada, no que eles elencaram como evento cultural: celebrações de missas, a semana santa que há uma encenação teatral, em frente ao adro do santuário.

O site da Câmara Municipal de Congonhas apresenta como festas, no calendário oficial da cidade: o jubileu do Senhor Bom Jesus de Matosinhos de Congonhas, que é uma comemoração religiosa; Semana Santa como uma outra significativa festa; semana do Aleijadinho.

O Jubileu do Senhor Bom Jesus de Matosinhos: trata-se de uma celebração que teve início no século XVIII. Caracteriza-se pela presença de vários romeiros, de todas as partes do país. Eles fazem pedidos, agradecem as graças alcançadas. No período, centenas de barraqueiros vendem artigos religiosos.

A Semana Santa conta com apresentações cênicas da Paixão e Morte de Jesus. A montagem dessa encenação envolve centenas de figurantes. O adro dos Profetas torna-se palco. A encenação atrai não somente expectadores locais, mas também, turistas de diversos países.

O Museu de Congonhas apresenta nos seus diversos vídeos, tótems digitais as festas da cidade, no entorno do Santuário, caracterizando-as como celebrações culturais. Todos esses dados serão analisados para a elaboração das ontologias do patrimônio e mostrará parâmetros para a criação de um modelo de gestão do conhecimento patrimonial.

A gestão do conhecimento patrimonial cultural é fundamental, no que tange à salvaguarda desse patrimônio e das informações sobre esse patrimônio. O compartilhamento de informações, a fácil recuperação da informação sobre o patrimônio Santuário Bom Jesus de Matosinhos, por sua comunidade local, deve ser considerada, no processo da pesquisa.

A disseminação da informação sobre a preservação do patrimônio deve ser um meio de se compreender e transformar a realidade. A qualificação da cultura em um processo interativo de construção do conhecimento sobre as ações voltadas para a preservação, a salvaguarda do patrimônio implica em uma gestão participativa.

CONCLUSÕES

O uso efetivo do local por parte dos alunos demonstrou o papel da função social do espaço urbano, enquanto memória viva e construída cotidianamente por todas as gerações, moradores do seu entorno, representando a potência afetiva destes símbolos.

Com base nesse resultado, podemos analisar as diversas maneiras de apropriação do patrimônio da região pelos jovens estudantes.

A preservação do patrimônio ambiental urbano de Congonhas é fruto das relações da comunidade para com esse patrimônio. O uso das edificações por parte da comunidade, do entorno como espaço lúdico, do conhecimento do Santuário Bom Jesus do Matosinhos como patrimônio da humanidade são aspectos relacionais significativos entre a comunidade e o patrimônio tombado que geram sentimento de pertencimento, de valor.

A utilização do espaço público pela comunidade é, de certa forma, um fator de validação do próprio patrimônio.

A pesquisa apontou indícios da intrínseca relação entre o papel da educação patrimonial, o papel da escola, na sua função de ensino-aprendizagem e a cultura local de Congonhas; o patrimônio da humanidade, o Santuário Bom Jesus de Matosinhos, o papel social do Museu de Congonhas e sobretudo um reconhecimento de identidades culturais, que há por parte dos moradores, no que tange ao Santuário e ao seu entorno.

Realizar o diagnóstico das diversas formas de apropriação do patrimônio da humanidade- O Santuário de Bom Jesus de Matosinhos é aplicar ao processo de gestão do patrimônio, uma visão crítica sobre a patrimonialização, perceber em que esses conhecimentos poderiam favorecer na elaboração de um modelo de gestão patrimonial, no sentido de relacionar o espaço ambiental- urbano com o espaço museológico.

A ideia de diagnóstico está associada a manter sua credibilidade junto ao público, que está intimamente ligada a vocabulários controlados para a criação de ontologias do patrimônio. O que significa e se propõe, ampliar o acesso ao conhecimento sobre o patrimônio da cidade de Congonhas e seu entorno.

Importante salientar que o diagnóstico apresentado torna-se parte indissociável de políticas de desenvolvimento e de defesa da integridades nacional, especificamente da formulação e execução de políticas culturais, muitas vezes criadas como instrumentos de preservação do patrimônio, preservação da identidade.

Para que um patrimônio seja preservado, é preciso, além dos atos jurídicos de tombamento, um reconhecimento das comunidades para com o seu patrimônio, o que é valor para os moradores. Assim, a implantação de um projeto sobre a preservação, recuperação, revitalização de um patrimônio deverá possuir um interesse social, local. Para que isso ocorra de forma cada vez mais democrática, o acesso à informação, no que tange ao mapeando das práticas culturais tradicionais e os termos usados sobre o patrimônio, pela comunidade da cidade de Congonhas é de fundamental relevância.

Um dos desafios a ser considerado é tomar os pontos significativos apontados pelos entrevistados como indicadores para a validação da criação das ontologias. Nesse sentido, consideramos que os nossos problemas podem estar no campo dos desafios tecnológicos, desenvolver um protótipo com especialistas da área da computação e que pesquisam na Ciência da Informação. Nossos problemas podem estar também no campo político- educacional, no que tange á identidade cultural. Portanto, a aplicação da Ontologia do Patrimônio, deve estar embasada na relação entre a teoria e a prática.

A preservação da identidade é necessária, devendo ser o suporte essencial para a criação do vocabulário controlado, no que se refere ao Patrimônio da Humanidade, o Santuário Bom Jesus de Matosinhos.

REFERÊNCIAS BIBLIOGRÁFICAS

Azevedo, F. L., Pires, J. R. F., Catão, L. P. (Orgs). (2009). *Cidadania, Memória e Patrimônio: as dimensões do museu no cenário atual*. Belo Horizonte: Crisálida.

Barbosa, C. R. (2010). *La muséologie à l'épreuve de La phénoménologie: La mise en scène des musées scientifiques et techniques*. Editions Universitaires Europeenes.

Brasil Enem por Escola. (2016). Recuperado em 16 de junho, 2016 de <http://www.brasil.gov.br/educacao/2016/10/inep-divulga-resultados-por-escola-do-enem-2015>

Brasil Ideb. Resultados e metas. (2017). Recuperado em 16 de junho, 2017 de <http://ideb.inep.gov.br/resultado/resultado/resultado.seam?cid=126048>.

Castriota, L.B. (2009). *Patrimônio Cultural: Conceitos, Políticas, Instrumentos*. Belo Horizonte: IEDS

Choay, F. (2006). *A alegoria do patrimônio.*, Trad.Machado,L.V. São Paulo: Estação Liberdade: UNESP.

Congonhas. História. (2011). Recuperado em 16 de junho, 2017 de <http://www.congonhas.mg.gov.br/detalhe-da-materia/info/historia/6495>

Costa, H.H.F.G. (2012). *Museologia e patrimônio nas cidades contemporâneas: uma tese sobre gestão de cidades sob a ótica da preservação da cultura e da memória*. Bol. Mus. Para. Emílio Goeldi. Cienc. Hum., Belém, v. 7, n. 1, p. 87-101, jan.-abr. 2012. Recuperado em 17 de agosto, 2017 de <http://www.scielo.br/pdf/bgoeldi/v7n1/a07v7n1.pdf>.

Fontanelli, S.A. (2005). *Centro de memória e ciência da informação: uma interação Necessária*. (Monografia, TCC). Escola de Comunicação e Artes. São Paulo: USP.

Gonçalves, J.R.S. (2009). O Patrimônio como categoria de pensamento. In: Abreu, R & Chagas, M. (Orgs.), *Memória e patrimônio: ensaios contemporâneos* (p.p.25-33). 2ª ed. Rio de Janeiro: Lamparina.

Instituto Brasileiro de Geografia e Estatística. (2016). Congonhas. Recuperado em 16 de junho de <http://www.ibge.gov.br/home/>

Instituto do Patrimônio Histórico e Artístico Nacional. (2017). *Santuário do Bom Jesus de Matozinhos - Congonhas (MG)*. Recuperado em 27 de março de <http://portal.iphan.gov.br/pagina/detalhes/46>

Jaques, P. B. (2003). Patrimônio Cultural Urbano: Espetáculo Contemporâneo? In: Revista de Urbanismo e Arquitetura, v. 6, n. 1.

Poulot, D. (2009). Uma História do patrimônio no ocidente. São Paulo: Estação Liberdade.



AVALIAÇÃO DOS MODELOS CLÁSSICOS DE RECUPERAÇÃO DA INFORMAÇÃO EMPREGADOS NA ÀREA DA SAÚDE

Josiane Mello¹, Angel Freddy Godoy Viera²

¹Doutoranda do Programa de Pós-Graduação em Ciência da Informação – UFSC, mellojosi@hotmail.com

²Professor Doutor do Programa de Pós-Graduação em Ciência da Informação – UFSC, Líder do Grupo de Pesquisa RITA – Recuperação da Informação e Tecnologias Avançadas, a.godoy@ufsc.br

RESUMO O presente artigo tem por objetivo avaliar o emprego dos modelos clássicos de recuperação da informação na área da saúde, por meio da publicação científica. Para alcance de tal objetivo, realizou-se um levantamento bibliográfico, no portal de periódicos da CAPES, em bases de dados da área da saúde. Trata-se de uma pesquisa exploratória-descritiva e bibliográfica, elaborada a partir de material já publicado e disponibilizado nas bases de pesquisa. Os resultados mostram que os modelos clássicos de recuperação da informação, se inserem de forma positiva na área estudada, contribuindo diretamente na vida da população, seja por meio de diagnósticos rápidos e precisos ou por métodos e técnicas inovadoras de prevenção e tratamento de enfermidades e etc. O modelo clássico Booleano pode ser empregado na Fonoaudiologia para classificar a gravidade de distúrbios articulatorios, na Oncologia e entre outras especialidades da área da saúde para acompanhar o tratamento de enfermidades, pela simplicidade e por se tratar de um modelo intuitivo com semântica precisa. Já o modelo clássico vetorial pode auxiliar no diagnóstico, prevenção, controle e tratamento de enfermidades. Por fim, o modelo probabilístico pode ser empregado na área da saúde, promovendo interação entre uma máquina e um sistema de informação de uma unidade médica e para auxiliar no processo de tomada de decisão médica. Recomenda-se que essa pesquisa seja ampliada, incluindo os modelos alternativos de RI e as tecnologias de máquinas de busca.

PALAVRAS-CHAVE *Recuperação da Informação, Sistemas de Recuperação da Informação, Modelos de Recuperação da Informação, Saúde.*

ABSTRACT The objective of this paper is to evaluate the use of classic models of information retrieval in the health area, through scientific publication. In order to reach this objective, a bibliographic survey was carried out, in the CAPES journals portal, in databases of the health area. It is an exploratory-descriptive and bibliographical research, elaborated from material already published and made available in the research bases. The results show that the classic models of information retrieval are inserted positively in the studied area, contributing directly to the population's life, either through rapid and precise diagnoses or innovative methods and techniques of prevention and treatment of diseases and etc. The classic Boolean model can be used in Speech Therapy to classify the severity of articulatory disorders, in Oncology and among other specialties of the health area to accompany the treatment of diseases, for simplicity and for being an intuitive model with precise semantics. The classic vector model can help in the diagnosis, prevention, control and treatment of diseases. Finally, the probabilistic model can be used in the health area, promoting interaction between a machine and an information system of a medical unit and to assist in the medical decision-making process. It is recommended that this research be expanded, including alternative models of RI and search engine technologies.

KEYWORDS *Information Retrieval, Information Retrieval Systems, Information Retrieval Models, Health.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A explosão informacional ocasionou vários problemas relacionados ao gerenciamento da informação, conforme descrito por Bush, em 1945, no artigo intitulado “As We May Think”. Segundo este autor, uma possível solução para tais problemas seria a automatização dos processos de recuperação da informação. Para tanto, ele chegou a imaginar um mecanismo específico, denominado de Memex (Memory Extension), que agregava as mais modernas tecnologias de informação existentes na época.

Na década seguinte, Mooers propôs a criação da disciplina de Recuperação da Informação, visando resolver os problemas oriundos do crescimento desenfreado do número de documentos produzidos e o acesso a estes.

Hoje, ainda enfrentam-se tais problemas e em uma escala muito maior, e, como Bush, busca-se na tecnologia a solução para tais problemas (FERNEDA, 2012).

Nos últimos anos, a área de Recuperação da Informação se desenvolveu significativamente, seus objetivos iniciais de indexação de textos e de busca por documentos úteis foram ampliados, incorporando, na atualidade, a pesquisa em modelagem, classificação de textos, arquitetura de sistemas, interfaces de usuário, filtragem, processamento de linguagem natural e entre outros.

Em se tratando de pesquisa, de acordo com Baeza-Yates e Ribeiro-Neto (2013) e Xie (2008), a referida área está contemplando estudos com dois enfoques distintos e complementares: um centrado no computador e outro centrado no usuário. Nos estudos com o enfoque centrado no computador, a recuperação da informação versa principalmente na construção de índices eficientes, no processamento de consultas com alto desempenho e no desenvolvimento de algoritmos de ranqueamento, objetivando melhorar os resultados. No que tange aos estudos com enfoque centrado no usuário, a recuperação da informação versa principalmente em investigar o comportamento do usuário, entender suas principais necessidades e determinar como esse entendimento afeta a organização e a operação do sistema de recuperação da informação.

No caso desta pesquisa, trata-se de um estudo centrado no computador, com o objetivo de investigar como os Modelos Clássicos de Recuperação da Informação são empregados na área da saúde e quais as contribuições para a área supracitada?

A escolha pela área da saúde se deu a partir de dois motivos. O primeiro em função de a referida área ter se desenvolvido substancialmente, sobretudo, depois do advento das tecnologias da informação e comunicação e as utilizaram amplamente. E o segundo, a partir do entendimento do papel e urgência da informação neste espaço. Sabe-se que, nesta área, a necessidade de informação para a tomada de decisão se faz presente constantemente, e a exatidão, pertinência, precisão e o tempo que esta leva para chegar às mãos do tomador de decisão fazem toda a diferença, podendo ter implicações no controle e no agravamento de enfermidades, reduzir o número de óbitos, trazer métodos e técnicas inovadoras para realizações de exames, diagnósticos e tratamentos, além de possibilitar a construção de novos conhecimentos aos profissionais desta área.

RECUPERAÇÃO DA INFORMAÇÃO

A Recuperação da Informação, doravante RI é uma área vasta dentro da Ciência da Computação e Ciência da Informação, pode-se entender a RI como um processo no qual um usuário converte sua necessidade de informação numa expressão de busca que submete ao sistema e o mesmo responde com uma listagem de registros que são relevantes para essa necessidade de informação.

Conforme relatos de Baeza-Yates e Ribeiro-Neto (2013), a preocupação em organizar para posteriormente buscar e recuperar a informação não é algo recente, as primeiras bibliotecas foram criadas para atender este objetivo, entre 3000 e 2000 a.C, sendo também, as primeiras instituições a adotarem sistemas de RI para recuperar informações.

Sobre isso, Foresti e Godoy Viera (2016, p. 6), pontuam que,

tradicionalmente a Recuperação da Informação (RI) acontecia nos ambientes escolares e acadêmicos em bibliotecas, local em que o usuário atuava na sua ânsia por informação e conhecimento novo manipulando as vetustas fichas catalográficas em formato papel. Suas necessidades eram grandemente orientadas pelo que acontecia em sala de aula. Mas não eram apenas as bibliotecas que permitiam a busca de informação em suas coleções, os arquivos e museus há muito atendem pesquisadores, empresas e a sociedade em geral através de seus inventários de demais instrumentos de pesquisa.

Com o passar do tempo, para acompanhar as transformações paradigmáticas da sociedade, a RI foi ampliando sua abrangência, com a banalização dos computadores e o advento das tecnologias móveis, a RI em dispositivos móveis cresceu significativamente, conforme pesquisa recente do IBGE, abordada no estudo de Foresti e Godoy Viera (2016).

Concomitante com a ampliação da abrangência, o conceito da RI sofreu alterações, deixando de ser vista como sinônimo de busca de literatura, conforme definição de Lancaster (1978), incorporando aspectos lógicos, Linguísticos e tecnológicos em sua conceituação.

Na atualidade, a RI é definida, conforme Baeza-Yates; Ribeiro-Neto (2013, p.1), como uma área, que,

trata de representação, armazenamento, organização e acesso a itens de informação, como documentos, páginas *web*, catálogos online, registros estruturados e semiestruturados, objetos multimídia, etc. A representação e a organização dos itens de informação devem fornecer aos usuários facilidade de acesso às informações de seu interesse (BAEZA-YATES; RIBEIRO-NETO, 2013, p.1).

Devido ao grande volume de documentos não estruturados armazenados na internet, surgem vários problemas relacionados à recuperação da informação. Assim, de acordo com Sharma (2013), a referida disciplina necessita de métodos que recuperem a informação de maneira eficaz, de forma automatizada.

Para atingir esses objetivos, os sistemas de RI fazem uso de modelos de recuperação da informação, que são apresentados na seguinte seção.

MODELOS DE RECUPERAÇÃO DA INFORMAÇÃO

Com as demandas e desafios impostos pela sociedade hodierna, os usuários necessitam cada vez mais de informações de alto nível de complexidade, a fim de empregá-las em suas atividades laborais ou em benefício próprio, suprimindo uma necessidade informacional imediata.

Essa necessidade de informação é traduzida pelo usuário em uma consulta, ou em uma sequência de consultas, que é submetida ao sistema de RI, sob um elenco de palavras chave, ou termos indexados, visando a obtenção de uma lista de documentos úteis e relevantes para ele.

Na percepção de Baeza-Yates e Ribeiro-Neto (2013) um sistema de RI deve ser capaz de recuperar todos os documentos que são relevantes à necessidade informacional do usuário e, ao mesmo, recuperar o menor número possível de documentos irrelevantes.

No entanto, é utopia esperar que um sistema de RI atenda plenamente tal objetivo, pois conforme observado pelos autores supracitados, a relevância está ligada a um julgamento pessoal, que varia de acordo com a tarefa a ser realizada e com o contexto desta. A relevância pode mudar com o tempo, à medida que novas informações estão disponíveis; com o local, talvez a resposta mais relevante seja a mais próxima e ainda com o dispositivo, considerando que a melhor resposta pode ser aquela disponibilizada em um documento pequeno, sendo baixada e visualizada mais rápido.

Diante da impossibilidade de um sistema de RI fornecer apenas respostas relevantes a todos os usuários o tempo todo, foram incorporados a estes, os modelos de RI, visando aproximar a necessidade de informação do usuário com o resultado obtido na busca, por meio do emprego de algoritmos e premissas de relevância definida em cada modelo de RI.

Os primeiros modelos de RI propostos foram os clássicos: Booleano, Vetorial e Probabilístico. Nestes modelos, “cada documento é descrito por um conjunto de palavras-chave representativas – também chamadas de termos de indexação – que busca representar o assunto do documento e sumarizar seu conteúdo de forma significativa”. (SOUZA, 2006, p.166).

De forma sucinta, apresentam-se a funcionalidade e desempenho dos modelos supracitados: O Modelo booleano é baseado na teoria de conjuntos e na álgebra Booleana, conforme descrição de Silva, Santos e Ferneda (2013, p.29-30),

o conteúdo informacional dos documentos é representado por um conjunto de termos de indexação. As buscas são formuladas por meio de uma expressão booleana composta por termos ligados através dos operadores lógicos (AND, OR e NOT). O resultado de uma busca é composto por um conjunto de documentos cuja representação satisfaz às restrições lógicas da expressão de busca. (SILVA, SANTOS E FERNEDA, 2013, p. 29 - 30).

Baeza-Yates e Ribeiro-Neto (2013, p.31) ressaltam como vantagens e desvantagens do modelo Booleano:

[...] o formalismo claro por trás do modelo e sua simplicidade, com a adoção de pesos binários para os termos de indexação. As principais desvantagens são o fato de não existir um ranqueamento, o que pode levar à recuperação de muitos ou de poucos documentos. Além disso, a formulação de consultas Booleanas é inconveniente para a maior parte dos usuários. Atualmente, sabe-se que a ponderação dos termos de indexação pode levar a uma melhoria substancial na qualidade da recuperação [...]. (BAEZA-YATES e RIBEIRO-NETO, 2013, p.31).

O Modelo vetorial de acordo com Baeza-Yates e Ribeiro-Neto (2013) considera casamentos parciais do termo de consulta, por meio de atribuição de pesos não binários aos termos de indexação das consultas e dos documentos, no qual, são usados para computar o grau de similaridade entre cada documento armazenado no sistema e a consulta do usuário.

A lista de resultados dos documentos recuperados é exibida em ordem decrescente, de acordo com o grau de similaridade destes, de modo a fornecer uma resposta mais precisa ao usuário, quando comparado ao modelo Booleano, em função do ranqueamento de documentos.

O modelo vetorial apresenta as seguintes vantagens, conforme Baeza-Yates e Ribeiro-Neto (2013, p.47),

- Seu esquema de ponderação de termos melhora a qualidade da recuperação da informação;
- Sua estratégia de casamento parcial permite a recuperação de documentos que aproximam as condições da consulta;
- A fórmula do cosseno ordena os documentos de acordo com o seu grau de similaridade em relação à consulta;
- A normalização pelo tamanho do documento está naturalmente embutida no modelo.

Como desvantagem o modelo não considera as dependências entre os termos de indexação. Apesar disso, o modelo não deixa de ser apontado como um excelente recurso para o ranqueamento de coleções genéricas.

Já o modelo probabilístico, desenvolvido por Robertson e Sparck Jones em 1976, faz uso da teoria probabilística para resolver o problema da RI. Souza (2006, p.167) diz que,

nesse modelo, supõe-se que exista um conjunto ideal de documentos que satisfaz a cada uma das consultas ao sistema, e que este conjunto pode ser recuperado. Através de tentativa inicial com um conjunto de documentos (para a qual se podem utilizar técnicas de outros modelos, como o vetorial) e do feedback do usuário em sucessivas interações, buscas e aproximar cada vez mais deste conjunto ideal, por meio de análise dos documentos considerados pertinentes pelo usuário. O valor desse modelo está em considerar a interação contínua com o usuário como um caminho para refinar o resultado continuamente.

No que tange as vantagens e desvantagens deste modelo, Baeza-Yates e Ribeiro-Neto (2013) pontuam algumas, respectivamente, que teoricamente, trata-se de um ótimo modelo de RI, por ranquear os documentos de forma decrescente, pautado na probabilidade de serem relevantes, baseado na informação disponível ao sistema. Embora, na prática, isso não funciona tão bem, pois a relevância do documento é afetada por variáveis externas ao sistema. Como inconveniente, o modelo necessita que seja estimado a separação inicial dos documentos em conjuntos de relevantes e não relevante. O tamanho dos documentos não é normalizado e o método também não leva em consideração a frequência na qual um termo de indexação ocorre em um documento.

METODOLOGIA

Esta pesquisa caracteriza-se quanto aos objetivos como exploratória-descritiva e quanto aos procedimentos técnicos como bibliográfica.

De acordo com Gil (2008, p. 27), as pesquisas exploratórias:

[...] são desenvolvidas com o objetivo de proporcionar visão geral, do tipo aproximativo, acerca de determinado fato. Este tipo de pesquisa é realizado especialmente quando o tema escolhido é pouco explorado e torna-se difícil sobre ele formular hipóteses precisas e operacionalizáveis.

Já as pesquisas descritivas “[...] têm como objetivo primordial a descrição das características de determinada população ou fenômeno ou o estabelecimento de relações entre variáveis.” (GIL, 2008, p. 28).

Quanto aos procedimentos técnicos, define-se como uma pesquisa bibliográfica, realizada nas bases de dados da área da saúde: Biblioteca Virtual da Saúde (BVS) - BIREME, MEDLINE, LILACS, IBECs do portal de periódicos da CAPES (<http://www.capes.gov.br/>), com o objetivo de levantar estudos, que empregam os modelos clássicos de recuperação da informação no domínio da saúde.

Para tal, a cobertura temporal utilizada foi de cinco anos (2012-2016), tipo de documento: artigo, idiomas: inglês, espanhol e português, campo de busca: título, resumo e assunto; assunto principal: armazenamento e recuperação da informação, limite da pesquisa: humanos e texto completo disponível.

Realizaram-se busca avançada, com auxílio de operadores booleanos. A estratégia de busca empregada é exibida no quadro 1.

Quadro 1 – estratégias de busca utilizadas

Estratégia de busca	Operador booleano	Total de documentos recuperados	Documentos relevantes
“information retrieval” AND model AND (Boolean OR Probabilistic OR vector) AND (medical OR medicine OR health)	AND/OR	12	07

Fonte: Mello e Godoy Viera (2017).

A pesquisa recuperou 12 documentos, destes, fez-se a leitura e excluíram-se 05 estudos, pois não estavam alinhados com o escopo desta pesquisa. Destes, 03 abordavam a técnica de aprendizado de máquina, 01 abordava o modelo alternativo de RI – Redes Bayesianas e 01 abordava o processo de elaboração de ontologia para auxiliar na atividade de processamento da linguagem natural. Os 07 documentos relevantes, que compõem o corpus desta pesquisa, encontram-se referenciados ao final deste resumo alargado, intitulado como referência do corpus da pesquisa.

RESULTADOS

Constatou-se por meio da literatura especializada o emprego de todos os modelos clássicos de recuperação da informação na área da saúde, os estudos foram enumerados de 01 até 07, conforme exibição no quadro 2

O modelo clássico Booleano foi identificado nos estudos 2 e 4. No primeiro mencionado, para

classificar a gravidade do desvio fonológico (DF). Já o segundo abordava o emprego do modelo na construção de uma rede para acompanhamento do tratamento de câncer de pulmão (SCLC). Verifica-

se com base nestas aplicações, que o modelo clássico Booleano pode ser empregado na Fonoaudiologia para classificar a gravidade de distúrbios articulatorios, na Oncologia e entre outras especialidades da área da saúde para acompanhar o tratamento de enfermidades, pela simplicidade e por se tratar de um modelo intuitivo com semântica precisa.

Já o modelo clássico vetorial foi sinalizado nos estudos 03 e 07. Onde o primeiro mencionado identificou o emprego do modelo supramencionado com foco na recuperação de caso de atendimento similar, para auxiliar no cuidado clínico e na pesquisa. Enquanto o segundo estudo foi aplicado para descobrir relações e outros significados entre drogas e doenças, reconhecimento de entidade multi-nível (gene, fármaco, doença, sintoma) e extração de interação (medicamento-gene, doença genética) da literatura. Verifica-se neste caso, uma importante contribuição do modelo supramencionado na área da saúde, podendo este, auxiliar no diagnóstico, prevenção, controle e tratamento de enfermidades.

Por fim, o modelo probabilístico foi identificado nos estudos 01, 05, 06. O primeiro estudo detectou o emprego do modelo em questão para o desenvolvimento de um componente de comunicação entre máquinas de anestesia e o sistema de informação de um determinado hospital. O segundo estudo supramencionado foi empregado para calcular e interpretar o atraso mútuo da informação (TDMI) para uma tarefa complexa, diversa e população escassamente medida, possivelmente não-estacionárias de séries temporais de composição de origem desconhecida. O terceiro estudo apresenta o emprego do modelo em tela para precisar se um documento biomédico acessado no passado poderá ser de interesse de acesso no futuro. De modo que, essa métrica de desejabilidade fornece um meio para melhorar a classificação de documentos para sistemas recuperação da informação e uma base para o avanço de modelos de recuperação da informação Bayesianos. Verifica-se que esse modelo pode ser empregado na área da saúde, promovendo interação entre uma máquina e um sistema de informação de uma unidade médica e para auxiliar no processo de tomada de decisão médica.

Quadro 2 – Resumo da aplicação dos modelos e tecnologias de RI

Número do Estudo	Modelos clássicos de RI
01 González Martínez, Y. ; Castillo García, M. ; Quevedo, Y. S. (2014)	Fez uso do modelo probabilístico com o apoio de agentes inteligentes, uso de padrões de vídeos e codificação de vídeos, metadados e técnicas de processamento de conteúdo para o desenvolvimento de um componente de comunicação entre máquinas de anestesia e o sistema de informação de um determinado hospital.
02 Brancalioni, A. R. ; Magnago, K. F. ; Keske-soares, M. (2012)	Empregou o modelo de recuperação da informação booleano , utilizando a teoria de conjunto Fuzzy para classificar a gravidade do desvio fonológico (DF).
03 Moen H; Ginter F; Marsi E; Peltonen L-M; Salakoski T; Salanterä S. (2015)	Relata o emprego de vários métodos para a recuperação de informação, com foco na recuperação de caso de atendimento, com base na similaridade textual, onde a similaridade é medida através da modelagem específica do domínio da semântica distributiva das palavras. Dentre os modelos empregados, identificou-se o modelo vetorial , com variantes de indexação aleatória e o modelo de rede neural semântica word2vec.
04 Udyavar, A. R; Wooten, D. J; Hoeksema, M. D; Bansal, M.; Califano, A.; Estrada, L; Chnell, S. ; Irish, J.	O câncer de pulmão (SCLC) é uma doença devastadora devido à sua propensão para a invasão precoce e recaída refratária após a resposta ao tratamento inicial. A topologia desta rede de SCLC fatores de transcrição (FT) foi derivada de conhecimento prévio e simulada usando modelagem booleana . Estas simulações predisseram que a rede se instala em atratores, ou padrões de expressão de (FT), que se correlacionam com os fenótipos (NE) ou (ML), sugerindo que a dinâmica da rede (FT) subjaz ao surgimento de fenótipos heterogêneos de SCLC.

<p>M; Massion, P. P; Quaranta, V. (2016)</p>	
<p>05 Albersa, D. J. ; Hripcsakb, G. (2012)</p>	<p>O estudo apresenta o emprego do modelo probabilístico, para calcular e interpretar o atraso mútuo da informação (TDMI) para uma tarefa complexa, diversa e população escassamente medida, possivelmente não-estacionárias de séries temporais de composição desconhecida e de origem. O principal veículo utilizado para esta análise é uma comparação entre a informação mútua temporizada, em média, a população e o atraso de informação mútua de uma população de agregados (aqui, agregação implica que a população é em conjunto antes de quaisquer estimativas estatísticas são implementadas). Através do uso de ferramentas de informação teórica, uma sequência de cálculos praticamente exequível são detalhados que permitem para a média e a agregação temporizada informação mútua ser interpretado. Além disso, esses cálculos também podem ser usados para compreender o grau de homo ou heterogeneidade presente na população. Os métodos foram aplicados e demonstrados na série de tempo das medições de glicose de duas subpopulações diferentes dos indivíduos do repositório de registro eletrônico de saúde da University Medical Center of Columbia, revelando uma imagem da composição da população, bem como características fisiológicas.</p>
<p>06 Goodwin, J. C. ; Johnson, T. R.; Cohen, T. ; Herskovic, J. R; Bernstam, E. V. (2012)</p>	<p>O artigo apresentou o emprego do modelo probabilístico para precisar se o documento biomédico acessado no passado poderá ser de interesse de acesso no futuro. Para tal, os autores avaliaram dois modelos de acesso ao documento, baseados em Anderson e Schooler (1991) o primeiro modelo baseado na frequência com que um documento foi acessado. O segundo baseado na frequência e relevância . O estudo conclui que é possível modelar com precisão a probabilidade de um documento ser acessado no futuro com base em acessos anteriores. O emprego deste modelo torna-se relevante diante da sobrecarga de informação no domínio biomédico, o esforço significativo de pesquisa centrou-se nos sistemas de recuperação de informação (RI). A métrica de desajustabilidade fornece um meio para melhorar a classificação de documentos para sistemas IR e uma base para o avanço de modelos bayesianos RI, fornecendo uma estimativa de probabilidade anterior teoricamente motivada.</p>
<p>07 Lee, S. ; Choi, J. ; Park, K. ; Song, M.; Lee, D. (2012)</p>	<p>O estudo utiliza o modelo vetorial, denominado de Swanson ABC para inferir relações ocultas na literatura biológica. Trata-se de um modelo consistente, no entanto, o modelo é inadequado para inferir relações com informações de contexto. Além disso, o modelo gera uma grande quantidade de candidatos a partir de texto biológico, e é uma técnica semi-automática, de mão-de-obra intensiva, que exige a entrada manual do especialista humano. Para enfrentar esses problemas, foram incorporado termos de contexto para inferir as relações entre as interações AB e as interações do BC. Foram propostas 3 etapas para descobrir relacionamentos ocultos significativos entre drogas e doenças: 1) reconhecimento de entidade multi-nível (gene, fármaco, doença, sintoma), 2) extração de interação (medicamento-gene, doença genética) da literatura, 3) contexto cálculo de pontuação de similaridade baseada em vetores. Posteriormente, foi avaliado a hipótese com os conjuntos de dados da "Doença de Alzheimer" relacionados 77.711 resumos da PubMed. Como padrões de ouro, os bancos de dados PharmGKB e CTD foram usados. Os resultados indicam que a inferência de relação baseada em contexto alcançou uma melhor precisão do que a abordagem do modelo ABC anterior. A análise da literatura também mostra que as interações inferidas pela abordagem baseada no contexto são mais significativas do que as interações pelo modelo ABC anterior.</p>

Fonte: Mello e Godoy Viera (2017)

CONSIDERAÇÕES FINAIS

A saúde é uma área que tem se desenvolvido substancialmente, sobretudo, com o auxílio das tecnologias da informação e comunicação.

Identificou-se o emprego de todos os modelos clássicos de recuperação da informação no domínio da saúde.

Pela pesquisa desenvolvida foi possível constatar que tais modelos, quando aplicados na área supracitada, proporcionam vários benefícios para a vida da população: diagnósticos rápidos e precisos, métodos e técnicas inovadoras de prevenção e tratamento de enfermidades e etc.

O modelo clássico Booleano pode ser empregado na Fonoaudiologia para classificar a gravidade de distúrbios articulatorios, na Oncologia e entre outras especificidades da área da saúde para

acompanhar o tratamento de enfermidades, pela simplicidade e por se tratar de um modelo intuitivo com semântica precisa.

Já o modelo clássico vetorial pode auxiliar no diagnóstico, prevenção, controle e tratamento de enfermidades.

Por fim, o modelo probabilístico pode ser empregado na área da saúde, promovendo interação entre uma máquina e um sistema de informação de uma unidade médica e para auxiliar no processo de tomada de decisão médica.

Recomenda-se ampliar o escopo da pesquisa, incluindo os modelos alternativos de recuperação da informação e as tecnologias de máquina de busca no domínio da saúde.

REFERÊNCIAS BIBLIOGRÁFICAS DO CORPUS DA PESQUISA

ESTUDO 1

González Martínez, Y. ; Castillo García, M. ; Quevedo, Y. S. (2014). Componente para la lectura de datos por alas-HIS desde máquinas de anestesia. *RCIM [online]*, 6 (1), n.1, 110 - 117.

ESTUDO 2

Brançalioni, A. R. ; Magnago, K. F. ; Keske-soares, M. (2012). Validation of a fuzzy linguistic model to classify the severity of phonological disorder, *Revista CEFAC: Atualização Científica em Fonoaudiologia e Educação*, 14 (3), 448-458.

ESTUDO 3

Moen H; Ginter F; Marsi E; Peltonen L-M; Salakoski T; Salanterä S. (2015). Care episode retrieval: distributional semantic models for information retrieval in the clinical domain. *BMC Medical Informatics and Decision Making*, 15 (2), 1-19.

ESTUDO 4

Udyavar, A. R; Wooten, D. J; Hoeksema, M. D; Bansal, M.; Califano, A.; Estrada, L; Chnell, S. ; Irish, J. M; Massion, P. P; Quaranta, V. (2016). Novel hybrid phenotype revealed in small cell lung cancer by a transcription factor network model that can explain tumor heterogeneity. *Cancer Res.* 77(5), 1063-1074.

ESTUDO 5

Albersa, D. J. ; Hripcsakb, G. (2012). Using time-delayed mutual information to discover and interpret temporal correlation structure in complex populations. *AIP Chaos:an Interdisciplinare Journal of Nonlinear Science*, 22(1), 1-26.

ESTUDO 6

Goodwin, J. C. ; Johnson, T. R.; Cohen, T. ; Herskovic, J. R; Bernstam, E. V. (2012). Predicting biomedical document access as a function of past use. *J Am Med Inform Assoc*; 19(3), 473- 478.

ESTUDO 7

Lee, S. ; Choi, J. ; Park, K. ; Song, M.; Lee, D. (2012). Discovering context-specific relationships from biological literature by using multi-level context terms. *BMC Medical Informatics And Decision Making*, 12(1), 1-12.

REFERÊNCIAS BIBLIOGRÁFICAS

Baeza-Yates, R.; Ribeiro-Neto, B. (2013). *Recuperação de informação: conceitos e tecnologia das máquinas de busca*. 2. ed. Porto Alegre: Bookman.

Ferneda, Edberto. (2012). *A. Introdução aos modelos computacionais de recuperação de informação*. Rio de Janeiro: Ciência Moderna.

Foresti, F. ; Godoy Vieira, A. F. (2016). A Recuperação da informação em dispositivos móveis, *Biblionline*, João Pessoa, 12 (2), 6 – 25.

Gil, Antônio Carlos. *Métodos e técnicas de pesquisa social*. 6. ed. São Paulo: Atlas, 2008.

Lancaster, F. W. (1978). *Information retrieval systems: characteristics, testing and evaluation*. 2. ed. New York, NY: Wiley.

Sharma, A. (2013). Intelligent information retrieval system: a survey. *Advance in Electronic and Electric Engineering*. 3 (1).

Silva, R. E. ; Santos, P. L. V. A. C. ; FERNEDA, E. (2013). Modelos de recuperação de informação e web semântica: a questão da relevância. *Inf. Inf.*, Londrina, 18 (3), 27 – 44.

Souza, R. R. (2016). Sistemas de recuperação de informações e mecanismos de busca na web: panorama atual e tendências. *Perspectivas em Ciência da Informação*, Belo Horizonte, 11 (2), 161-173.

Xie, I. (2008). *Interactive information retrieval in digital environments*. New York: IGI Publishing.



REPRESENTAÇÃO DOCUMENTAL DE GRAFFITI COMO PATRIMÔNIO CULTURAL: PROPOSTA METODOLÓGICO-CONCEITUAL

Zaira Regina Zafalon¹, Fábio Rogério Batista Lima², Paula Regina Ventura Amorim Gonzalez³,
Paula Regina Dal'Evedove⁴

¹Universidade Federal de São Carlos, zaira@ufscar.br

²Universidade Estadual Paulista, fabio.robal@yahoo.com.br

³Universidade Estadual de Londrina, paulaventuramorim@gmail.com

⁴Universidade Federal de São Carlos, dalevedove@ufscar.br

RESUMO Questões acerca do patrimônio cultural têm sido amplamente discutidas tanto por conta da ampliação de conceitos relacionados ao patrimônio quanto aqueles vinculados à cultura. O *graffiti* é considerado uma forma de manifestação artística em espaços públicos, popularmente definidos como um tipo de inscrição feita em paredes. Em decorrência do aspecto efêmero do *graffiti* a pesquisa é apresentada com o objetivo de propor um modelo de representação documental de *graffiti* tendo em vista a garantia de sua consolidação enquanto patrimônio cultural, com suas características descritivas, culturais, espaço-temporais e ideológicas. Como método foram utilizados elementos previstos na análise semiótica aplicada (cores, linhas, superfícies, formas e luzes) e na análise facetada rangathaniana (personalidade, matéria, energia, espaço e tempo). O modelo proposto para a Representação Documental de *Grffiti* apresenta a representação imagética do *graffiti* em fotografias, o que viabiliza sua preservação como patrimônio cultural, e nove metadados, como elementos da representação documental: Título, Artista, Criação, Citação cultural, Citação espaço-temporal, Citação ideológica, Descrição, Resumo e Palavras-chave.

PALAVRAS-CHAVE *Grffiti, Patrimônio Cultural, Representação Documental, Representação Imagética.*

ABSTRACT Issues concerning cultural heritage have been widely discussed due to the broadening of concepts related to heritage as well as those related to culture. *Grffiti* is considered a form of artistic expression in public spaces, popularly defined as a type of inscription done on walls. Due to the ephemeral aspect of *graffiti*, the study is presented with the objective of proposing a model of documentary representation of *graffiti*, in order to guarantee its consolidation as a cultural heritage, with its descriptive, cultural, spatiotemporal and ideological characteristics. The elements used in the methodology were applied semiotic analysis (colors, lines, surfaces, shapes and lights) and Ranganathanian faceted analysis (personality, matter, energy, space and time). The proposed model for the Documentary Representation of *Grffiti* presents the image representation of the *graffiti* in photographs, which allows its preservation as a cultural heritage, and nine types of metadata, as elements of documentary representation: Title, Artist, Creation, Cultural citation, Spatiotemporal citation, Ideological citation, Description, Summary and Keywords.

KEYWORDS *Grffiti, Cultural Heritage, Documentary Representation, Image Representation.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A relação entre patrimônio e cultura há muito se discute, principalmente quando se considera a expressão *bens culturais*, haja vista o fato de que remete à materialidade, à permanência, à preservação.

Nesse universo, as questões conceituais relacionadas ao patrimônio cultural têm sido ampliadas. Se, em meados do século XX o patrimônio cultural relacionava-se às “[...] obras monumentais, obras de arte consagradas, propriedades de grande luxo associadas às classes dominantes, [...]”, tomou-se, depois, outra perspectiva, com aspectos além da materialidade, tais como “a música, a literatura, a dança” e o “legado cultural das minorias [...] contado através da história social”. (Corso, 2004, p. 30).

A Convenção sobre a Salvaguarda do Patrimônio Mundial, Cultural e Natural (UNESCO, 1972) considera como patrimônio cultural:

Os monumentos. – Obras arquitectónicas, de escultura ou de pintura monumentais, elementos de estruturas de carácter arqueológico, inscrições, grutas e grupos de elementos com valor universal excepcional do ponto de vista da história, da arte ou da ciência;

Os conjuntos. – Grupos de construções isoladas ou reunidos que, em virtude da sua arquitectura, unidade ou integração na paisagem têm valor universal excepcional do ponto de vista da história, da arte ou da ciência;

Os locais de interesse. – Obras do homem, ou obras conjugadas do homem e da natureza, e as zonas, incluindo os locais de interesse arqueológico, com um valor universal excepcional do ponto de vista histórico, estético, etnológico ou antropológico.

No mesmo documento está definida a obrigação do Estado em “assegurar a identificação, protecção, conservação, valorização e transmissão às gerações futuras do patrimônio cultural e natural [...]” (UNESCO, 1972). Tendo em vista que a Convenção indica que os Estados devem envidar esforços para “desenvolver os estudos e as pesquisas científicas e técnica e aperfeiçoar os métodos de intervenção que permitem a um Estado enfrentar os perigos que ameaçam o seu patrimônio cultural e natural” (UNESCO, 1972), apresenta-se a seguinte questão de pesquisa: como garantir a recuperação do *graffiti* enquanto patrimônio cultural, tendo em vista sua característica efêmera?

Apresenta-se, assim, como objetivo desta pesquisa propor um modelo de representação documental de *graffiti* tendo em vista a garantia de sua consolidação enquanto patrimônio cultural, com suas características descritivas, culturais, espaço-temporais e ideológicas.

Interessa-se pelo *graffiti* enquanto documento objeto desta pesquisa visto que, ao longo da história da humanidade, a arte esteve sempre presente como forma de expressão, embora as técnicas de produção criadas pelo homem variassem bastante em épocas e lugares. “Nenhuma sociedade deixou de produzir arte, por mais baixo que tenha sido seu nível de existência material” (Farthing, 2011). Sob a influência de diferentes circunstâncias, a arte e toda a sociedade sofreram e ainda sofrem transformações, pois estão interligadas (Farthing, 2011; Lima & Santos, 2014).

Pinturas rupestres, esculturas rudimentares, pinturas ornadas por diversos tipos de pigmentos orgânicos, o uso da tecnologia da informática para produção e reprodução de pinturas ‘eletroeletrônicas’, ‘imagens de terceira geração’ e também o *graffiti* (*street art*) levaram o homem comum e o artista a se sentirem instigados a registrar e, através de seus impulsos criativos, transformar em arte, formas e imagens, assentadas em diversos tipos de suportes (Plaza, 1993; Couchot, 2003; Lima, Santos & Francisco, 2016).

A origem do *graffiti* remonta à pré-história, das pinturas rupestres aos escritos e desenhos das civilizações antigas (Grécia, Roma, Egito, Pompéia, Monumentos Maias e Astecas, etc.), passando pelas inscrições corporais (escarificações e tatuagens) encontradas em diversas tribos indígenas (Riout et al., 1985). Sua etimologia deriva do Grego *graphein*, verbo que significa escrever, desenhar. O conceito moderno de *graffiti* vem do italiano *graffito* (no plural *graffiti*). Esse tipo de manifestação artística atravessou séculos, mostrou-se uma prática presente nas mais diferentes civilizações e vem ganhando novas formas, estilos e suportes (Lima, Santos & Francisco, 2016).

O *graffiti* é considerado uma forma de manifestação artística em espaços públicos, popularmente definidos como um tipo de inscrição feita em paredes. De acordo com a literatura e alguns teóricos da área, o *graffiti* surgiu no início da década de 1970, em Nova York, Estados Unidos (Cooper & Chalfant, 1984; Castleman, 1982; Lima, Santos & Francisco, 2016). O *graffiti* norte-americano, segundo Campos (2013), deriva da cultura *hip-hop* emergente no período, a qual correspondia à expressão visual de um movimento composto igualmente pela vertente musical (Dj e Mc) e de dança (*breakdance*). Este movimento é o veículo de comunicação usado para expor a realidade das ruas.

No Brasil, o *graffiti* começou a se fixar no final da década de 1970, em São Paulo. Os artistas brasileiros, não satisfeitos com o *graffiti* norte-americano, começaram a incrementar a arte com um toque brasileiro. Nos dias de hoje, o estilo do *graffiti* brasileiro é reconhecido entre os melhores do mundo. A figura 1 apresenta o painel *Gigante de Boston e Greenway*, de Os Gêmeos, pintado em Boston, a convite do Instituto de Arte Contemporânea. O local do mural é particularmente especial por estar situado diretamente em frente à principal estação ferroviária e do edifício da Reserva Federal, bem como na Greenway (Spinelli, 2007; Ventura, 2009; Campos, 2009; Lima, Santos & Francisco, 2016).



Figura 1. Gigante de Boston e Greenway (OS GÊMEOS, 2012).

Muitas polêmicas giram em torno desse movimento artístico, pois, de um lado, o *graffiti* é desempenhado com qualidade artística e, por outro lado, não passa de poluição visual e vandalismo. No Brasil, é comum distinguir *graffiti* de pichação, sendo que o primeiro está associado a uma prática de natureza artística, enquanto o segundo é, basicamente, entendido como uma expressão ilegal sem motivação estética (Spinelli, 2007; Ventura, 2009; Campos, 2009; Lima, Santos & Francisco, 2016).

A questão entre *graffiti* e pichação não está em definir o que é ou não arte. Isto se configuraria como uma arbitrariedade. Do ponto de vista do patrimônio compreende-se que danificá-lo é crime, mas a questão vai além, visto que os que vivem em centros urbanos modernos estão expostos tanto às

diversidades culturais, étnicas, de gênero, etc., quanto aos problemas de falta de regulamentação relacionada à depredação de patrimônios e ao empilhamento e massificações propagandistas de *outdoors*, faixas, cartazes, *banners*, com imagens de contradições sociais até comuns nas megalópoles (Lima, Santos & Francisco, 2016).

O *graffiti*, a pintura em mural, o estêncil, as *tags* (assinaturas), a pichação, o *skate*, o *hip-hop*, o *rap*, etc., são marcas culturais das metrópoles, que se configuram com situações limítrofes entre o mundo lícito e idealizado e o ilícito e brutal. A grafia do *graffiti* pode ter características que variam do protesto, da ironia, da frase de resistência político-cultural, à recuperação física e mental de uma identidade perdida (Lima, Santos & Francisco, 2016). De todo modo, o *graffiti* significa a maneira que alguns encontram para expressar sua indignação.

Compreende-se que a representação documental, por sua característica de voltar-se para fins de recuperação e acesso aos documentos, seja o processo que preveja mecanismos para que o *graffiti* possa ser garantido como patrimônio cultural. A representação documental configura-se como uma proposta para garantir a recuperação da informação e da consolidação do *graffiti* enquanto patrimônio cultural haja vista o fato de recorrer à representação imagética do *graffiti* (com a fotografia como representação de nível simples) e aos aspectos inerentes à própria representação documental (como representação de nível complexo, por favorecer a recuperação e o acesso aos documentos).

Nesse contexto, entende-se que a Semiótica na concepção do teórico e filósofo Charles Sanders Peirce (1839-1914), aplicada aos estudos de representação documental, favorece uma melhor sistematização do processo de significação e sentido do *graffiti* na qualidade de signo que representa o mundo concreto dos objetos visuais. Enquanto registro sígnico, o *graffiti* congrega informações intimamente relacionadas às questões socioculturais, o que o torna produto da cultura contemporânea. Assim, se materializa como uma representação visual, fruto de uma representação mental. Isto porque, “não há imagens como representações visuais que não tenham surgido de imagens na mente daqueles que as produziram, do mesmo modo que não há imagens mentais que não tenham origem no mundo concreto dos objetos visuais” (Santaella & Noth, 2005, p. 15).

Uma vez registrado em um suporte passível de acesso e recuperação, torna-se possível descrever a natureza sígnica do *graffiti* mediante aplicação de procedimentos de significação e interpretação, com fins de representação documental.

Nas palavras de Pato (2014, p. 488),

As características físicas dos documentos imagéticos guardam um primeiro nível de informação, tais como tipo, cor, dimensões. A indexação por esses atributos intrínsecos é baseada em conteúdo. No segundo nível, há um processo de “tradução” da substância intelectual dos suportes e da representação das mensagens por meio de signos. Essa indexação é produzida por humanos e baseada em conceitos [...].

Transposto para um documento fotográfico, as estratégias empregadas para a representação imagética podem ser consideradas em uma proposta metodológica própria para o *graffiti*. Na literatura especializada da área, a fotografia é fundamentada em estudos que oferecem técnicas de análise e de interpretação de imagens para fins documentais e que consideram o contexto de produção e recepção, elementos que vão além do significado manifesto (Panofsky, 1979; Smit, 1989; Smit, 1996; Manini, 2002; Manini, 2004). Sontag (1981) discute as múltiplas significações possíveis do documento fotográfico. Considera-se, nesse cenário, que as categorias rangianthianas ampliem as capacidades de

análise e observação dos elementos concretos, tal como observado por Costa (2008), que expandiu a perspectiva da análise de conteúdo de imagens ao congregar as categorias básicas da didática retórica (quem, como, que, onde, quando), empregadas para a análise documental de imagens, às cinco categorias essenciais de Ranganathan – personalidade, matéria, energia, espaço e tempo.

As semelhanças conceituais entre as categorias rangathanianas e retóricas viabilizam “[...] um alargamento teórico e uma indicação de parâmetros para se estabelecer procedimentos metodológicos e fazer análise de documento iconográfico ou imagético” (Costa, 2008, p. 99). Este recurso de ampliação categórica tem potencial de gerar uma representação documental em que as funções sociais do objeto ou ideia categorizada sejam contempladas, característica imprescindível para a atividade representacional do *graffiti*, enquanto signo humano produzido socialmente, reflexo de determinada cultura.

É importante destacar que o modelo aqui proposto reflete um direcionamento conceitual e metodológico que revisita as abordagens tradicionais da Ciência da Informação, que trabalham com a leitura e análise de documentos imagéticos, sendo oportuno pesquisas complementares que aprofundem o entendimento acerca das relações entre os signos semióticos de documentos não verbais, conforme proposto por Pato (2014).

Reconhece-se que as unidades de informação, merecidamente aquelas consagradas como instituições de património cultural, assumem-se como agentes sociais de comunicação. Assim, entende-se que a representação documental seja atividade fundamental para o estabelecimento do processo comunicativo entre tais unidades e o público a que se destina, pautando-se, para tanto, na conveniência do usuário.

Zafalon e Dal’Evedove (2016) afirmam que

A representação documental requer, portanto, conhecimento do público a que se destina, das especificidades do gênero e do formato dos documentos de que dispõe e dos procedimentos, metodologias e instrumentos a serem utilizados na representação dado que, por meio de elementos representativos do documento, é possível promover busca e recuperação dos mesmos de modo a manter o acesso garantido.

Desse modo, é por meio de uma proposta de representação documental, que recorre a todos os aspectos dos recursos informacionais, tanto descritivos quanto de conteúdo cultural, espaço-temporal e ideológico, que se torna possível a materialização e a imanência dos *graffiti*.

Garrido Arilla (1999) considera a representação documental como o processo de descrição de elementos que permitem identificar e estabelecer formas de recuperação de um recurso informacional. Assim, a descrição bibliográfica é a ação na qual se expõem elementos de unidades informativas que permitem reconhecer e particularizar formalmente um documento em referências específicas que aludem os dados externos do documento e se distinguem dos restantes. Zafalon (2012, p. 68) afirma que “representação da informação é, portanto, o ato de articular formas de descrição a partir de instrumentos que permitam tornar cognoscível um recurso informacional sem que seja necessário recorrer ao documento original para identificá-lo.”

METODOLOGIA

A presente pesquisa, de natureza aplicada, propõe um modelo de representação do *graffiti*, para que seja garantida sua recuperação como uma arte que compõe o patrimônio cultural brasileiro.

Nesse contexto, o desenvolvimento do trabalho consistiu em um estudo teórico exploratório e descritivo. O caráter exploratório e descritivo permitiu a análise do tema e a construção do conhecimento teórico. Segundo Köche (2003, p. 126), os estudos exploratórios desencadeiam “[...] um processo de investigação que identifica a natureza do fenômeno e apontam características essenciais das variáveis que se quer estudar”.

Para o desenvolvimento da pesquisa recorreu-se aos procedimentos de pesquisa bibliográfica e documental, visto que se explorou o trabalho publicado sobre o tema estudado e às características inerentes ao documento em análise. (Lakatos & Marconi, 2001).

Na perspectiva de abordagem do problema, a pesquisa é qualitativa, pois trabalha com observações, descrições e interpretações que, segundo Gerhardt e Silveira (2009), ao se utilizar o método qualitativo o pesquisador busca explicar o porquê das coisas, exprimindo o que convém ser feito.

Para a análise dos dados coletados, utilizou-se a análise de conteúdo que se configura como: “[...] um conjunto de técnicas de análise das comunicações, que utiliza procedimentos sistemáticos e objetivos de descrição do conteúdo das mensagens.” (Bardin, 2006, p. 38). Bardin (2006) pontua, ainda, que a análise de conteúdo desenvolve-se em três fases, as quais foram contempladas nesse trabalho: 1. pré-análise: consistiu no contato e organização dos dados relativos ao *graffiti* possibilitando sua sistematização; 2. exploração do material: feita a partir da exploração do material coletado, definida sua categoria e identificado seu contexto; e 3. tratamento dos resultados, inferências e interpretações: desenvolvido a partir da análise do *graffiti* e pela proposição de sua representação documental, pois, ainda que uma arte com caráter efêmero, deve ter garantida sua recuperação e consolidação como patrimônio cultural.

RESULTADOS

Para que fosse possível alcançar o objetivo de propor um modelo de representação documental para *graffiti*, foram identificados os elementos previstos na análise semiótica aplicada (cores, linhas, superfícies, formas e luzes), proposta por Santaella (2010), e na análise facetada proposta por Ranganathan: personalidade, matéria, energia, espaço e tempo. A construção desta pesquisa foi baseada no entendimento de que o emprego das categorias essenciais propostas por Ranganathan viabiliza um aprofundamento na análise documental de recursos imagéticos. Neste sentido, as categorias ranganathanianas foram relacionadas às categorias informacionais que sustentam a análise documental de imagens (COSTA, 2008). De modo prático, o modelo “Representação documental de *graffiti*” apresenta nove elementos de metadados, a saber: Título, Artista, Criação, Citação cultural, Citação espaço-temporal, Citação ideológica, Descrição, Resumo e Palavras-chave.

Em “título” acrescenta-se uma designação que identifica o *graffiti* pelas suas características principais na cena, de modo que seja objetivo e de fácil identificação do contexto. No campo “artista” insere-se o autor da obra. Em “criação”, são preenchidos dados relativos à data específica de criação, ou possível

década ou século. O preenchimento deste campo foi baseado no Código de Catalogação Anglo-Americano (AACR2), por reunir o conjunto de regras para a descrição bibliográfica, especificamente no capítulo 1 “regras gerais para descrição”. No campo referente à “citação cultural” o pesquisador insere trechos que permitem especificar o contexto de criação do *graffiti*. Em “citação espaço-temporal” são identificadas questões referentes ao período de criação do *graffiti*, com informações de localização do tipo de prédio no qual o *graffiti* foi inscrito. Este campo foi criado porque o *graffiti* não se encontra somente em locais pré-determinados (somente em prédios, podendo constituir murais em túneis ou em arcos de pontes, por exemplo) e este dado de localização facilita a identificação de qual ponto da rua ele está instalado. Sugere-se a inclusão de informações de geolocalização, uma vez que as coordenadas geográficas favorecem a localização do *graffiti* nas ruas, e que podem ser identificadas a partir de dados EXIF da imagem fotográfica do *graffiti*. Em “citação ideológica” identificam-se questões que envolvem a concepção crítica e ideológica da proposta artística do *graffiti*. Para o campo “descrição” adotou-se o AACR2 para informações acerca da especificação do *graffiti*, como quantidade de imagens que compõem o *graffiti* em si, bem como as especificações do mesmo, como a indicação de cores (em relação às regras de descrição física, agrupadas em regras do tipo x.5, onde ‘x’ designa o capítulo a ser consultado). Em “resumo” é relatado de forma minuciosa a característica física do *graffiti*, com os dados que reportam especificamente às categorias rangianthianas: personagem, matéria, energia, espaço e tempo que participam da cena. Por fim, no campo “palavras-chave” indicam-se os principais termos correspondentes à cena, vistos no campo “resumo” e também pelas terminologias extraídas mediante a análise facetada.

Para fins de exemplificação, a proposta de representação documental do *graffiti* pode ser analisada no quadro abaixo:

<i>Documento imagético</i>	<i>Elementos de metadados</i>	<i>Metadados</i>
	<i>Título</i>	<i>Gigante de Boston e Greenway</i>
	<i>Artista</i>	<i>Os Gêmeos</i>
	<i>Criação</i>	<i>2012</i>
	<i>Citação cultural</i>	<p>“Como muitas das obras de arte famosas que a precederam, o colorido mural que domina um parque em frente da principal estação ferroviária de Boston, na costa leste dos Estados Unidos, provoca alguma controvérsia. A pintura de 21 metros por 21 metros [...] mostra uma figura usando roupas brilhantes que não combinam entre si e com o rosto envolto no que parece ser um véu, mostrando apenas seus olhos vesgos. O trabalho [...] se tornou motivo de polémica quando a estação local da rede de TV Fox divulgou declarações de pedestres que passavam pelo local e criticavam a figura mascarada, que lembra um cartum. Alguns a consideraram ameaçadora e pediram que fosse removida do local.” (Fonte 1). “A figura enorme de um menino descalço, grafitada no sistema de ventilação do túnel Dewey, no centro de Boston, tem inspirado um movimento popular contrário ao mural assinado pelos grafiteiros brasileiros Otavio e Gustavo Pandolfo, conhecidos como Os Gêmeos. De acordo com a agência de notícias Reuters, cidadãos exigem a extinção de The Giant of Boston (O Gigante de Boston, na tradução livre) por acreditarem que o menino descalço se parece com um terrorista, já que está com o rosto todo, com exceção dos olhos, coberto por uma camiseta.” (Fonte 2). Fonte 1: http://br.reuters.com/article/entertainmentNews/idBRSPPE87800720120809</p>

		Fonte 2: http://veja.abril.com.br/entretenimento/grafite-de-os-gemeos-e-chamado-de-terrorista-nos-eua/
	Citação espaço-temporal	Praça Dewey, centro de Boston (100 Cambridge St., Boston, MA, EUA) “EUA, 2012. OS GEMEOS foram convidados pelo Instituto de Arte Contemporânea em Boston para pintar um mural e participar de uma exposição com curadoria de Pedro Alozo. Esta foi a primeira exposição solo de OS GEMEOS no museu ICABoston. O local do mural é particularmente especial por estar situado diretamente em frente à principal estação ferroviária e do edifício da Reserva Federal, bem como na Greenway.” (Fonte 1) Fonte 1: http://www.osgemeos.com.br/pt/projetos/instituto-de-arte-contemporanea-gigante-de-boston-e-greenway-eua/
	Citação ideológica	“As obras de Os Gêmeos apresentam características como o tom lúdico, as cores intensas, as formas com superfícies decoradas com xadrezes, listras, estampas e floridos, o uso da cor amarela para efeito de chamar a atenção do público, como usada pelas placas de trânsito se fazem presentes em seus personagens.” (Fonte: LIMA, 2012). LIMA, F. R. B. Imagem e tecnologia: webmuseu de arte. 2012. Dissertação (Mestrado em Ciência da Informação)- Universidade Estadual Paulista, Marília, 2012.

	Descrição	1 graffiti : látex : spray : molde vasado : color. ; 21 x 21 m
	Resumo	Graffiti, com dimensões de 21x21m. Retrata um personagem, com os olhos semicerrados, pupila pequena, sentado, com cotovelos apoiados no chão, pernas dobradas. Na cabeça está uma camisa ou um casaco vermelho, o qual cobre, parcialmente, o nariz e a boca. A camisa, de manga longa, com punho marrom, apresenta a cor verde como predominante no tecido estampado. O cinto é vermelho. A calça, com punho azul próximo aos tornozelos, é xadrez, com azul e vermelho predominantes. Está descalço, com o pé direito sobreposto ao pé esquerdo.
	Palavra-chave	Gigante. Boston. Os Gêmeos. Instituto de Arte Contemporânea.

CONCLUSÕES

Ao agregar as características da representação intrínseca e extrínseca do *graffiti* enquanto recurso imagético com vistas à representação documental, o modelo proposto alia as exigências necessárias para a interoperabilidade e a troca de registros bibliográficos, estabelecendo uma dinâmica expressiva entre os metadados descritivos, os metadados de assunto, além de metadados culturais, espaço-temporais e ideológicos.

O modelo “Representação documental de *graffiti*” apresenta nove elementos de metadados, a saber: Título, Artista, Criação, Citação cultural, Citação espaço-temporal, Citação ideológica, Descrição, Resumo e Palavras-chave.

Esta relação viabiliza a preservação e perpetuidade do *graffiti* como patrimônio cultural, por conta da representação imagética do *graffiti* em fotografias, e de mecanismos de recuperação, acesso, uso e apropriação do recurso imagético pelos sujeitos nos diferentes sistemas de informação, por conta da representação documental.

REFERÊNCIAS BIBLIOGRÁFICAS

- Bardin, L. (2002). *Análise de conteúdo*. Lisboa: Edições Setenta.
- Campos, R. (2009) Movimentos da imagem no graffiti: das ruas da cidade para os circuitos digitais. In: Carmo, R.; Simões, J. *A produção das mobilidades: redes, espacialidades e trajectos*. Lisboa: Imprensa de Ciências Sociais.
- Castleman, C. (1982) *Getting up: subway graffiti in New York*. Cambridge: Mit Press, 1982.
- Cooper, M. & Chalfant, H. (1984) *Subway Art*. Londres: Thames & Hudson.
- Corsi, E. (2004). Patrimônios histórico-culturais: uma nova perspectiva para o urbano e o rural através do turismo sustentável. *Caminhos de Geografia* (V. 2, ed. 10, pp. 22-34). Uberlândia. Recuperado em 12 janeiro, 2017 de <http://www.seer.ufu.br/index.php/caminhosdegeografia/article/viewFile/15323/8622>
- Costa, L. S. F (2008). *Uma contribuição da Teoria Literária para a análise de conteúdo de imagens publicitárias do fim do século XIX e primeira metade do século XX, contemplando aspectos da natureza brasileira*. Tese de doutorado, Universidade Estadual Paulista, Marília, SP, Brasil.
- Couchot, E. (2003) *A tecnologia na arte: da fotografia a realidade virtual*. Porto Alegre: Ed. UFRS.
- Farthing, S. (2011) *Tudo sobre arte: os movimentos e as obras mais importantes de todos os tempos*. Rio de Janeiro: Sextante.
- Garrido Arilla, M. R. (1999) *Teoría e historia de la catalogación de documentos*. Madrid: Síntesis.
- OS GÊMEOS. [*Gigante de Boston e Greenway*]. (2012) Recuperado em 01 de abril de 2017, de <http://www.osgemeos.com.br/pt/projetos/instituto-de-arte-contemporanea-gigante-de-boston-e-greenway-eua/#!/5588>.
- Gerhardt, T. E. & Silveira, D. T. (Org.). (2009). *Métodos de pesquisa*. Porto Alegre: Ed. UFRGS.
- Köch, J. C. (2003). *Fundamento da metodologia científica: Teoria da ciência e prática da pesquisa*. (21.ed.). Petrópolis: Vozes.
- Lakatos, E. M. & Marconi, M. A. (2001). *Fundamentos de metodologia científica*. (4.ed.). São Paulo: Atlas.
- Lima, F. R. B. (2012). *Imagem e tecnologia: webmuseu de arte*. Dissertação de mestrado, Universidade Estadual Paulista, Marília, SP, Brasil.
- Lima, F. R. B., Santos, P. L. V. A. C. & Francisco, J. B. (2016, Novembro) Superfícies alteradas: a condição dos grafites nos espaços urbanos de São Paulo. *Anais do Encontro Nacional de Pesquisa em Ciência da Informação*. Salvador, BA, Brasil, 17. Recuperado em 30 março, 2017, de <http://basessibi.c3sl.ufpr.br/brapci/index.php/article/download/50180>.
- Lima, F. R. B. & Santos, P. L. V. A. C. (2014) Museu e suas tipologias: o webmuseu em destaque. *Informação & Sociedade: Estudos*, (V. 24, ed. 2, pp. 57-68) João Pessoa. Recuperado em 25 fevereiro, 2016, de <http://www.ies.ufpb.br/ojs/index.php/ies/article/view/16244/11491>

- Manini, M. P. (2004). Análise documentária de fotografias: leitura de imagens incluindo sua dimensão expressiva. *Cenário Arquivístico: Revista da Associação Brasileira de Arquivologia*, (V. 3, ed. 1, pp. 16-28) Brasília.
- Manini, M. P. (2002). *Análise documentária de fotografias: um referencial de leitura de imagens fotográficas para fins documentários*. Tese de doutorado, Universidade de São Paulo, São Paulo, Brasil. Recuperado em 14 junho, 2017, de <http://www.teses.usp.br/teses/disponiveis/27/27143/tde-23032007-111516/pt-br.php>.
- Pato, P. R. G. (2014). Ícone, índice e símbolo, fundamentos para ler e organizar a informação em imagens. *Anais do Encontro Nacional de Pesquisa em Ciência da Informação*, Belo Horizonte, MG, Brasil.
- Ponofsky, E. (1979). *Significado nas artes visuais*. Tradução de Maria Clara F. Kneese e J. Guinsburg. (2.ed). São Paulo: Perspectiva.
- Riout, D. et al. (1985) *Le livre Du Graffiti*. Paris: Alternatives.
- Santaella, L. & Nöth, W. (2005). *Imagem, cognição, semiótica, mídia*. São Paulo: Iluminuras.
- Spinelli, L. (2007) Pichação e comunicação: um código sem regra. *Logos: Comunicação & Universidade*, (V. 14, ed. 26, pp. 111-121) Rio de Janeiro. Recuperado em 19 junho, 2016, de <http://www.logos.uerj.br/PDFS/26/08lucianospen.pdf>
- Smit, J. W (1989). *A análise da imagem: um primeiro plano*. In: Smit, Johanna Wilhelmina (coord.). *Análise documentária: a análise da síntese*. (2.ed). Brasília: IBICT.
- Smit, J. W (1996). *A representação da imagem*. *INFORMARE: Cadernos do Programa de Pós-Graduação em Ciência da Informação*, (V. 2, ed. 2, pp. 28-36) Rio de Janeiro.
- Sontag, S. (1981). *Ensaio sobre a fotografia*. Tradução de Joaquim Paiva. (2.ed). Rio de Janeiro: Arbor.
- UNESCO. (1972) *Convenção sobre a Salvaguarda do Patrimônio Mundial, Cultural e Natural*. 1972. Recuperado em 30 março, 2017, de <http://whc.unesco.org/archive/convention-pt.pdf>.
- Ventura, T. (2009). Hip-hop e graffiti: uma abordagem comparativa entre Rio de Janeiro e São Paulo. *Análise Social*, (192), p. 605-634. Recuperado em 19 junho, 2016, de http://www.scielo.mec.pt/scielo.php?script=sci_arttext&pid=S0003-25732009000300007&lng=pt&tlng=pt.
- Zafalon, Z. R. (2012) *Scan for Marc: princípios sintáticos e semânticos de registros bibliográficos aplicados à conversão de dados analógicos para o Formato MARC21 bibliográfico*. Tese de Doutorado, Universidade Estadual Paulista, Marília, SP, Brasil.
- Zafalon, Z. R. & Dal'Evedove, P. R. (2016, Novembro) Representação documental: pesquisa e ensino. *Anais do Encontro Nacional de Pesquisa em Ciência da Informação*, Salvador, BA, Brasil, 17. Recuperado em 30 março, 2017, de <http://www.ufpb.br/evento/lti/ocs/index.php/enancib2016/enancib2016/paper/viewFile/4176/2391>



DESAFIOS À ORGANIZAÇÃO E ACESSO AO PATRIMÓNIO CULTURAL

U. PORTO DIGITAL MUSEUM PROJECT: TOWARDS CONVERGENCE IN UNIVERSITY'S INFORMATION MANAGEMENT

Maria Manuela Pinto¹, Susana Medina², Paula Menino Homem², Augusto Ribeiro³, Marisa Monteiro³, Rodolfo Matos³, Alexandre Lourenço³, Sara Oliveira⁴

¹University of Porto / CIC.DIGITAL Porto (PORTUGAL), mmpinto@letras.up.pt

²University of Porto / CITCEM (PORTUGAL), smedina@fe.up.pt; phomem@letras.up.pt

³University of Porto (PORTUGAL), aecr@uporto.pt; mmonteir@fc.up.pt;
rodolfo@uporto.pt; aafonso@reit.up.pt

⁴University of Porto / Master in Information Science FEUP-FLUP (PORTUGAL), saraoliveira878@gmail.com

ABSTRACT The University of Porto (U.Porto) aims to promote its scientific and cultural heritage through digital media by providing the integrated aggregation of digital information and related metadata which is managed and preserved in archives, libraries, museums and other technology information systems and repositories. This goal is meant to be reached by a theoretical based approach and an information management intervention model – the MGSIU-AP (Active and Permanent University Information System Management Model), which support an intervention toolkit directed to U.Porto Museums. This toolkit is complemented by a learning through projects educational approach, a digital platform which considers, among others, semantics and systems' interoperability, a long term digital preservation repository and an online heritage discovery portal to be developed in the scope of the U.Porto Digital Museum, itself an ongoing interdisciplinary project. This paper describes and discusses the development and subsequent process' model which has an inherent value of standardisation. It considers tools as CIDOC Conceptual Reference Model (CRM), providing an extensible ontology for concepts and information in cultural heritage and museum documentation, the standard SPECTRUM 4.0, describing 21 management procedures in workflow format, along with the policies and minimum features required for the standards to be met. Here will also be presented the work in progress in the field of domain ontologies, framed by an interdisciplinary theoretical approach which crosses information science, information systems and museology in the context of a service-oriented approach. Both goals and issues will also be examined.

KEYWORDS *University heritage, university museum, digital museum, information management, MGSIU-AP.*

COPYRIGHT This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

The University of Porto (U.Porto) is a centenary public institution whose Mission is to create scientific, cultural and artistic knowledge, superior education strongly anchored in research, the social and economic recognition of knowledge and the active participation in the development of the communities where it is embedded. Its roots are in scientific and cultural entities of the eighteenth and nineteenth

centuries, and concentrates, in its museological unities, the testimony of the fruitful link between art and research, collecting and teaching, personal introspection and recognition of the world.

Referring to European university museums, Lourenço (2006) pointed to existing critical issues such as facilities, a real diagnosis, the study and interpretation of collections, and Rorschach, in 2012, stressed the management and audiences of the Museum, recognizing the duality of a focus between the internal and the external public. Sharing with Lourenço the focus on the identity, role and mission of the university museum Rorschach stated:

“[...] university museum must never forget that it is in fact a very different kind of beast¹, with a different mission and access to a different, and very strong, array of intellectual resources within its parent university. University art museums can do things that larger municipal museums cannot do, and in my view they ought to do these things, indeed must do them.

What kinds of things? More intellectually risk-taking exhibitions; real engagement of students in creative and meaningful ways that have the potential to nurture life-long lovers and supporters of the arts; meaningful involvement of faculty across disciplines, that can lead to a broader understanding and appreciation of the key importance of art and visual culture in civilizations and cultures throughout human history; and new ways of thinking about collections, including long-term loans from underused collections in larger museums, experimentation with new media in partnership with related university disciplines and resources, and the building of important collections in new areas not yet recognized by the major museums.” (RORSCHACH, 2012).

In fact, the university’ museum distinguishes itself by the unique characteristics of its collection(s): the disciplinary diversity, the richness of the collections typology, the exploration of multiple combinations in educational function, the transdisciplinary investigation and interpretation, either museological or non-museological.

Nonetheless, the identity crisis acknowledged by researchers and professionals has marked the university museums in the last decades. Resulting from and supporting the educational process and the research, these museums see themselves between a custodial and technicist paradigm that roams, in a transverse way, archives, libraries and museums, as well as a tendency to a “public” positioning as cultural service. This means that they tend to move away from the focus on its natural prosumer (producer and consumer) inside the academy to external communities. Meanwhile, academic prosumers increasingly value technological mediation and forget ‘traditional’ spaces which are facing diverse constraints, namely financial.

From a theoretical point of view, the paradigmatic change in emergent scientific areas – such as Information Science (IS) – point to an ongoing shift based on interdisciplinary and trans disciplinary approaches that entail interscientific fields as Information Systems and takes on the Archivistics, Librarianship and even Museology as a branch or applied disciplines of IS. Nevertheless, it is a movement under construction and the increasing complexity of the initiatives in which the university’s cultural and scientific heritage services are involved entails abilities and competencies that go beyond their current framework of resources and expertise and suffers with the absence of a common work environment with shared practices.

¹ Termo aplicado ao Museu universitário por Otto Freundlich, na primeira metade do século XX.

Focusing the mission of University of Porto's museums it is oriented towards the study, representation, description, safeguard, promotion and dissemination of collections of educational and research artifacts. It considers those that are significant to the academy and the scientific universe, to the institutional and "eminent figures" of the University celebration, and to pedagogical and research purposes, through collections specifically conceived with those goals in mind or naturally resulting from them.

Regarding the incorporation of institutional contexts, the U.Porto museums protect material evidence with origin in predecessor institutions to the University since the 18th century until the present, as well as artifacts and information of personal and professional routes provided by teachers, researchers and *alumni*, along with others related to the institution's activities, to the city, to the region, to the country and international interactions.

Thus, U.Porto, as other higher education institutions, faces a unique role in this field pointing to four main levels:

- producing new knowledge, namely epistemological thinking and models (at research and intervention levels);
- empowering its information system and other resources, information/heritage services and technical expertise, through a systemic approach, the integrated management and collaborative practices;
- narrowing the gap between heritage services, technology and classrooms focusing on the dynamic co-production and (re)use of digital contents and tools and providing the opportunity of using as learning projects the real internal and external communities' needs / problems, including heritage ones, encouraging professors, students and researchers to find solutions inside the teaching-learning process and through educational practices, namely assessment works;
- transferring new knowledge to society, enabling the innovation process through platforms, services and activities;
- Building networks.

This paper discuss one of its main issues, the integrated management of the museum' collections representation and description as well as any other related metadata and information (born-digital and digitized) controlled by institutional services, namely information services, either traditional (archives, libraries, museums) or emerging (information management and information technologies).

This goal is meant to be reached, despite information services providers and technological platforms, by a theoretical based approach, an intervention model – the Active and Permanent University Information System Management Model (MGSIU-AP) –, a learning by projects educational concept / platform – the U.OpenLab – and a digital infrastructure which considers, among others, semantics and systems' interoperability, a shared long term digital preservation repository and online discovery portal in the scope of the U. Porto Digital Museum, itself an ongoing interdisciplinary project.

THEORETICAL CONTEXT AND METHODOLOGY

The present paper aims to present a project which emerged within university museums and directed to the university scientific and cultural heritage and its dissemination and promotion through technologies,

by providing the aggregation of existing “digital collections” (archives, libraries and museums’ digital assets and other technological information systems or repositories).

This purpose is rewritten when it considers the ongoing paradigm shift conceptually based on the main concepts of *information*, *communication* and *information system*, reshaping the concept of *heritage* and reviewing the concept of *culture*.

Information and *communication* refer directly to the human capacity of mentally shaping ideas and impressions and share them with the beings of the same species (phenomenon):

Information: In the scope of the trans and interdisciplinary INFORMATION SCIENCE proposed and defended by us, it has got a double semantic functionality. It refers to a human and social phenomenon that includes shaping ideas and emotions (to inform) and the exchange, the effective interaction of those ideas and emotions between human beings (to communicate). It identifies a scientific object: structured set of codified mental and emotional representations (signs and symbols) shaped by social interaction, which can be recorded in any material medium (paper, film, compact disc, magnetic tape, etc.) and, therefore, communicated in an asynchronous and multidirected way (DELTCI, 2017).

Communication: Since the middle of the 20th century this term is mixed in or mistaken for mass journalism, namely in radio, television and, more recently, in multimedia channels, although in the available literature there plenty of examples of the broad spectrum of the use of such a plastic and rich concept. It is a synonym of human and social interaction and assumes, necessarily, the existence of information in the form of messages or contents which are transmitted, shared, in brief, communicated (Dictionnaire encyclopédique des sciences de l’information et de la communication, 1997: 120-123). The term info-communicational completely identifies a human and social phenomenon represented partially or partly by the concepts Information and Communication (DELTCI, 2017).

Conversely, *heritage* is still connected to the descriptive need, inserted in Modernity, to aggregate distinct objects, human and natural, referring to different phenomena, whose common features are conventional (to value the old, rare, handcrafted and pre-industrial are mental and affective dispositions which arise and change with the evolution of the historical process):

Heritage: a construct from the 1800s filled, mostly, with romanticism, nationalism, historicism and reaction to industrialism (capitalist industrial revolution with its successive waves since the 18th century). It is inseparable from the ethnographic and historical-archeological conception of CULTURE that flourished in the 19th century. It encompasses all the material and immaterial goods, identifiers of a culture, of a community, of a people and essential to the ideological cohesion of the Nation State. It is a concept that has gained a strong ideological density and impact through time, which does not help, but hinders instead, its inclusion in a scientific and epistemological field. (DELTCI, 2017)

Culture: Having Edgar Morin’s definition as an inspiration, culture is understood as a human and social dimension with two capitals: a cognitive, mythological and ritual capital (knowledge, beliefs, norms, prohibitions, values) and a technical capital (know-how, practices, rules). The cultural materializations (objects and behaviors) result from the latter, while the former absorbs the strict sense of culture (letters, arts and sciences) accepted by most «schools» and theories and questioned in essays of great value. In a strict sense, culture and information are confused with each other and become synonyms” (DELTCI, 2017).

Finally the concept of *information system*, which cannot be confused with the *technological information system*, although inseparable:

Information System: is a totality formed by the dynamic interaction of the parts, in other words, it has a lasting structure with a flow of states in time. So being, an Information System is constituted by the different types of INFORMATION recorded or not externally to a subject (that each person has in his/her memory is an information of the system), it does not matter what is the MEDIUM (material and technological), in accordance with a structure (producing / receiving entity) extended by the action in the line of time. [...] an ARCHIVE or about a LIBRARY [or a Museum]: one and other are Services and Institutions and in this measure they can be seen as a system; or they are an IS [Information system], which corresponds to all the Information produced / received or incorporated in the System Archive or in the System Library [or in the System Museum] (DELTCI, 2017).

The scientific and theoretical bases point to a new informational paradigm and a convergence based on a systemic, informational and disciplinary perspective of university information system, comprehending Museology in its disciplinary relationship with Archivistics, Librarianship and Documentation, and the Museum as an information (and metadata) producing and managing entity, deepening theoretically convergence movements as LAM (Libraries, Archives and Museums).

Clearly, knowledge-driven societies need new tools but we believe that archives, libraries, and museums are likely sources for innovations in knowledge management. Likewise, these institutions can become important sites for teaching information literacy and developing skills in information retrieval and evaluation.

We also contend that archives, libraries, and museums will continue to fill a broad social need to decrease disparities between “information haves” and “information have-nots” especially as skills in acquiring, evaluating, manipulating, and generating information become more fundamental to individual and social well-being. [...] LAM must work at the center of epistemological frameworks that are simultaneously coherent and destabilizing, for there is no alternative mechanism for progress in knowledge (HEDSTROM e KING, s.d.).

The assumed positioning, applied to the U.Porto Digital Museum project, has the contribution of the epistemological and theoretical thinking (Silva & Ribeiro, 2002) and was framed by the intervention model MGSIU-AP, typified as systemic-informational and organizational/managerial model and pointing to a service-oriented approach (Pinto, 2015). This model was based on the research and diagnosis made at Portuguese Public Universities, specifically at University of Porto, and involving traditional services (Archives, Libraries, Documentation Centers and Museums), as well as the emerging ones such as Informatics and the Information Management function (Pinto, 2015).

Thus, the U. Porto Digital Museum project assumes a holistic, systemic and integrated approach, regarding information management and the impact of ICT. This reflection occurs on two levels: a) advantages and difficulties of standardization efforts on informational production, flow, organization and representation at the Museum; and b) addressing information sharing among information services within the university, leading to collaborative efforts directed at information management improvement in attaining a common institutional mission.

Given this conceptual model and the fact that the U.Porto Digital Museum must be able to integrate information from multiple disciplines and various sources, the authority control and descriptive metadata working group was formed in order to: a) build up a map of existing data elements in online

catalogues; b) analyse metadata requirements; c) adopt and implement descriptive metadata standards, such as CIDOC-CRM, which provides an extensible ontology for concepts and information in cultural heritage and museum documentation; d) and test the domain frameworks, developed with the direct involvement of a sample of U.Porto Museums and most significant collections. This sample differentially represented the disciplines, the collections and the user communities within the U. Porto from which the collections were drawn. It comprised the Museum of Natural History and Science of U.Porto, FEUPmuseum - Museum of the Faculty of Engineering of U.Porto, Abel Salazar House Museum and Museum of Medical History “Maximiano Lemos”- Faculty of Medicine U.Porto.

Initially we focused our attention on existing online catalogues in order to identify metadata structures, data elements and information representation and retrieval tools (classification schema, thesaurus, etc) that had been used to describe artefacts in the referred museums. Since 2008, every museum manages its collections using Index Rerum (<http://indexrerum.com>), presented as an integrated inventory, management and dissemination web based software, except FEUPmuseu which adopted In Arte (http://sistemasfuturo.pt/index_en.html), presented as a system in compliance with international standards of management and documentation of cultural heritage.

At this stage, we were also engaged in the process of mapping those elements, structures and tools whose heterogeneous nature soon became evident due to their varied provenance. Some of the key findings provide the background to the argument that, in order to build a future metadata structure(s) that can deal with this variety, it will be a top priority to fulfil two main goals: assure an answer to specific needs and clearly adopt international standards of management and documentation of cultural heritage, namely those developed by the Documentation Committee of the International Council of Museums (ICOM-CIDOC) and related areas, as Libraries and Archives.

However, additional metadata standards should be also considered, especially those that are capable to support the specificities of disciplines such as Natural History or Anatomy.

In order to provide an appropriate representation of U.Porto museums’ collections, the second stage was based on interviews in context with those museum curators for whom a data system that supports interoperability and access across collections (internal and external of the U.Porto management system) is a relevant need. Together with the curators, the team focused on evidence of the required types of descriptive metadata and controlled vocabularies.

The final phase comprises the cross evaluation of the collected data, regarding both metadata structures and values, terminology consistency and data transformations that should be accomplished before exporting data to a new system, in order to design an appropriate metadata infrastructure for such a multidisciplinary universe.

Additionally, the *U.Porto Museums Procedures Manual* is being developed with the direct collaboration of the U.Porto Museums, aiming at guaranteeing, through a set of clear instructions, the museological procedures/practices standardization and an adequacy to the specificity of each museum and collection. Based on SPECTRUM 4.0, a tool that implements and expands this work to the service management and procedures field, this manual consists of a group of procedures, activities and guidelines for U.Porto museums, describing how these tasks are to be carried out, in line with the museum diagnosis and good practices. This manual describes 21 collections management procedures in workflow format, along with the policies and minimum requirements. It contributes to the: preventing of errors, clarifying of the processes; coordinating activities; developing an orientation tool for the

service and its monitoring, including organization and representation procedures; fulfilling management's goals; and supporting the integration of new collaborators (Ferreira, 2016).

These activities were complemented by a *Software Requirements Document* directed at services and collections management, aiming to serve as a basis to a software evaluation, selection and acquisition. It was based on the previous standards and resulting frameworks and documents, as well as on the *Collections Management Software Criteria Checklist – 2012*, from the Canadian Information Heritage Network (CHIN). Two other guide documents were delivered: the *U.Porto Digitization Guide* and the *Digital Preservation Requirements Guidelines* (Rua, 2016).

RESULTS

ICOM guidelines suggest that museum documentation is concerned with the development and use of information about the objects within a museum collection and the procedures which support the management of the collection and will be recorded in a museum documentation system, underlining its importance:

“With effective documentation, a museum should be able to facilitate:

- collection policies;*
- collection care and accountability;*
- collection access, interpretation and use; collection research.*

[...]

4. The documentation system and the information it holds should conform to appropriate standards developed by national and international organisations, while taking account of local needs (Code 2.20). The documentation chapter in the ICOM guide to Running a Museum: a Practical Handbook has a comparison of a number of these standards, including the CIDOC Conceptual Reference Model (CRM), the CIDOC Guidelines, the AFRICOM Handbook and SPECTRUM. The LIDO standard should be taken into account when planning the contribution of data to, and interchange with, other systems (ICOM-CIDOC 2012).

Notwithstanding, the main challenge is not the implementation of a unique museum documentation system but, depending on the nature of the artefact and related information, it is how represent and describe it in the “right” form and with the suitable system, under the principle of interoperability and multilevel aggregation. As stated by Fernanda Ribeiro regarding archives field, we should discuss the interest and the use of classification in a transverse way (across collections and services), “either in what concerns intellectual organization of information, in a hierarchical structure of organic-functional or thematic type, or at information representation and retrieval level, regarding the informational content” (Ribeiro 2013). It is a question of improving knowledge and experience of archives and “how classification has been used in order to establish organizational schedules, which provide context to the informational production of the organisms (persons, families, corporate bodies) that generate the archives and how such schedules represent (or not) with scientific objectivity the informational reality being considered” (Ribeiro 2013) and the use of “indexing languages in other information systems, namely libraries, where subject classifications have a generalized application since the last decades of 19th century” (Ribeiro 2013).

Thus, the achieved results of the presented project are related with the empowerment of the university information system, its information services, specifically museums, framed by a trans and

interdisciplinary theoretical approach crossing Information Science and Museology with contributions of Information Systems in the context of a service-oriented approach.

As main result we point an operationalization toolkit which comprises:

- a theoretical framework and an intervention model (MGSIU-AP);
- an interdisciplinary authority control and descriptive metadata working group (curators, museologists, information managers - archivists, librarians -, computer specialists domain specialists);
- a transversal metadata set profile and domain metadata set profiles;
- the *U.Porto Museums Software Requirements Document*;
- the *U.Porto Museums Procedures Manual*
- and the *U.Porto Digitization Guide*.

Considered, but not yet concluded, are the *U.Porto Museums Metadata Guidelines*.

As an important work in progress, conducted by Sara Oliveira (2017), we point towards an experimental study directed at the heritage audiovisual content aggregation process, thus providing enriched and high quality data and metadata. It was based on the results of EUscreen and EUscreenXL and its main contribution, in an experimentation context, is the construction of an application profile - Dublin Core Application Profile (DCAP) - according to the rules defined by the Singapore Framework, a reference model for the description of resources that aims to implement semantics interoperability between different communities of practice (Malta & Baptista 2014). The created DCAP, named *Audiovisual Content Application Profile (ACAP)* contains functional requirements, a domain model, a description set profile, user guides and syntax guides (Oliveira 2017). This profile will improve both the information management within the audiovisual production of TVU., a U.Porto service targeting recordings, documentation and the communication of scientific knowledge and heritage, as well as its aggregation by EUscreen, a professional audiovisual content portal, and its consequent access through Europeana portal. The role of audiovisual content aggregators is essential not only to enable the availability of content on Europeana, but also to map and standardize metadata models from different sources into a common scheme, in this case Europeana Data Model (EDM) (Oliveira 2017).

Last, but not least, there is the development of a prototype - U.OpenLab - which provides support to the first experiences related with digital contents production based on museum's collections in the context of a *learning by projects* methodology applied to the real internal museum's needs / problems, encouraging professors, students and researchers to find solutions inside the teaching-learning process and through educational practices, namely assessment works (Pinto et al. 2016).

CONCLUSION

It is questioned, today, in which way the university museum expresses the faculties' mission and how they are part of the museum, especially when there is a diversity of materials, whose trace, was usually lost in the memory but obeys a logic of production and information recording. These artefacts arrived in the museum through donation, acquisition or transfer between institutions and involve personal, institutional and research paths. We know enough to understand that transferring networks of national and international scientific knowledge were built and should be promoted and preserved through rich and meaningful structures, descriptions and access points.

Only a systematic study allows the recognition of the processes of Science and Heritage production, reaching university individuals, artifacts and research paths, accumulators of an unrepeatably story and evidence of an institution activity. Only a systematic evaluation of a disaggregated collection (Zoology, Mineralogy, Botanic and Anthropology's materials, among others) can provide sense to a very rich universe. One must remember that one thing is the history of the institution and its communities and another is the way how they should be considered in a preservation, interpretation and information management policy, with the first usually being carried out within a legal and regulatory framework, whereas the second shelters a narrative that demands systematic information, resources crossing and a domain based description, technically accurate and which follows international standards and good practices.

These specificities of the university museums reveal a territory in which the borders between the education, the investigation and the museological practice blur. They are also associated to the concept of a digital space that establishes relationships between the museums, the libraries, the archives, the more recent "university information systems" and the different "communities" and collections. This digital space will allow the connection with the global networks, that go beyond the academy, building a meeting "locus" with individuals from "everywhere", independently of education level, acquired knowledge and singular interests.

To do so, it imposes, along with a technical and technological perspective, a more epistemological position, i.e., the information valuing as a human and social process and the historicity (organic-contextual production) of the cultural heritage and richness, instead of applying neutral rules of classification and inventory of the decontextualized collections that are numerous times only comprehended by the domain's specialists.

With the Digital Museum Project, the University of Porto proposes to:

- dynamically enrich the delivering of digital contents by the U.Porto Museums with the study, representation and description of the collections, as well as the efficiency of the information management processes and services (Pinto, 2015), with positive expression on the information seeking, retrieval and reuse;
- induce networking, collaboration and knowledge sharing among communities and disciplinary areas inside academia, as information management communities of practice, crossing functional, technical and scientific areas.

We believe that a "Digital Museum" will enable enhancing the shared creation, the open experimentation, the endeavours of volunteers around Heritage and the continuous update of the collections, the existences and the indicators of past existences, now gone, and from "today", according it is made the work to discovering, describing, representing, disseminating and developing them with the "unexpected digital media".

ACKNOWLEDGEMENTS

The authors would like to acknowledge the most valuable support in the translation of the current article to Sandra Luna.

REFERENCIAS

- Ferreira, F. (2016). *Uma plataforma comum para os museus da U.PORTO : gestão, processos e tecnologia*. (Dissertação de Mestrado), Universidade do Porto.
- Hedstrom, M., King, J. L. On the LAM : Library, Archive, and Museum Collections in the Creation and Maintenance of Knowledge Communities. Retrieved from <http://www.oecd.org/education/country-studies/32126054.pdf>
- ICOM CIDOC (2012). Statement of principles of museum documentation [Version 6.2, June 2012]. Retrieved from <http://network.icom.museum/cidoc/working-groups/documentation-standards/principles-of-museum-documentation/>
- Lourenço, M. C. (2006). Nota breve sobre os museus e colecções das universidades europeias. *Boletim da Rede Portuguesa de Museus*, 19 (mar 2006), 13-18.
- Lourenço, M. C. (2008). Where past, present and future knowledge meet : an overview of university museums and collections in Europe. *Museologia Scientifica Memorie*, 2, 321-329. Retrieved from <http://www.anms.it/upload/rivistefiles/105.PDF>
- Malta, M. C., & Baptista, A. A. (2014). A panoramic view on metadata application profiles of the last decade *Int. J. Metadata, Semantics and Ontologies*, 9(1), 58-73.
- Medina, S., Pêgo, J. P., Machado, C., et al. (2009). Building a collaborative network for the digital representation of engineering collections. In *Putting University Collections to Work in Teaching and Research – Proceedings of the 9th Conference of the International Committee of ICOM for University Museums and Collections (UMAC)*. Berkeley, USA, 10 a 13 de September de 2009.
- Monteiro, M., Bernardo, L. M. & Araújo, J. M. (2008). Preserving memory in the University of Porto: the Physics Collection of the Faculty of Science. *Sci. Instrum. Soc. Bull.* 97 (June 2008), 27-30.
- Monteiro, Marisa L. & Soares, Miguel F.O. (2010). Meteorological (and other) instruments revealed: the Collection of the Geophysical Institute of Porto University. *Sci. Instrum. Soc. Bull.* 104 (Mars 2010), 17-21.
- Oliveira, T. C. M. (2010). *User generated content audiovisual para instituições de ensino superior: proposta de boas práticas*. (Dissertação de Mestrado), Universidade de Aveiro.
- Oliveira, S. (2017). *Disseminação de conteúdos audiovisuais na web : uso de um perfil de aplicação para a gestão e agregação dos recursos da TVU*. (Dissertação de Mestrado), Universidade do Porto.
- Pinto, M. M. (2015). The Portuguese University: Knowledge Leverage towards Innovation. In G. Jamil, S. Lopes, A. Malheiro da Silva, & F. Ribeiro (Eds.) *Handbook of Research on Effective Project*

Management through the Integration of Knowledge and Innovation (466-490). Hershey, PA. Retrieved from doi:10.4018/978-1-4666-7536-0.ch024

Pinto, M. M. (2015). *A Gestão da Informação nas Universidades Públicas Portuguesas : Reequacionamento e proposta de modelo* (Tese de Doutoramento em Informação e Comunicação em Plataformas Digitais), Universidade do Porto, Universidade de Aveiro.

Pinto, M. M., Matos, R., Medina, S., Abreu, R., Sousa, A., Faria, L., Amorim, J., Paiva, S., Martins, N., Barbosa, T., Figueiredo, T., Feio, P., Magalhães, D., & Almeida, M. (2016). Narrowing the gap between museums, classrooms and technology: the U.OpenLab Initiative Prototype . *ICERI2016 Proceedings*. (pp. 4442-4448). IATED Publications. Retrieved from doi: 10.21125/iceri.2016.2052

Ribeiro, F. (2003). *O Acesso à informação nos arquivos*. Lisboa : Fundação Calouste Gulbenkian; Fundação para a Ciência e a Tecnologia, Ministério da Ciência e do Ensino Superior. ISBN 972- 31-1017-2.

Ribeiro, F. (2007). “An integrated perspective for professional education in libraries, archives and museums: a new paradigm, a new training model,” *Journal of Education for Library and Information Science*, Spring, 48(2).

Ribeiro, F. (2013). O uso da classificação nos arquivos como instrumento de organização, representação e recuperação da informação. In *Informação e/ou Conhecimento: as duas faces de Jano – Atas do I Congresso ISKO Espanha e Portugal / XI Congreso ISKO España*. Faculdade de Letras da Universidade do Porto - CETAC.MEDIA. ISBN: 978-989-8648-10-5

Ribeiro, L. M., David, G., Azevedo, A., & Santos, J. C. M. (2001, March 28-30). Developing an Information System at the Engineering Faculty of Porto University. In J. Knop, P. Schirmbacher. *The Changing Universities. The Role of Technology: Proceedings*. Paper presented at International Conference of European University Information Systems, 7th, EUNIS, Berlin, Humboldt University. Retrieved from <http://hdl.handle.net/10216/606>

Rorschach, K. (2012). Why do universities have museums? Duke News. [Talk given at Nasher Museum of Art, 10.11.2004]. Duke University. Retrieved from http://www.nasher.duke.edu/about_from-director.php

Rua, J. (2016). *Digitalização, preservação digital e acesso em instituições de memória : contributos para o projeto Museu Digital da U.PORTO*. (Dissertação de Mestrado), Universidade do Porto.

Silva, A. M. da, & Ribeiro, F. (2002). *Das ciências documentais à ciência da informação : ensaio epistemológico para um novo modelo curricular*. Afrontamento.

Simões, M. G. (2011). *Classificações bibliográficas : percurso de uma teoria*. Coimbra : Almedina. ISBN 978-972-40-4608-2.

Stewart, Michael (2015). The University of Porto Science Museum. In Report of the Annual Study Conference: Portugal 3rd-8th May 2015, Sci. Instrum. Soc. Bull. 127 (December 2015), 21.

POSTERS

PRÁTICA INFORMACIONAL EM DOMÍNIOS INTERDISCIPLINARES: UM ESTUDO SOBRE A PRODUÇÃO DO CONHECIMENTO NA GOVERNANÇA DA ÁGUA

Marianna Zattar¹, Regina Maria Marteleto², Marta Pedro Varanda³

¹Universidade Federal do Rio de Janeiro, 0000-0002-3328-3591, mzattar@facc.ufrj.br

²Instituto Brasileiro de Informação em Ciência e Tecnologia, 0000-0002-3439-0217, regina.mar2@gmail.com

³Universidade de Lisboa, 0000-0002-9762-2724, marta@iseg.ulisboa.pt

RESUMO Esta comunicação é uma apresentação dos resultados de uma tese de doutorado elaborada no Programa de Pós-Graduação em Ciência da Informação do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), em convênio com a Universidade Federal do Rio de Janeiro (UFRJ). Estuda o processo de produção do conhecimento e as práticas informacionais no domínio do conhecimento e campo científico da Governança da água, a partir da análise das interações estabelecidas pelos membros dos grupos de pesquisa, por meio das configurações das redes de coautoria de integrantes de três grupos de pesquisa selecionados. Parte da compreensão da Governança da água como domínio complexo, de caráter teórico-prático inter e transdisciplinar, ao requerer a participação integrada de diversas disciplinas na reflexão das questões atinentes à gestão das águas, tanto quanto de outros atores do próprio campo científico e da sociedade. Utiliza como referencial teórico-conceitual os conceitos de campo científico (Pierre Bourdieu) e de domínio do conhecimento (Birger Hjørland e outros). Emprega os conceitos de rede social e de prática informacional como categorias operacionais na análise empreendida no campo empírico da pesquisa. Adota as metodologias qualitativa e de análise de redes sociais (ars) nos procedimentos metodológicos. Indica nos resultados que as dinâmicas e os processos de produção, mediação e apropriação de conhecimentos no domínio do conhecimento da Governança da água devem extrapolar as tradicionais fronteiras disciplinares da organização do conhecimento e rompe com a compreensão do fazer científico como exclusivo dos pesquisadores e dos saberes da ciência. Conclui que as dinâmicas de representação e organização do conhecimento exigem a superação dos limites disciplinares a partir da abordagem crítica de fenômenos complexos

PALAVRAS-CHAVE *Produção do conhecimento. Prática informacional. Domínio do conhecimento. Campo científico.*

ABSTRACT This presentation is a compilation of the results of a doctoral thesis elaborated in the Graduate Program in Information Science of the Brazilian Institute of Information in Science and Technology (IBICT), in partnership with the Federal University of Rio de Janeiro (UFRJ). We investigated the process of knowledge production and informational practices in the knowledge domain and scientific field of Water Governance, based on the analysis of the interactions established by the members of the research groups and through the configurations of the coauthoring networks of members of the three selected research groups. The thesis stemmed from the understanding of Water Governance as a complex domain, composed of a theoretical and practical character that is interdisciplinary and transdisciplinary, by requiring the integrated participation of several disciplines in the reflection of issues related to water management, as well as other actors in the scientific field and society. It used as theoretical-conceptual reference the concepts of scientific field (Pierre Bourdieu) and knowledge domain (Birger Hjørland et al.). It used the concepts of social network and

informational practice as operational categories in the analysis undertaken in the empirical field of research. It applied qualitative methodologies and analysis of social networks (ars) in the methodological procedures. The results indicate that the dynamics and the processes of production, mediation and appropriation of knowledge in the knowledge domain of Water Governance must extrapolate the traditional disciplinary boundaries of the organization of knowledge, and the results break with the understanding of scientific doing as unique to researchers and to science. We concluded that the dynamics of representation and organization of knowledge demand the overcoming of the disciplinary limits from the critical approach of complex phenomena.

KEYWORDS *Knowledge production, Informational practice, Knowledge domain, Scientific field.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

No âmbito das exigências relacionadas à água está a Governança da água como uma possibilidade de organização das discussões e decisões em diferentes contextos, caracterizando-se como um domínio complexo, de caráter teórico-prático inter e transdisciplinar, ao requerer a participação integrada de diversas disciplinas na reflexão das questões atinentes à gestão das águas, tanto quanto de outros atores do próprio campo científico e da sociedade. Assim, esta comunicação é uma apresentação dos resultados de uma tese de doutorado que tem como objetivo central o estudo o processo de produção do conhecimento e as práticas informacionais no domínio do conhecimento e campo científico da Governança da água. Para isso, destaca-se como ponto de partida que a Governança da água é também uma governança ambiental, o que exige a participação de diferentes atores em iniciativas coletivas em um contexto complexo e plural (Empinotti & Jacobi, 2012).

O referencial teórico que orienta a pesquisa foi sistematizado conceitualmente e operacionalmente. O referencial teórico-conceitual parte dos conceitos de campo científico (Pierre Bourdieu) e de domínio do conhecimento (Birger Hjørland e outros) no âmbito dos eixos epistemológicos, temáticos, conceituais e metodológicos e tem como objetivo construir uma base que possibilite a análise da comunidade e do ambiente de produção do conhecimento sobre a Governança da água sob a perspectiva relacional e interacionista. A relação de complementaridade parte da perspectiva do paradigma social da informação para o estudo das bases sociais e culturais historicamente estabelecidas pelas comunidades discursivas na produção e organização do conhecimento. A fim de estabelecer um diálogo entre o campo teórico e o terreno da pesquisa, empregam-se os conceitos de redes sociais e práticas informacionais como operadores empíricos no estudo das questões que envolvem as dinâmicas da construção da informação e das estruturas informacionais nos processos de produção do conhecimento no campo/ domínio da Governança da água a partir dos conceitos de rede social e de prática informacional como categorias operacionais na análise empreendida no campo empírico da pesquisa.

Os métodos utilizados para coleta de dados são a pesquisa documental e a entrevista e para análise dos dados as metodologias qualitativa e de análise de redes sociais a partir da análise das interações estabelecidas pelos membros dos grupos de pesquisa, identificadas por meio das configurações das redes de coautoria de integrantes de três grupos de pesquisa selecionados.

Os principais resultados da pesquisa demonstram que as dinâmicas de representação e organização do conhecimento exigem a superação dos limites disciplinares que extrapolam as tradicionais formas de

produção do conhecimento por meio das interações e integrações que são estabelecidas nos mais diversos contextos e elos entre diferentes atores. Trata-se, portanto, de uma alternativa teórico-metodológica para a compreensão crítica de fenômenos complexos.

METODOLOGIA

A pesquisa foi realizada com os pesquisadores e líderes de três Grupos de Pesquisa da região sudeste do Brasil, que estavam cadastrados na Plataforma Lattes do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), uma vez que essa instituição está inserida no Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC) e no Sistema Nacional de Ciência, Tecnologia e Inovação (SNCTI).

No passos metodológicos foram percorridas nove etapas que buscaram reconhecer os grupos, os seus respectivos membros (pesquisadores e líderes) e as publicações (artigos científicos) de forma que pudessem ser estudadas os aspectos relacionados às práticas informacionais do campo/ domínio estudado, a Governança da água, a partir da identificação das redes de coautoria das publicações dos pesquisadores de cada um dos grupos e as demais relações estabelecidas no processo de produção do conhecimento.

A coleta de dados teve como unidade de análise as relações estabelecidas entre os atores dos grupos de pesquisa selecionados na definição do campo empírico. Desse modo, a coleta dos dados foi realizada com a utilização da triangulação de métodos, que integra a análise das estruturas, dos processos e dos resultados ao incorporar os atores sob a perspectiva da ação (Minayo et al., 2005). O esquema escolhido nesta pesquisa foi composto pela identificação e seleção dos grupos de pesquisa estudados, a análise temática da produção científica dos pesquisadores, a identificação das coautorias dos artigos científicos selecionados, a identificação das informações dos periódicos dos artigos e a entrevista roteirizada. A análise dos dados coletados foram realizados com a análise da rede de coautoria e com a leitura exaustiva e transversal que pretendeu a interpretação das transcrições das entrevistas para que pudesse orientar as conexões entre os níveis teóricos e práticos desta pesquisa e alcançar os objetivos desta tese.

RESULTADOS

Os resultados mostram que a relação entre campo e domínio e a compreensão teórico-prática da inter e transdisciplinaridade podem ser uma alternativa na abordagem crítica de fenômenos complexos que extrapolam os limites disciplinares e a participação exclusiva de atores acadêmicos na produção do conhecimento.

Na análise dos dados contidos nos currículos Lattes dos pesquisadores que compõem os grupos de pesquisas pode-se notar a diversidade dos níveis de formação (graduação, mestrado e doutorado) e das áreas de conhecimento na indicação das áreas de formação e de atuação. Nas redes de coautorias percebe-se a participação de atores não acadêmicos como autores das publicações e de atores acadêmicos que não são membros do grupo de pesquisa ou à instituição à qual o grupo encontra-se vinculado. Nas entrevistas foram percebidas as parcerias entre os grupos com outros grupos, nos planos nacional e internacional; a atuação em comunidades externas à universidade (movimentos sociais,

comitês de bacias e instituições governamentais), diferentes formas de comunicação e de produção do conhecimento intra grupos.

CONCLUSÕES

Indica-se, como conclusão, que os processos de produção, mediação e apropriação de conhecimentos e da participação de novos atores sociais na produção do conhecimento extrapolam as fronteiras disciplinares e organizacionais tradicionais e o entendimento do fazer científico como exclusivo dos pesquisadores e dos saberes da ciência.

Por fim, indica-se que o estudo da Governança da água apresenta-se como um contexto ou objeto privilegiado nas dinâmicas entre ciência e sociedade. Sob essa perspectiva, diversos são os potenciais de desenvolvimento de pesquisas futuras a partir desta tese. Para apresentá-los de forma sintética, serão elencadas quatro perspectivas. A primeira é o desenvolvimento de pesquisas que relacionem os dados sobre Governança da água ou Gestão de Recursos Hídricos em diferentes países, como por exemplo Brasil e Portugal. A segunda se encontraria no desenvolvimento de estudos inter, multi (ou pluri) e transdisciplinares em contextos diferentes da Governança da água ou Gestão de Recursos Hídricos. A terceira possibilidade está na ordem do referencial teórico-conceitual com o desenvolvimento de estudos que contemplem os conceitos de campo de Pierre Bourdieu e de domínio de conhecimentos de Birger Hjørland e outros autores. A quarta perspectiva está nos conceitos operacionais de prática informacional de Reijo Savolainen e outros autores e de redes sociais de Linton C. Freeman e outros autores.

REFERÊNCIAS BIBLIOGRÁFICAS

- Bourdieu, P. (1983a). Esboço de uma teoria da prática. In R. Ortiz (Org.), *Pierre Bourdieu: Sociologia* (p. 46-81). São Paulo, SP: Ática.
- Bourdieu, P. (1983b). O campo científico. In R. Ortiz. (Org.), *Pierre Bourdieu: Sociologia* (pp. 122-155). São Paulo, SP: Ática.
- Bourdieu, P. (2004). *Os usos sociais da ciência: por uma sociologia clínica do campo científico*. São Paulo, SP: UNESP.
- Castro, J. E. (2007). Water Governance in the 21st century. *Ambiente e Sociedade*, 10(2), 97-118.
- Empinotti, V., & Jacobi, P. R. (2012). Pensando as questões da água de uma nova forma. In V Empinotti, & P R Jacobi (Org.), *Pegada hídrica: Inovação, corresponsabilização e os desafios de sua aplicação* (pp. 7-14). São Paulo, SP: Annablume.
- Hjørland, B. (2002). Domain analysis in information science: eleven approaches traditional as well as innovative. *Journal of Documentation*, 58(4), 422-462. Recuperado de <http://www.emeraldinsight.com/doi/abs/10.1108/00220410210431136>.
- Hjørland, B. (1998). Theory and metatheory of information science: a new interpretation. *Journal of Documentation*, 54(5), 606-621.

Hjørland, B., & Albrechtsen, H. (1995). Toward a new horizon in information science: domain-analysis. *Journal of the American Society for Information Science*, 46(6), 400-425. Recuperado de <http://goo.gl/LngNmI>.

Minayo, M. C. S., Souza, E. R., Constantino, P., & Santos, N. C. Método, técnicas e relações em triangulação. In M C S Minayo, S G Assis, & E R Souza. (Orgs.). *Avaliação por triangulação de métodos: Abordagem de programas sociais* (pp. 71-104). Rio de Janeiro, RJ: Fiocruz.

Savolainen, R. (2007). Information behavior and information practice: reviewing the "umbrella concepts" of information-seeking studies. *The Library Quarterly*, 77(2), 109-132.

CLASSIFICAÇÃO DA INFORMAÇÃO ARQUIVÍSTICA SEGUNDO MÉTODOS FILOMEMÉTICOS: METADADOS COMO UNIDADES FENOTÍPICAS?

L. S. Ascensão de Macedo

Universidade de Coimbra, 0000-0001-7251-7314, laureanomacedo@student.fl.uc.pt

RESUMO Este artigo explora métodos filogenéticos aplicados à classificação da informação arquivística. Analisam-se paralelismos discursivos baseados na ‘metáfora biológica’ existente na literatura arquivística e na CI, bem como as aplicações deste método nas Humanidades Digitais, conhecida como filomemética. Para obter resultados que nos permitam analisar do ponto de vista evolucionista a estrutura de um sistema de informação arquivístico, analisamos uma amostra descrita numa matriz de distâncias/similaridade baseada em propriedades do Records in Context. Em resultado disto, efetua-se neste artigo uma análise em torno dos outputs das aplicações utilizadas no âmbito da análise filogenética. O contributo deste estudo exploratório visa contrapor os modelos top-down (base funcional) de representação da informação arquivística face a um modelo de representação científica e evolucionista dos conjuntos informacionais. Conclui-se que os metadados podem constituir-se como unidades fenotípicas por permitirem apresentar padrões de (dis)similaridade entre as unidades informacionais analisadas.

PALAVRAS-CHAVE *Classificação arquivística, metáfora biológica, filomemética, metadados, traços fenotípicos.*

ABSTRACT This paper explores phylogenetic methods applied to archival information and knowledge classification. We analyse discourses about the “biological metaphor” in the archival and information science literature, and their methodological applications in the Digital Humanities, known also as phylomemetics. We adopt an exploratory method using metadata as phenotypic traits and RiC properties to build distance/similarity matrices with dendritic visualizations based on set of taxa (typological records). This study analyses the outputs of phylogenetic tools to support our interpretation of the data. The importance of this study shows us alternative methods to functional analysis in archival classification theory to an evolutionary classification of archival information systems.

KEYWORDS *Archival classification, biological metaphor, phylomemetics, metadata, phenotypic traits.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

A classificação arquivística consiste numa atividade de organização sistemática e estruturada de conjuntos informacionais arquivísticos, com o objetivo de controlar e proporcionar meios para descrever, representar e recuperar informação (Simões & Freitas, 2013; Simões, Freitas, & Rodríguez-Bravo, 2016). A representação da informação arquivística materializa-se em instrumentos de pesquisa,

u. g., roteiros, guias, inventários, catálogos, índices, *thesauri* e bases de dados relacionais (MacNeil, 2012), podendo conter *instrumenta* sintéticos como planos, grelhas, quadros ou tabelas de seleção, com objetivos distintos de utilização. No entanto, a classificação arquivística não se cinge apenas às necessidades de recuperação de informação, mas constitui a ‘coluna vertebral’ tanto para a tomada de decisão no âmbito da avaliação da informação como na definição de níveis de acesso, de proteção de dados, de confidencialidade, de preservação e conservação, inclusa a (re)engenharia dos processos de trabalho (Mokhtar & Yusof, 2015). Ainda assim, a maioria dos instrumentos de pesquisa adotou uma visualização hierarquizada e multinível das classes informacionais ou documentais (Gartner, 2016), de acordo com determinadas regras e convenções adotadas a nível internacional e nacional para a sua descrição (cf. DGARQ, 2011), com forte penetração em sistemas de informação eletrónicos (Niu, 2013). A representação da informação arquivística, segundo as convenções normativas supramencionadas, permanece configurada de forma hierarquizada, estática, unidimensional, idiossincrásica, onde o fundo/arquivo é descrito como um sistema fechado, qual “silo”, e tratado como resíduo da atividade das organizações. Os mais recentes estudos apontam para a representação multidimensional da informação arquivística segundo modelos conceituais e sua transposição para ontologias formais para suporte à descrição. Esta perspetiva aproxima-se dos pressupostos da ISO 21.127 (2014), como é o caso do *RiC: Records in Context* (Gueguen et al., 2013; ICA, 2016). Apesar de a literatura arquivística referir-se à ‘metáfora biológica’ como um traço distintivo em relação aos modelos de classificação em bibliotecas, esta perspetiva carece de uma análise à luz da Teoria Evolucionista.

Este artigo efetua um estudo exploratório aos métodos filomeméticos enquanto modelos de classificação científica aplicáveis à representação da informação arquivística. Analisar-se-ão os conceitos operatórios associados à ‘metáfora biológica’, com uma contextualização das aplicações de modelos filomeméticos aplicados nas Humanidades Digitais, seguido de uma exemplificação de aplicação em torno de unidades informacionais.

CONTEXTUALIZAÇÃO

A 'METÁFORA BIOLÓGICA' NA ARQUIVÍSTICA E O DISCURSO EVOLUCIONISTA EM CI

As comparações inspiradas em processos biológicos para explicar fenómenos na arquivística não são espúrias. Autores fundamentais da arquivística, como Muller, Feith & Fruin (1898), Jenkinson (1922), Casanova (1928) estabeleciam analogias em torno do processo de formação e natureza dos arquivos baseadas em 'metáforas biológicas' (Adell, 2010; Bak, 2012; Hurley, 1993). Na terminologia arquivística, existem vários conceitos definidos com base no símile biológico, materializado em termos como arquivo, fundo, série e, designadamente, organicidade (cf. Quadro 1).

Quadro 1. *Arquivo e Organicidade na terminologia arquivística (português).*

Entrada terminológica	Definição*	Fonte
Arquivo	<i>Conjunto orgânico de documentos, independentemente da sua data, forma e suporte material, produzidos ou recebidos por uma pessoa jurídica, singular ou colectiva, ou por um organismo público ou privado, no exercício da sua actividade e conservados a título de prova ou informação. É a mais ampla unidade arquivística. A cada proveniência corresponde um arquivo.</i>	(BNP, 2010, p. 379)
Organicidade	<i>Característica que decorre do facto de os arquivos reflectirem, enquanto produto natural da actividade de uma administração – no seu todo ou em cada uma das suas unidades, bem como nas relações entre elas – a vontade e o funcionamento dessa administração. Fundamenta os princípios da proveniência e do respeito pela ordem original.</i>	(BNP, 2010, p. 379)
	<i>Relação natural entre documentos de um arquivo em decorrência das atividades da entidade produtora.</i>	(Brasil. Arquivo Nacional, 2005)

* Sublinhados nossos.

Na terminologia inglesa, vários termos evidenciam esta perspectiva biológica, *u. g. records aggregation* (IP2, 2008), *archival bond* (Duranti, 2015), *archival sedimentation* (Mata Caravaca, 2015), *archival nature*, *accumulation*, *organic collection* (Pearce-Moses, 2005), incluindo equívocos, como expressões “archival darwinism” (Ham, 1993) ou “neo-Darwinian construction” (Cook, 2011).

A estrutura hierarquizada das classes arquivísticas – arquivo/fundo, secção, série, documento composto/simples – tem sido entendida como uma conexão entre unidades informacionais (Adell, 2010). O símile entre série e a *species* lineana também não é espúrio (Adell, 2010; Hurley, 1993; Janzen, 1994; Mai, 2011; Rieppel, 2010; Wilkins, 2009). Podemos encontrar em Jenkinson, em 1943, a formulação de uma estrutura inspirada na taxonomia lineana:

In the case of Archives Species, Genera and Orders are represented by Classes, Groups and Categories of Records: which are, to an almost equal extent, natural, not artificial, divisions.
(Jenkinson, 1943/2003, p. 200)

O discurso evolucionista no âmbito da CI conta com importantes contributos que permitem dispor de uma base teórica e interdisciplinar no âmbito da filogenia cultural (Gnoli, 2006). Assiste-se a uma aplicação de modelos filomeméticos para explicar processos evolutivos em sistemas de informação enquanto abordagens emergentes, designadamente

- da curadoria digital (Thomer & Weber, 2014; Weber, 2015; Weber & Thomer, 2016);
- da computação forense (Costa et al., 2015; Lanterna & Barili, 2017; Oliveira et al., 2016);
- da estematologia/crítica textual (Andrews & Macé, 2012; Heikkilä, 2007), em alternativa aos métodos de *document clustering* (Huang, 2008).

FILOGENÉTICA / FILOMEMÉTICA NAS HUMANIDADES E CIÊNCIAS SOCIAIS

Darwin defendeu que *all true classification is genealogical* (1859, p. 420), para reforçar a validade dos princípios da descendência com modificação e da hereditariedade. Tais princípios darwinianos constituem um dos pilares da filogenética sistemática (Lecointre, 2015).

No entanto, a filogenética não se tornou numa disciplina exclusiva da Biologia Evolutiva: as suas potencialidades transdisciplinares tiveram larga aplicação, por exemplo, na classificação de macrofamílias linguísticas (Atkinson, 2006; Fangerau, Geisler, Halling, & Martin, 2013), na crítica textual (Andrews & Macé, 2012; Heikkilä, 2007), na arqueologia evolutiva (García Rivero, 2013, 2016; O'Brien & Lyman, 2003), na teoria organizacional (McCarthy, 2005), nos processos evolutivos em tecnologias industriais (Lee, Jung, Yoon, Kim, & Jung, 2013) e na evolução cultural (Currie & Mace, 2011; Mesoudi, 2011, 2016). A transposição de métodos filogenéticos para o domínio das Humanidades e Ciências Sociais constitui uma abordagem emergente sobre a aplicação de métodos baseados na epistemologia evolucionista (Gontier, 2012) e na exploração de modelos algorítmicos meta-heurísticos (Eiben & Smith, 2012).

Os modelos de classificação da Sistemática Filogenética resultaram em três escolas filosóficas: a fenética, a cladística e a taxonomia filogenética (García Rivero, 2010). A diferença entre estas três escolas, exemplificadas sinteticamente *infra* (Quadro 2), consiste no facto de a fenética não se basear no princípio de descendência com modificação, ao contrário da cladística e da taxonomia evolutiva. A visualização das relações entre os *taxa* em dendrogramas, os quais constituem hipóteses de visualização (Morrison, 2014).

Quadro 2. Escolas filogenéticas pós-síntese

<i>Escola</i>	<i>Método de classificação</i>	<i>Modelo de visualização</i>	<i>Procedimentos</i>	<i>Autores fundamentais</i>	<i>Problemas</i>
<i>Fenética (numerical taxonomy)</i>	<i>Similaridade morfológica</i> <i>Unidade: OTU</i>	<i>Fenograma</i>	<i>Estatístico (análise fatorial, análise de conglomerados, análise coordenadas principais).</i>	(Sneath & Sokal, 1973; Sokal, 1986)	<i>- não explica a história evolutiva;</i> <i>- gera paralelismos, convergências evolutivas.</i>
<i>Cladística (sistemática filogenética)</i>	<i>abordagem “descendência com modificação”:</i> <i>carateres análogos, homologia, sinapomorfias, autapomorfias, simplesiomorfia.</i>	<i>Cladograma</i>	<i>Estatístico (máxima verossimilhança, de inferência bayesiana, método dos mínimos quadrados, parsimónia)</i>	(Hennig, 1975)	<i>- critério da parsimónia insuficiente;</i> <i>- desvalorização da anagénesse</i> <i>- ausência de critérios de ranking.</i>
<i>Taxonomia evolutiva</i>	<i>Similitude entre classes, grupos monofiléticos e parafiléticos, anagéneses.</i>	<i>Filograma</i>	<i>Estatístico (máxima verossimilhança, de inferência bayesiana, método dos mínimos quadrados, parcimónia)</i>	(Mayr, 1981; Mayr & Bock, 2002)	<i>- critério de semelhança e genealogia não é passível de ser representada simultaneamente.</i>

Fonte: adaptado a partir de García Rivero (2010).

Sendo os fenótipos o conjunto de propriedades morfológicas de um organismo (Mahner & Kary, 1997), em CI tem como possíveis equivalentes os metadados, ativos semânticos e propriedades significativas, uma vez que constituem elementos descritivos de artefactos informacionais (Gartner, 2016).

MATERIAIS E MÉTODOS

Sendo o documento de arquivo a unidade mínima de um fundo na hierarquia de classes, a análise documental através dos métodos da diplomática contemporânea consiste na identificação de carateres externos e internos das tipologias documentais (Rogers, 2015, 2016).

Os esquemas de metadados – tanto de tipo descritivo, administrativo, técnico como de preservação,

além da variedade e diversidade de esquemas (cf. Riley, 2010) – podem ser analisados segundo processos evolutivos: *some metadata are re-used (inheritance)*, *other metadata are eliminated (extinction)* and *still others are updated or newly generated (mutation)* (Niu, 2013). De modo similar, os traços fenotípicos culturais são objeto de *recombination, copying error, and [...] can be the foundation for the production of new traits* (O'Brien, Lyman, Mesoudi, & VanPool, 2010).

Propomos a seguinte questão de investigação: os metadados possibilitam reconstituir um sistema de informação arquivístico segundo uma perspectiva evolucionista, constituindo-se como hipótese de classificação científica desse sistema informacional? Adota-se um método misto (qualitativo e quantitativo) segundo uma abordagem de descrição *bottom-up* (documento a documento). Para possibilitar um estudo filomemético, estabeleceu-se o seguinte fluxo de trabalho:

- (i) **Coleta de dados e processamento.** Seleccionámos aleatoriamente 9 documentos compostos (N= 41 peças documentais) de tipologia diversa classificados segundo MEF/ASIA (DGLAB, 2016) proveniente do sistema de informação do Governo Regional da Madeira (GRM), referente ao período de registo de 2014 a 2016.

Quadro 3. Classes MEF/ASIA

Classe MEF/ASIA	Nota de definição	Relação com outros PN
300.10.302 Eliminação de documentos e informação	Destruição controlada de documentação e informação administrativa, independentemente do suporte, decorridos os prazos legais. Inicia com a proposta e termina com a comunicação da eliminação à entidade competente. Inclui análise da documentação ou informação, verificação de prazos legais de conservação administrativa e de destino final, autorização de eliminação e elaboração de auto de eliminação.	Complementar#300.30.008#Registo de documentos e informação# Complementar#300.30.602#Avaliação de informação arquivística#
300.10.600 Depósito de bens	Aquisição da responsabilidade da gestão de bens de terceiros. Inicia com o pedido de depósito ou entrega de bens e termina com a devolução ou decisão sobre o destino dos bens. Inclui o registo dos bens a depositar, a sua receção e confirmação de depósito.	300.30.008#Registo de documentos e informação# Síntese(sintetizado)
300.30.008 Registo de documentos e informação	Inserção de dados relativos a documentos e informação, independentemente da sua natureza ou suporte. Inicia com a recolha de dados e termina com o seu registo. Inclui atualização de dados.	Complementar#300.30.602#Avaliação de informação arquivística# Complementar#300.10.302 Eliminação de documentos e informação#
300.30.602 Avaliação de informação arquivística	Atribuição de um valor informacional, de acordo com critérios de valorimetria aplicáveis, para efeitos de determinação do destino final da informação/documentação. Inicia com a identificação do objeto a avaliar e termina com a produção de instrumento de avaliação. Inclui estudo do contexto de produção, criação de estrutura de organização intelectual e justificação dos valores atribuídos.	Complementar#300.10.302 Eliminação de documentos e informação# Complementar#300.30.008 Registo de documentos e informação# Complementar#300.40.501 Conservação e proteção de informação#

Fonte: extraído de DGALB (DGLAB, 2016).

O tamanho da amostra é meramente experimental e constituímos uma matriz binária, acessível em Macedo (2017), onde se descreve cada documento simples pelas suas tipologias documentais identificadas no Quadro 4. Na matriz binária atribuíram-se valores numéricos: 0 significa que uma tipologia documental não possui determinada propriedade *RiC* e 1 possui essa propriedade, conforme Quadros 5.

Quadro 4. Relação de tipologias documentais

Abreviatura	Descrição	Tradição	Taxon	Obs.
auto_ceden	Auto de cedência a título devolutivo	original	I4	
auto_elim_d_pdg	Auto de eliminação, por via de portaria de gestão de documentos	cópia	C2, H2	
auto_elim_o_pdg	Auto de eliminação, por via de portaria de gestão de documentos	original	C1, H1	
auto_elim_rada_d	Auto de eliminação, por via de relatório de avaliação	duplicado	B2	
auto_elim_rada_o	Auto de eliminação, por via de relatório de avaliação	original	B1	
auto_de_entre	Auto de entrega	original	E1	
auto_incorp_d	Auto de incorporação	duplicado	F2	
auto_incorp_o	Auto de incorporação	original	F1	
guia-r_d_rada	Guia de remessa, por via de relatório de avaliação	duplicado	D2	
guia-r_o_rada	Guia de remessa, por via de relatório de avaliação	original	D1	
inf_int_o	Informação interna	original	A3, B3, C3, D3, G3, H3, I1	
idd_d	Instrumento de pesquisa	cópia	G2	
idd_o	Instrumento de pesquisa	original	G1	
ofic_c	Ofício	cópia	A5, B5, C5, D5, F5, G5, H5	
ofic_o	Ofício	original	A4, B4, C4, D4, F4, G4, H4, I2	
ofic_rem	Ofício	original	I3, F3, A6	Entidade externa
oficio_rem_o				
O_resp_oGA				
rada_d	Relatório de avaliação	duplicado	A2	
rada_o	Relatório de avaliação	original	A1	

Fonte: extraído de Macedo (Macedo, 2017).

Quadro 5. Extrato de matriz binária

Taxa	A1	A2	A3	A4	A5	A6	B1	B2	B3	B4	B5	C1	C2	C3	C4	C5	D1	D2	D3	D4	D5	E1	F1	F2	F3	F4	F5	G1	G2	G3	G4	G5	H1	H2	H3	H4	H5	I1	I2	I3	I4	
	rada_o	rada_d	inf_int_o	ofic_c	ofic_o	O_resp_oGA	auto_elim_rada_o	auto_elim_rada_d	inf_int_o	ofic_c	ofic_o	auto_elim_o_pdg	auto_elim_d_pdg	inf_int_o	ofic_c	ofic_o	guia-r_o_rada	guia-r_d_rada	inf_int_o	ofic_c	ofic_o	auto_de_entre	auto_incorp_o	auto_incorp_d	oficio_rem_o	ofic_exp_o	ofic_exp_d	idd_o	idd_d	inf_int_o	ofic_c	ofic_o	auto_elim_o_pdg	auto_elim_d_pdg	inf_int_o	ofic_c	ofic_o	inf_int_o	ofic_c	ofic_o	ofic_rem	auto_ceden
RiC-R074	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	
RiC-R075	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
RiC-R078	1	1	0	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	0	0	0	0	1
RiC-R079	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RiC-R080	1	1	0	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	0	0	0	0	
RiC-R082	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
RiC-R085	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
RiC-R086	1	1	1	0	0	1	1	1	0	1	0	1	1	1	0	0	1	1	0	0	1	1	1	1	0	0	1	1	1	0	0	1	1	1	0	0	1	0	0	0	0	
...	

Fonte: extraído de Macedo (Macedo, 2017), através de doi:10.7910/DVN/ERE8HL.

Quadro 6. Relação de entidades e de propriedades do RiC

Set of Properties	Properties	Set of Properties	Properties
Shared Properties of all Entities	RiC-P1 Global Persistent Identifier	Additional Property Specific to Person and Person Assumed Identity	RiC-P36 Gender
	RiC-P2 Local Identifier	Additional Property Specific to Delegate-Agent	RiC-P37 Technical Characteristics
	RiC-P3 Name	Additional Properties Specific to Corporate Body and Corporate Body Assumed Identity	RiC-P38 Services to the Public
	RiC-P4 General Note		RiC-P39 Contact Information
RiC-P5 Authenticity and Integrity Note	RiC-P40 Operating Hours		
Content Information	RiC-P6 Content Type	Properties of Position	RiC-P41 Facilities
	RiC-P7 Content Extent		RiC-P42 Type
	RiC-P8 Quality of Information		RiC-P43 Description
	RiC-P9 Scope and Content		RiC-P44 History
	RiC-P10 Encoding Format	Properties of Function	RiC-P45 Type
	RiC-P11 Language Information		RiC-P46 Description
	RiC-P12 Media Type		RiC-P47 History
	RiC-P13 Production Technique		RiC-P48 Type
Information about Carrier	RiC-P14 Medium	Property of Function (Abstract)	RiC-P49 Description
	RiC-P15 Physical or Logical Extent		RiC-P50 History
Information about Management and Use	RiC-P16 Physical Characteristics Note	Properties of Activity	RiC-P51 Description
	RiC-P17 Classification		RiC-P52 Type
	RiC-P18 Conditions of Access		RiC-P53 Description
	RiC-P19 Conditions of Use	RiC-P54 History	
	RiC-P20 History	Properties of Mandate	RiC-P55 Type
RiC-P21 Record State	RiC-P56 Description		
Properties of Record Set	RiC-P22 Authenticity and Integrity Note	Properties of Documentary Form	RiC-P57 History
	RiC-P23 Type		RiC-P58 Type
	RiC-P24 Accrual Note		RiC-P59 Description
	RiC-P25 Accrual Status	Properties of Date	RiC-P60 History
	RiC-P26 Arrangement		RiC-P61 Type
	RiC-P27 Classification		RiC-P62 Calendar
Properties Summarizing the Members of a Record Set	RiC-P28 History	Properties of Place	RiC-P63 Type
	RiC-P29 Content Extent		RiC-P64 Geographic Coordinates
	RiC-P30 Physical or Logical Extent		RiC-P65 Address
Properties Shared by All Agents	RiC-P31 Scope and Content	Properties of Concept/Thing	RiC-P66 Type
	RiC-P32 Type		RiC-P67 Description
	RiC-P33 Identity Type		
	RiC-P34 Language Information		
	RiC-P35 History		

Fonte: extraído de ICA.EGAD (ICA, 2016)

Refira-se que o esquema *RiC* possui 67 propriedades (Quadro 6) com 792 relações multientidades. As propriedades *RiC* foram reduzidas a 143 propriedades (marcadas em inglês pelo presente do indicativo).

De acordo com os Quadros 4 e 5 *supra* exarados, as referências aos documentos (*taxa*) são representadas na sequência de dois dígitos alfanuméricos em que a letra maiúscula representa o processo documental pertencente a um processo de negócio concreto e o número a tipologia documental (*u. g.*, *A1* decompõe-se em *A* equivale à classe 300.30.602 *Avaliação da Informação Arquivística*, e 1 corresponde a tipologia documental *relatório de avaliação*; *A4*: *A* = 300.30.602 *Avaliação da Informação Arquivística*; 4 = *ofício*).

Para aplicar relações de propriedades *RiC-R* com valor de 1, então *A1* consiste num *Record Component* (*RiC-R075*) que é parte de (*is part of*) de *Record* (*RiC-R14*), e inversamente (*RiC-R14-Record has part*

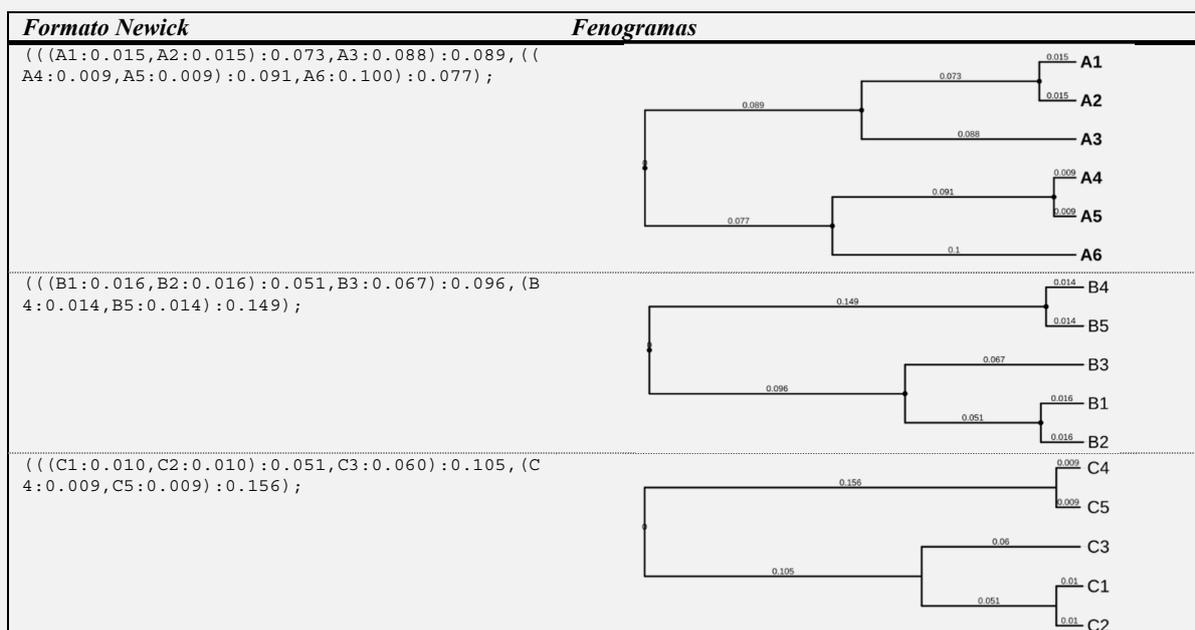
RiC-R075-Record Component). No caso de não ser explícita a presença de uma propriedade, recebe valor de 0. Obteríamos, por exemplo, uma sequência binária do tipo:

```
> A1
01101101111010111011011111110101011010011111101010100101010111110
1010000011010000110000000000110110111000011111100101101001111110
00111111000001000000000000100000000010000000000000000000000000100010
100010.
```

(ii) **Análise e visualização de dados.** Seleccionámos as aplicações *web DendroUPGMA* (Garcia-Vallve, Palau, & Romeu, 1999) para processamento de dados binários e para extração de *outputs* e *iTOL* (Letunic & Bork, 2016), para visualização dendrífica. Convertemos automaticamente os *datasets* binários em matrizes de distâncias segundo o coeficiente de Sørensen–Dice e exportado para o formato Newick (Cardona, Rosselló, & Valiente, 2008), cujos *outputs* são acessíveis em Macedo (2017). Adotámos um método de agrupamento hierarquizado UPGMA (*Unweighted Pair Group Method with Arithmetic Mean*), cujo resultado consiste num dendrograma baseado na (dis)similaridade entre *taxa* (Pavlopoulos, Soldatos, Barbosa-Silva, & Schneider, 2010).

RESULTADOS E DISCUSSÃO

Analisada cada tipologia documental, obtivemos numa matriz de binária 201 variáveis (*RiC* propriedades-relações) em 41 *taxa* (documentos simples). Obtivemos três tipos de *outputs*, calculado automaticamente por *DendroUPGMA* através do coeficiente de Sørensen-Dice, que resultou (i) numa matriz de distâncias, (ii) numa matriz de similaridade e (iii) correspondente formato Newick e respetivo dendrograma (Figura 1). O algoritmo UPGMA constituiu um parâmetro pré-definido em *DendroUPGMA* para a constituição de *clusters* e visualização dendrífica.



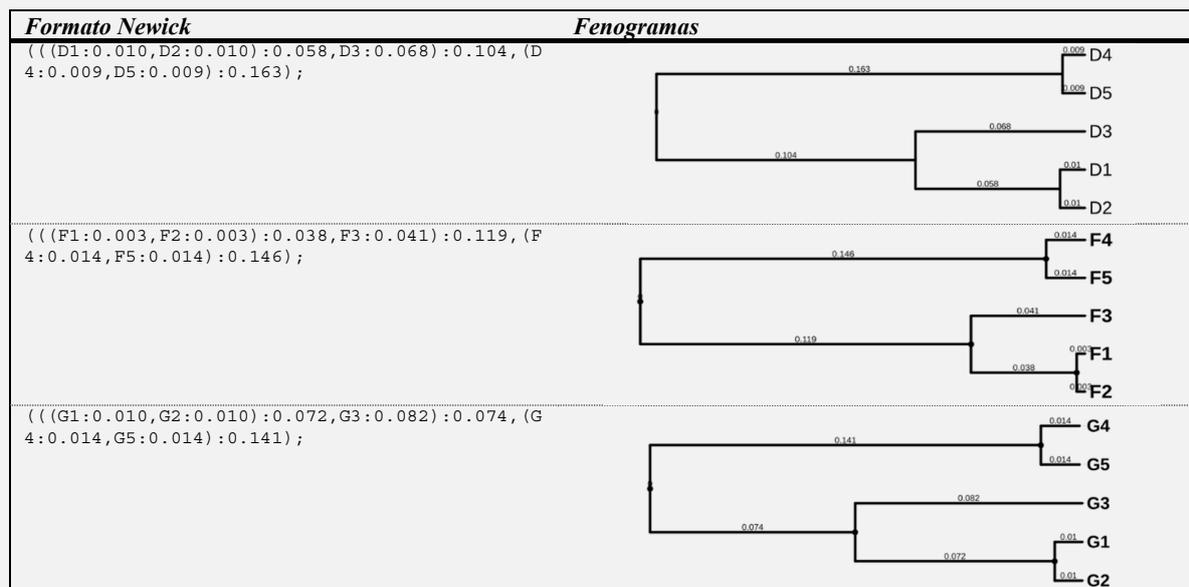
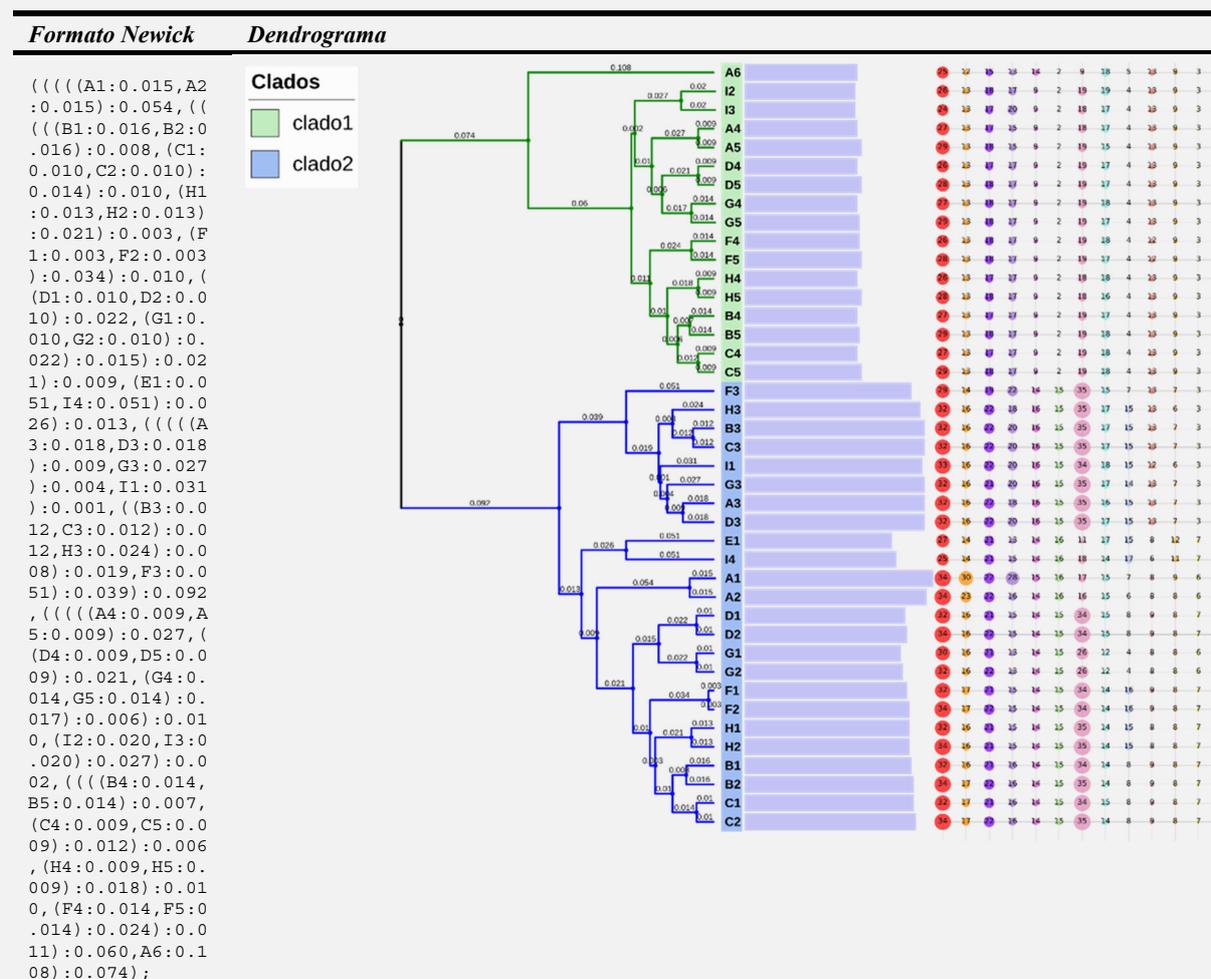


Figura 1. Fenogramas de relações entre taxa (tipologias documentais), por processos documentais, processado em *DendroUPGMA* (Garcia-Vallve et al., 1999)

A partir da Figura 1, verifica-se que a metainformação do *RiC*, enquanto traços fenotípicos, permite uma diferenciação entre os *taxa*. As relações entre os *taxa* de cada dendrograma evidenciam similaridades, como A1-A2, B1-B2, C1-C2, D1-D2, F1-F2 e G1-G2 (cf. *supra* Quadro 4), distinguindo-se entre si ao nível das tipologias documentais e ao nível do suporte e da tradição documental (analogico/digital/híbrido). A mesma leitura pode ser feita para os pares A4-A5, B4-B5, C4-C5, D4-D5, F4-F5, G4-G5 (ofícios). No caso das informações internas (A3, B3, C3, D3 e F3), verifica-se que se encontram em ramos isolados mas vinculados a um *cluster*, respetivamente A1-A2, B1-B2, C1-C2, D1-D2, F1-F2 e G1-G2, numa relação, por exemplo, de tipo “*RiC-R18-Record*” *is member of* “*RiC-R13-Record Set*”, que determina a formação de agregações de documentos. Refira-se, contudo, que os dendrogramas apresentados constituem fenogramas, *i. e.*, baseado apenas na similaridade morfológica entre os *taxa*, portanto não indiciador de filogenia (Khanafiah & Situngkir, 2006; Morrison, 2014).



As barras do cladograma correspondem ao somatório das propriedades RiC por *taxon* e desdobrados em 13 esferas coloridas na seguinte sequência (esquerda/direita): *Record, Record Component, Record Set, Agent, Occupation, Position, Function, Activity, Mandate, Documentary Form, Date, Place, Concept/Thing*.

Figura 2. Dendrograma de relações entre *taxa* (tipologias documentais), processado em *iTOL* (Letunic & Bork, 2016)

Se agregarmos os *taxa* num único dendrograma, é possível observar a constituição de dois ramos, visível na Figura 2. Em primeiro lugar, o *clado 1* representa a comunicação interinstitucional por meio de ofícios, verificando-se A6 como a correspondência recebida (em ramo isolado) e os restantes como correspondência expedida. No *clado 2*, verifica-se a preparação da comunicação interna e da informação de suporte. A título exemplificativo, o caso do par A4-A5 (300.30.602) aparece vinculado aos *taxa* D4-D5 (*idem*) e G4-G5 (300.30.008), constituindo um *continuum* que nasce da avaliação de informação sobre conjuntos informacionais acumulados, em que o auto de entrega e guia de remessa e os instrumentos de pesquisa partilham propriedades similares entre si. Igualmente, no *clado 2*, verificam-se relações de similaridade entre pares D1-D2 (300.30.602) e G1-G2 (300.30.008), pelo facto de o relatório de avaliação partilhar propriedades comuns aos instrumentos de pesquisa. Os pares B1-B2, C1-C2, H1-H2, por seu turno, derivam de procedimentos similares de eliminação de informação arquivística (300.10.302), independentemente se resultante de um relatório de avaliação ou da execução de uma portaria de gestão de documentos.

Estes resultados constituem, contudo, uma aproximação preliminar aos métodos filomeméticos, contando com aspetos positivos e com limitações:

- (i) O recurso a uma matriz binária permite-nos, em princípio, descrever a informação a partir da identificação de propriedades observáveis em tipologias documentais de arquivo. Tendo como exemplo a Figura 1, verifica-se a presença de relações hierárquicas entre os *taxa* (documentos simples), que poderão indiciar uma hipótese de relação de ordem física entre tipologias documentais obtida a partir da matriz binária. As propriedades *RiC*, contudo, não significam que sejam suficientes para estabelecer relações de filogenia. Conforme Figura 2, os *clusters* representam apenas (dis)similaridades morfológicas entre os *taxa*, mais próximo dos métodos fenéticos do que cladísticos, por o algoritmo UPGMA proporcionar este tipo de resultados (Khanafiah & Situngkir, 2006). A título de exemplo, autos de eliminação (H1-H2, B1-B2, C1-C2) e os autos de incorporação (F1-F2) dissemelham-se em relação ao *cluster* H3, B3, C3, I1, G3, A3, D3 (informações internas) e não se encontram vinculados entre si conforme a ordem “natural” ou “original” atribuída pelo produtor.
- (ii) Os ramos constituídos pelo algoritmo UPGMA revelam aspetos pertinentes sobre as relações funcionais entre os *taxa*. Thomer e Weber utilizaram algoritmos distintos (máxima verossimilhança) para identificar “eventos de especiação” a partir de metadados com fins de preservação digital (Thomer & Weber, 2014; Weber & Thomer, 2016). O recurso a outros algoritmos depende da granularidade informativa dos metadados e dos objetivos de análise filogenética, que terá resultados distintos em função da questão de investigação e da natureza do objeto de estudo (Yu & Gen, 2010). A principal limitação deriva do facto de a maior parte das aplicações informáticas para o estudo filogenético está desenhada apenas para processar informação sobre entidades bióticas.

CONCLUSÃO

Este artigo efetuou uma aproximação preliminar aos métodos filomeméticos aplicados à classificação da informação arquivística. Procuramos responder a uma questão atual que afeta à organização do conhecimento em sistemas de informação arquivísticos, na ausência de outros modelos disponíveis: de que modo podemos aplicar modelos de classificação científica, de base evolucionista, em sistemas de classificação arquivísticos? Propusemos uma análise contextual em torno da “metáfora biológica” baseada na literatura científica. A questão fundamental consiste em determinar que metainformação, perante a diversidade de esquemas, poderá ser significativa para uma classificação filomemética da informação arquivística baseada no princípio de hereditariedade (Niu, 2013). Recorremos, assim, a uma amostra de documentos de arquivo a partir da análise das tipologias documentais baseada numa matriz binária de propriedades *RiC*. Obtivemos várias hipóteses de vínculos de (dis)similaridade entre as unidades informacionais de nível de documento (*taxa*). As potencialidades deste modelo de prospeção de informação, independentemente da sua materialização, permitirão analisar conjuntos informacionais arquivísticos sob uma perspetiva evolucionista em torno de ecossistemas informacionais distintos.

Várias técnicas de classificação continuam a ser exploradas *ad haustum* no âmbito da aglomeração de dados por inferência computacional (Janssens, Zhang, Moor, & Glänzel, 2009; Jung, Bae, & Liu, 2008). No entanto, os modelos baseados em algoritmos evolucionistas estão mais próximos dos princípios arquivísticos, como proveniência, organicidade e ordem original. Importa, por conseguinte, indagar as

potencialidades e as limitações sobre os algoritmos meta-heurísticos aplicáveis à prospeção e reconstituição evolucionista da informação arquivística (Ventura & Luna, 2016; Yu & Gen, 2010).

Em última análise, procurámos explorar uma forma distinta das abordagens interpretivistas (base funcional) e encetámos uma metodologia filomemética. Revela-se cada vez mais necessário explorar na arquivística um discurso mais computacional no âmbito da organização do conhecimento, por permitir ambientes controlados para o processamento e interpretação da informação.

REFERÊNCIAS BIBLIOGRÁFICAS

- Adell, J. I. (2010). Clasificación de documentos y clasificación en biología. *Tábula*, 13, 105–124.
- Andrews, T. L., & Macé, C. Trees of Texts – Models and methods for an updated theory of medieval text stemmatology | Digital Humanities 2012, Digital Humanities (2012). Retrieved from <http://www.dh2012.uni-hamburg.de/conference/programme/abstracts/trees-of-texts-models-and-methods-for-an-updated-theory-of-medieval-text-stemmatology.1.html>
- Atkinson, Q. D. (2006). *From species to languages : a phylogenetic approach to human prehistory*. ResearchSpace@Auckland.
- Bak, G. (2012). Continuous classification: Capturing dynamic relationships among information resources. *Archival Science*. <https://doi.org/10.1007/s10502-012-9171-8>
- Cardona, G., Rosselló, F., & Valiente, G. (2008). Extended Newick: it is time for a standard representation of phylogenetic networks. *BMC Bioinformatics*, 9(532). <https://doi.org/doi.org/10.1186/1471-2105-9-532>
- Casanova, E. (1928). *Archivistica*. Siena: Stab. arti grafiche Lazzeri.
- Cook, T. (2011). We Are What We Keep; We Keep What We Are’: Archival Appraisal Past, Present and Future. *Journal of the Society of Archivists*, 32(2), 173–189. <https://doi.org/10.1080/00379816.2011.619688>
- Costa, F. de O., Lameri, S., Bestagini, P., Dias, Z., Rocha, A., Tagliasacchi, M., & Tubaro, S. (2015). Phylogeny reconstruction for misaligned and compressed video sequences. In *Image Processing (ICIP), 2015 IEEE International Conference on* (pp. 301–305).
- Currie, T. E., & Mace, R. (2011). Mode and tempo in the evolution of socio-political organization: reconciling “Darwinian” and “Spencerian” evolutionary approaches in anthropology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567), 1108–1117. <https://doi.org/10.1098/rstb.2010.0318>
- Darwin, C. R. (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London: John Murray.
- DGARQ. (2011). *Orientações para a Descrição Arquivística 3.^a Versão*. Lisboa: DGARQ.

DGLAB. (2016). Projeto ASIA – documento em consulta pública [dataset]. Lisboa: DGLAB. Retrieved from http://arquivos.dglab.gov.pt/wp-content/uploads/sites/16/2016/03/ASIA_divulgacao_2016-03-11.xlsx

Duranti, L. (2015). Archival bond. In L. Duranti & P. Franks (Eds.), *Encyclopedia of Archival Science* (Luciana Du, pp. 28–29). Rowman & Littlefield.

Eiben, Á. E., & Smith, J. E. (2012). Evolutionary Algorithms. In F. Neri, C. Cotta, & P. Moscato (Eds.), *Handbook of Memetic Algorithms* (pp. 9–27). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-23247-3_2

Fangerau, H., Geisler, H., Halling, T., & Martin, W. (Eds.). (2013). *Classification and Evolution in Biology, Linguistics and the History of Science Concepts – Methods – Visualization*. Stuttgart: Franz Steiner Verlag.

Garcia-Vallve, S., Palau, J., & Romeu, A. (1999). Horizontal gene transfer in glycosyl hydrolases inferred from codon usage in *Escherichia coli* and *Bacillus subtilis*. *Molecular Biology and Evolution*, 16(9), 1125–1134.

García Rivero, D. (2010). Introducción a la teoría de la clasificación ya las escuelas taxonómicas (Fenética, Cladística y Taxonomía evolutiva). In et al. Escacena (Ed.), *Clasificación y Arqueología: Enfoques y métodos taxonómicos a la luz de la evolución darwiniana* (pp. 61–94). Sevilla: Universidad de Sevilla.

García Rivero, D. (2013). *Arqueología y Evolución: a la búsqueda de filogenias culturales*. Sevilla: Universidad de Sevilla.

García Rivero, D. (2016). Darwinian Archaeology and Cultural Phylogenetics (pp. 43–72). https://doi.org/10.1007/978-3-319-25928-4_3

Gartner, R. (2016). *Metadata Shaping Knowledge from Antiquity to the Semantic Web*. Heidelberg: Springer. <https://doi.org/10.1007/978-3-319-40893-4>

Gnoli, C. (2006). Phylogenetic classification. *Knowledge Organization*, 33(3), 138–152.

Gontier, N. (2012). Applied Evolutionary Epistemology: a new methodology to enhance interdisciplinary research between the life and human sciences”. *Kairos: Revista de Filosofia & Ciência*, 4, 7–49. Retrieved from http://www.academia.edu/download/30865351/Applied_Evolutionary_Epistemology.pdf

Gueguen, G., Fonseca, V. M. M. da, Pitti, D. V., Grimouard, C. S., Marques, M., Fonseca, D., ... Sibille-De Grimouard, C. (2013). Toward an International Conceptual Model for Archival Description: A Preliminary Report from the International Council on Archives’ Experts Group on Archival Description. *The American Archivist*, 76(2), 566–583.

Ham, F. G. (1993). *Selecting and appraising archives and manuscripts*. Chicago, IL: Society of American Archivists.

- Heikkilä, T. (2007). The Evolution of Texts: Confronting Stemmatological and Genetical Methods - Research database Tuhat - University of Helsinki. *Variants - Journal of the European Society for Textual Scholarship*.
- Huang, A. (2008). Similarity Measures for Text Document Clustering. In *NZCSRSC 2008, April 2008, Christchurch, New Zealand* (pp. 49–56).
- Hurley, C. (1993). What, If Anything, Is A Function? *Archives & Manuscripts*, 21(2), 208–220.
- ICA. (2016). *RiC-CM-0.1: Records in Contexts: a conceptual model for archival description*. Paris: ICA.
- IP2. (2008). The InterPARES 2 Project Dictionary. Retrieved from http://www.interpares.org/ip2/ip2_terminology_db.cfm
- ISO. (2014). *ISO 21127:2014 - Information and documentation -- A reference ontology for the interchange of cultural heritage information*. Geneva: International Organization for Standardization.
- Janssens, F., Zhang, L., Moor, B. De, & Glänzel, W. (2009). Hybrid clustering for validation and improvement of subject-classification schemes. *Information Processing and Management*. <https://doi.org/10.1016/j.ipm.2009.06.003>
- Janzen, L. M. (1994). *Series: History, Theory and Practice*. University of British Columbia. <https://doi.org/10.14288/1.0087494>
- Jenkinson, H. (1922). *A manual of archive administration including the problems of war archives and archive making*. Oxford: The Clarendon Press.
- Jenkinson, H. (2003). *Selected writings of Sir Hilary Jenkinson*. Chicago, IL: Society of American Archivists.
- Jung, J. Y., Bae, J., & Liu, L. (2008). Hierarchical business process clustering. In *Proceedings - 2008 IEEE International Conference on Services Computing, SCC 2008*. <https://doi.org/10.1109/SCC.2008.69>
- Khanafiah, D., & Situngkir, H. (2006). Visualizing the Phylomemetic Tree: Innovation as Evolutionary Process. *Journal of Social Complexity*, 22(22), 20–30.
- Lantern, D., & Barili, A. (2017). Forensic analysis of deduplicated file systems. *Digital Investigation*, 20, S99–S106. <https://doi.org/10.1016/j.diin.2017.01.008>
- Lecointre, G. (2015). Descent (Filiation). In T. Heams, P. Huneman, G. Lecointre, & M. Silberstein (Eds.), *Handbook of Evolutionary Thinking in the Sciences* (pp. 159–207). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-017-9014-7_9
- Lee, J., Jung, E., Yoon, J., Kim, S., & Jung, H. (2013). Evolutionary Patterns of an Artifact : The Mobile Phone, 1469–1478.
- Letunic, I., & Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*, 1. <https://doi.org/10.1093/nar/gkw290>

- Macedo, L. S. A. de. (2017). Matriz binária RiC: classificação filogenética da informação arquivística. Harvard Dataverse. <https://doi.org/10.7910/DVN/ERE8HL>
- MacNeil, H. (2012). What finding aids do: Archival description as rhetorical genre in traditional and web-based environments. *Archival Science*, 12(4), 485–500. <https://doi.org/10.1007/s10502-012-9175-4>
- Mahner, M., & Kary, M. (1997). What Exactly Are Genomes, Genotypes and Phenotypes? And What About Phenomes? *Journal of Theoretical Biology*, 186(1), 55–63. <https://doi.org/10.1006/jtbi.1996.0335>
- Mai, J.-E. (2011). The modernity of classification. *Journal of Documentation*, 67(5), 710–730. <https://doi.org/10.1108/00220411111145061>
- Mata Caravaca, M. (2015). The concept of archival “sedimentation”: its meaning and use in the Italian context. *Archival Science*, 1–12. <https://doi.org/10.1007/s10502-015-9256-2>
- McCarthy, I. P. (2005). Toward a Phylogenetic Reconstruction of Organizational Life. *Journal of Bioeconomics*, 7(3), 271–307. <https://doi.org/10.1007/s10818-005-5245-5>
- Mesoudi, A. (2011). *Cultural Evolution. How Darwinian Theory can Explain Human Culture & Synthesize the Social Sciences*. University of Chicago Press (Vol. 53). Chicago and London: University of Chicago Press. <https://doi.org/10.1017/CBO9781107415324.004>
- Mesoudi, A. (2016). Cultural Evolution: A Review of Theory, Findings and Controversies. *Evolutionary Biology*, 43, 481–497. <https://doi.org/10.1007/s11692-015-9320-0>
- Mokhtar, U. A., & Yusof, Z. M. (2015). Classification: The understudied concept. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2014.12.002>
- Morrison, D. A. (2014). Is the tree of life the best metaphor, model, or heuristic for phylogenetics? *Systematic Biology*, 63(4), 628–38. <https://doi.org/10.1093/sysbio/syu026>
- Muller, S., Feith, J. A., & Fruin, R. (1898). *Handleiding voor het ordenen en beschrijven van archieven: ontworpen in opdracht van de Vereeniging van Archivarissen in Nederland*. Groningen: Erven B. van der Kamp.
- Niu, J. (2013). Recordkeeping metadata and archival description: a revisit. *Archives and Manuscripts*, 41(3), 203–215. <https://doi.org/10.1080/01576895.2013.829751>
- O’Brien, M., & Lyman, R. (2003). *Cladistics and archaeology*. Salt Lake City: University of Utah Press.
- O’Brien, M., Lyman, R., Mesoudi, A., & VanPool, T. (2010). Cultural traits as units of analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365, 3797–3806. <https://doi.org/10.1098/rstb.2010.0012>
- Oliveira, A. A. de, Ferrara, P., De Rosa, A., Piva, A., Barni, M., Goldenstein, S., ... Rocha, A. (2016). Multiple parenting phylogeny relationships in digital images. *IEEE Transactions on Information Forensics and Security*, 11(2), 328–343.

- Pavlopoulos, G. A., Soldatos, T. G., Barbosa-Silva, A., & Schneider, R. (2010). A reference guide for tree analysis and visualization. *BioData Mining*, 3(1), 1. <https://doi.org/10.1186/1756-0381-3-1>
- Pearce-Moses, R. (2005). *A glossary of archival and records terminology*. Chicago, IL: Society of American Archivists.
- Rieppel, O. (2010). The series, the network, and the tree: changing metaphors of order in nature. *Biology & Philosophy*, 25(4), 475–496. <https://doi.org/10.1007/s10539-010-9216-4>
- Riley, J. (2010). *Glossary of Metadata Standards*. Indiana University Libraries White Professional Development Award.
- Rogers, C. (2015). Diplomatics of born digital documents – considering documentary form in a digital environment. *Records Management Journal*, 25(1), 6–20. <https://doi.org/http://dx.doi.org/10.1108/RMJ-03-2014-0021>
- Rogers, C. (2016). A Literature Review of Authenticity of Records in Digital Systems: From “Machine-Readable” to Records in the Cloud. *Acervo*, 29, 16–44. Retrieved from <http://revista.arquivonacional.gov.br/index.php/revistaacervo/article/view/715>
- Simões, M. da G., & Freitas, M. C. V. de. (2013). A Classificação em Arquivos e em Bibliotecas à Luz da Teoria da Classificação: pontos de convergência e de divergência. *Ponto de Acesso*, 7(1), 81–115.
- Simões, M. da G., Freitas, M. C. V. de, & Rodríguez-Bravo, B. (2016). Theory of Classification and Classification in Libraries and Archives: Convergences and Divergences. *Knowl. Org*, 43(7), 530–539.
- Thomer, A. K., & Weber, N. M. (2014). The phylogeny of a dataset. In *Proceedings of the ASIST Annual Meeting*. <https://doi.org/10.1002/meet.2014.14505101064>
- Ventura, S., & Luna, J. M. (2016). *Pattern Mining with Evolutionary Algorithms*. <https://doi.org/10.1007/978-3-319-33858-3>
- Weber, N. M. (2015). *A framework for analyzing the sustainability of peer produced science commons*. University of Illinois at Urbana-Champaign. Retrieved from <https://www.ideals.illinois.edu/bitstream/handle/2142/88062/WEBER-DISSERTATION-2015.pdf?sequence=1>
- Weber, N. M., & Thomer, A. K. (2016). Modeling Cultural Evolution with Metadata Collections: A Phylomemetic Approach. In *International Conference on Dublin Core and Metadata Applications DC-2016, Copenhagen, Denmark*. Copenhagen.
- Wilkins, J. S. (2009). *Species: A History of the Idea (Species and Systematics)*. Berkeley and Los Angeles, CA: University of California Press.
- Yu, X., & Gen, M. (2010). *Introduction to evolutionary algorithms*. London: Springer. <https://doi.org/10.1007/978-1-84996-129-5>

O PAPEL DOS SITES NOTICIOSOS NA DIVULGAÇÃO DA INFORMAÇÃO CIENTÍFICA E SEU USO NOS PROCESSOS DE EDUCAÇÃO INFORMAL DE CTI

Benjamín Camacho Vega¹, Zeny Duarte de Miranda²

¹Universidade Federal da Bahia, 0000-0002-6416-7905, bennvcv_23@hotmail.com

²Universidade Federal da Bahia, zenydu@gmail.com

RESUMO Tem como objetivo refletir sobre a transição da divulgação de informação científica da mídia tradicional para as novas mídias digitais a partir do início do século XXI, graças ao desenvolvimento das Tecnologias de Informação e Comunicação e a influência exercida pela Internet sobre as pessoas. Neste esquema, os Sites Noticiosos disponibilizam uma imensa quantidade de informação na rede para os indivíduos, daí a sua importância na "sociedade da informação", já que as questões da Ciência, Tecnologia e Inovação, despertam especial interesse nos cidadãos pela oportunidade que lhes oferece de estar bem informados no momento de tomar decisões no cotidiano e permitir-lhes participar no desenvolvimento que ocorrem ao seu redor. Para tanto, o método utilizado foi um estudo bibliográfico a partir de textos tanto da Ciência da Informação como de Divulgação Científica e da Mídia de Massa. Conclui-se que a importância da divulgação por meio dos sites noticiosos é relevante devido ao alcance que têm na sociedade, uma vez que utilizam o potencial da Internet para transmitir informação rápida e massivamente, promovendo assim, a cultura da alfabetização científica.

PALAVRAS-CHAVE *Divulgação de informação científica, Sites noticiosos, Cultura, Mídia de Massa, Sociedade.*

ABSTRACT It aims to reflect about the transition of scientific information divulgation from traditional media to new digital media from the beginning of XXI century, thanks to the development of Information and Communication Technologies and the influence exerted of Internet on people. In this scheme, the News Web Sites offer a huge amount of information on the web for individuals, hence its importance in the "information society", as the issues of Science, Technology and Innovation, arouse special interest in the citizens to the opportunity to offer them to be well informed when making decisions in everyday life and allow them to participate in the development taking place around them. Therefore, the method used was a literature study from texts both of Information Science as Scientific Divulgation and Mass Media. It is concluded that the importance of divulgation through the News Web Sites is relevant because of the range they have on society, since they use the potential of the Internet to transmit information quickly and massively, thus promoting a culture of scientific literacy.

KEYWORDS *Scientific Information Divulgation, News Web Sites, Culture, Mass Media, Society.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

O estudo sobre divulgação da ciência é de extrema importância para a sociedade atual, já que trata de uma área do saber da qual dependemos para promoção de pesquisas relacionadas com a saúde e a cura de doenças, as discussões sobre temas em destaque na sociedade, as tecnologias para o nosso cotidiano, os progressos na computação e na informática, e na mídia contemporânea. Quando a ciência é retirada de seu lugar de produção habitual entre a comunidade científica, e os seus resultados são apresentados ao público geral, surge um interesse especial no conhecimento que dela emana. Assim, as pessoas analisam cuidadosamente o que acontece ao seu redor para criar novos sentidos para suas vidas e reestruturar a sua participação na tomada de decisões na sociedade. Por isso, é necessário refletir sobre o papel que desempenha a mídia como divulgadora destes saberes, porque quando a informação científica é transmitida por estes canais gera-se um impacto maior em comparação com quando se comunica só por os meios tradicionais, enquanto que ajuda a diversas instituições para circular a sua mensagem entre novos públicos. A mídia transforma os resultados das pesquisas em dados digeríveis para todos, já que a sua linguagem é adaptada para as diferentes audiências.

Com ajuda da mídia, as sociedades modernas se transformaram em sociedades da informação, uma vez que as suas práticas comunicativas permitiram trocar ideias, difundir mensagens e produzir novos saberes, os quais se expandem radicalmente a baixo custo a partir do uso dos computadores e da Internet para documentar, interpretar, traduzir e até criar quadros de referência compartilhados úteis para todos (Estrada, 2003). A Internet é o meio divulgador mais importante e os sites de notícias são lugares onde a população pode encontrar informações para estar ciente do que acontece no mundo. No entanto, o desconhecimento no uso desta ferramenta para divulgar, propicia que o cidadão fique distante dos avanços científicos, as consequências e a possibilidade de participar ativamente na tomada de decisões.

A informação científica é um mecanismo que permite a criação de novos conhecimentos e, juntamente com os processos de transferência tecnológica e inovação (CTI), promove o desenvolvimento de uma sociedade ou de um determinado país. Como a ciência tornou-se parte das forças produtivas em muito maior medida do que nunca, hoje se trata de um agente estratégico da mudança nos planos de desenvolvimento econômico, social, educativo e de divulgação. O presente artigo pretende alcançar a divulgação de informação científica adaptada aos novos tempos para evitar um isolamento de conhecimento que prejudique o desenvolvimento comunitário. E, e por sua vez, adotar os sites noticiosos como uma fonte de saberes de ciência, para que assim os indivíduos tomem partido nos acontecimentos diários que determinam a evolução de uma sociedade melhor.

A IMPORTÂNCIA DA DIVULGAÇÃO CIENTÍFICA NA MÍDIA DE MASSA

A divulgação científica propicia ao público conhecer, de primeira mão, a obra feita por pesquisadores nas diferentes áreas do conhecimento, graças a uma variedade de canais, a interpretação básica de seus resultados, além do uso acessível da linguagem para aqueles que não são especialistas. Estrada (2003) a define a divulgação científica como um conjunto de atividades que permitem levar o entendimento científico à sociedade, com o objetivo de que o maior número de pessoas interessadas em se informar, ou entender sobre este tipo de acontecimentos, se beneficiem com mensagens facilmente assimiláveis. Essas ações são uma expressão polivalente que incluem todos os tipos de explicações, transmissões de mensagens e desenvolvimento do pensamento científico, buscando impactar a maior quantidade de

audiência não qualificada para gerar pessoas inquietas com interesse nos progressos da sociedade (Meadows, 1999).

Uma verdadeira divulgação não se centra apenas nas descobertas mediáticas do momento, como as alterações climáticas ou o genoma humano, mas que abrange áreas com teorias bem estabelecidas e socialmente aceites, incluindo temas como os assentamentos humanos e seus problemas diários, até mesmo alguns campos inteiros de estudo, como a física ou a biologia (Haro, 2010). A divulgação científica, através da mídia, não pretende substituir as bases de uma educação escolar rígida que treina especialistas, nem aperfeiçoar seu campo de conhecimento. Ela busca direcionar o conhecimento científico aos seus receptores. Assim, a divulgação massiva da ciência pode ser feita através de qualquer formato existente na mídia e transmitir informações geradas por diversos especialistas (Moirand, 2000).

Por isso, nos últimos anos a informação científica vem proliferando, na medida em que se criou uma vasta gama de canais referentes ao tema, com conteúdo exclusivo que tentam apresentar os estudos realizados por cientistas em todo o mundo (Moirand, 2000). Este processo de massificação é considerado uma alteração cognitiva das opiniões, crenças e práticas que exige tanto da busca de concordância, como da integração dos meios corretos para reduzir a dissonância dos resultados das pesquisas usando códigos que reforcem os modelos e façam chegar a mensagem em sua forma original para os destinatários (Estrada, 2003)

Por sua parte, a mídia digital cria uma ação de divulgar a ciência complementada por diferentes ferramentas que ajudem à população a entender claramente as pesquisas, como o uso de imagens, vídeos e outros materiais multimídia que acurtam a distância espaço-temporal entre transmissores e receptores (Sanz, 2002). As inovações tecnológicas permitem à ciência ganhar espaço na sociedade e reduzir a brecha entre o campo da pesquisa e a geração de novos trabalhos, além de que pode ser tratada cada vez por um número maior de pessoas.

A CULTURA DA CIÊNCIA CRIADA PELA MÍDIA

Até finais do século XX, a transmissão de conhecimento científico era realizada pela publicação de artigos em magazines, 'journals' ou memórias de congressos, mas muitas eram de acesso restrito só para os usuários (Jianmin, 2005). Com ajuda da mídia, os pesquisadores e centros acadêmicos, fizeram chegar a sua mensagem aos diferentes públicos, reduzindo a 'ciência perdida', que é aquela que passa despercebida, principalmente das ciências humanas e sociais (Sanz, 2002). A rápida aceitação da mídia no processo de transmitir notícias científicas à cidadania está ligada ao interesse demonstrado pelos destinatários sobre este tema, já que, segundo Jianmin (2005), 85% dos habitantes das grandes cidades recebem informações de CTI pelo menos uma vez por semana através de jornais, revistas, rádio, televisão ou Internet.

A divulgação massiva da ciência se tornou uma prática comum para difundir as pesquisas, tanto entre a própria comunidade científica, como entre a população, pois assim é possível construir conhecimentos mais abrangentes entre todas as partes, diversificando a transmissão de conhecimento para a sociedade (Pontes, 2005). Mas, para que a mídia cumpra como divulgadora e potencialize o uso da informação, primeiramente deve contextualizar aquilo que vai publicar para ter elementos de certeza, fundamentar suas ideias indo à origem da fonte, criar um discurso científico verídico com material que seja aplicável no dia a dia das pessoas, afastando-se de razões comerciais para que o público não o consuma só por

morbidade. Também, se devem incluir códigos na informação para ampliar a noção da realidade, utilizando no processo todos os recursos disponíveis, inclusos os tecnológicos (Orlandi, 1999).

A mídia precisa ser capaz de demonstrar veracidade nas notícias, não é confiar simplesmente em sua aparência. Cada um dos espaços deve ganhar o reconhecimento a partir do material que difunde, tornando-o digno de seguir e demonstrando o seu valor no desenvolvimento científico. Só assim, a divulgação da ciência, por canais midiáticos, pode ter uma utilidade na sociedade, provocando o "efeito da informação científica", quando os conteúdos de um meio de circulação tradicional são removidos para ocupar um lugar no cotidiano da população, transformando os dados em conhecimento útil (Orlandi, 1999).

INTERNET: O NOVO MEIO DE DIVULGAÇÃO DA CIÊNCIA

Segundo Haro (2010), o fato de que as revistas especializadas continuem a sua política de exclusão e aumentando seu preço em razão de 15% anual, enquanto que a mídia dá pouco espaço à informação científica por suas restrições de formato, gerou a opção de pesquisadores e institutos pela busca de outros canais, evitando o 'empobrecimento do conhecimento' (Castells, 2009). O agente que está revolucionando esse campo, e aproveitando todas as oportunidades possíveis, é a Internet, uma vez que desde o surgimento da World Wide Web é possível uma maior conexão e acessibilidade de conteúdos (Guerrero, 2003). A rede não deve ser vista apenas como um veículo essencial para divulgar, mas também como alternativa para os modelos existentes e aproximar à ciência e o conhecimento aos indivíduos.

A Internet comunica, transmite e amplia a gama dos saberes gerais, mas também favorece o surgimento de novos canais para atingir um público maior, graças ao desenvolvimento das novas tecnologias, potencializando sua prática entre os diversos grupos sociais, incluindo os acadêmicos e os cientistas (Pontes, 2005). A própria UNESCO (2000) garante que a Internet é a ferramenta que constitui uma inovação fundamental na divulgação da informação científica, capaz de melhorar a 'sociedade do conhecimento' e trazer novas opções na busca por aumentar o nível de participação dos cidadãos.

A rede é um meio com uma carga valiosa de conhecimento e apoio cultural inesgotável, com espaços abertos para a heterogeneidade de vozes e pensamentos. É o lugar da controvérsia, com geração de conflitos entre as partes que a integram, uma vez que aqui convergem quem defende a pluralidade de ideias e os críticos por desinformar ou propagar notícias que não são fáceis de confiar (Guerrero, 2003). Para Castells (2009), a Internet é tão real como a própria vida, e pode ser descrita a partir dos problemas que surgem diariamente, menosprezando alguns valores humanos, mas também é usada infinitamente em benefício das pessoas, pois é uma extensa fonte de dados que permite agir no mundo de forma criativa.

Para considerar a Internet como o principal meio de informação científica, é necessário trabalhar na educação e na cultura dos internautas, já que passa por bombardeio informativo, na maioria das vezes sem preparos tanto para receber quanto para identificar, podendo cair facilmente no erro (Castells, 2009). O internauta, com bases para lidar com esses problemas pode participar no processo de difusão através da identificação dos materiais que são para seu benefício, permitindo-lhe gerar suas próprias conclusões e reproduzir materiais que o ajudem para interagir com outros usuários, discutir suas ideias e promover o ciclo da divulgação científica (Codina, 2009). Assim, o internauta tem a possibilidade de escolher

entre uma vasta gama de conteúdos e entrar em espaços que lhe resultem de maior importância de acordo com seus interesses e curiosidades, ajudando a Internet a se expandir e representando um salto qualitativo em relação aos seus competidores, deixando de lado a forma passiva de divulgar graças à possibilidade de interagir entre todas as partes (Codina, 2009).

OS SITES NOTICIOSOS TORNAM POPULAR O CONHECIMENTO

Os sites noticiosos são páginas ligadas às agências ou empresas de jornalismo que têm a finalidade de transmitir notas de atualidade e interesse geral buscando manter informado ao público do que acontece diariamente na sociedade (Codina, 2009). Com o boom da Web 2.0, a divulgação científica tem mais oportunidades de atingir um público maior e penetrar em setores, distantes, há poucos anos atrás, já que permite compartilhar, comparar e promover a ciência em novos contextos. Internet World Stats (2010), indica que o 54% dos internautas, aproximadamente 3.010 milhões, usam os sites noticiosos como sua principal fonte informativa, seguidos pelas redes sociais com um 39%. Por isso, pode-se dizer que estes canais são parte da contínua expansão da rede e ajudam as pessoas a refletirem sobre as pesquisas científicas. Na esfera da divulgação, os sites noticiosos são inovadores, usam uma linguagem fácil de seguir para os não especialistas e sem limitações, mantendo grande potencial para mudar a forma como as pessoas obtêm informação (Sanz, 2002).

O trabalho desses sites torna mais fácil a tarefa de encontrar informações na rede, sua presença é maior do que outro tipo de páginas, possuem maior número de seguidores e seus arquivos são fáceis de consultar (Pontes, 2005). Entre suas vantagens está a adaptabilidade e o fato de que apresenta informações de forma criativa devido à variedade de públicos que os visitam, usando uma interação que procura captar a atenção do usuário para oferecer uma variedade de opções dentro do seu próprio espaço. Com essa versatilidade, podem divulgar a ciência, criar conteúdos relacionados, e incentivar ao usuário a expressar suas ideias participando em fóruns de discussão ou gerando debates com outras pessoas, permitindo aumentar a curiosidade, a convivência e a cultura científica da cidadania.

Como tal, os sites não são uma fonte em si. São um canal que facilita enormemente o acesso à informação, assim que as notas nesses espaços devem estar preparadas para ser fácil de encontrar pelos usuários que o desejem, e é preciso cuidar de todos os aspectos para evitar que se tornem o maior divulgador de pseudociência (Guerrero, 2003). Os sites podem atuar como os principais divulgadores da informação gerada pelos pesquisadores, uma vez que eles são os encarregados de acercar ao público para aquilo que produzem (Jianmin, 2005). O maior impacto dos sites na tarefa de divulgar é o fato de que eles podem até definir a informação mais importante e digna de destaque, por serem líderes de opinião, pela popularidade com que contam, pelos seguidores confiantes no que publicam.

CONSIDERAÇÕES FINAIS

A Internet está chamada a ser a grande divulgadora de informação científica e tecnológica. Com o boom das TIC, cada vez há mais pessoas gerando propostas para obter o máximo benefício dos cenários digitais. A publicação na rede de todo tipo de pesquisas que abrangem as áreas da vida diária, tem incentivado o elo do cidadão comum por consumir esse conhecimento, uma vez que a partir dele é possível participar das tomadas de decisões.

Para que os sites noticiosos sejam utilizados como ferramentas de divulgação científica, tem-se que desenvolver conteúdos culturais pertos dos cidadãos, com recursos multimídia que permitam alcançar a maior quantidade de público e explicar o contexto sem cair no superficial. Assim, as notas especializadas podem ser usadas como um complemento no desenvolvimento social das pessoas. Segundo Pontes (2005), entre as principais ações que devem ser realizadas para melhorar a divulgação científica através dos sites noticiosos, está a de visibilizar a grande quantidade de pesquisas que são esquecidas por ser consideradas complementares e criar códigos de uso comum a todas as pessoas para que a informação seja compreensível.

Estes espaços são uma ferramenta inovadora para criar um novo processo cultural, sendo necessário evitar a saturação de informação como fontes pouco confiáveis, sem contribuição ao conhecimento ou, que, unicamente têm infinidade de dados a colocarem o usuário perto da pseudociência. O ato de divulgar por esses canais deve observar o cumprimento do papel integrador de processos democráticos de um país e a discussão sobre os progressos sociais e a ética nos dados científicos. Portanto, é importante que nos diferentes âmbitos a procura desta informação se adapte ao dia a dia para evitar um isolamento de conhecimento que prejudique o desenvolvimento comunitário, e por sua vez se apoie nos sites noticiosos como uma verdadeira fonte de saberes da ciência, tudo com o objetivo de que os indivíduos tomem partido nos acontecimentos da sociedade que determinam as suas perspectivas para a evolução de uma melhor nação.

REFERÊNCIAS BIBLIOGRÁFICAS

- Castells, M. (2009). *A sociedade em rede. A era da informação*. São Paulo, Brasil: Paz e Terra.
- Codina, L. (2009). Ciencia 2.0: Redes sociales y aplicaciones en línea para académicos. *Hipertext.net*, (7), 13.
- Estrada, L. (2003). La Divulgación de la Ciencia. Educación, apostolado o...? Em J. Tonda, A. Sánchez, N. Chávez (Ed.). *Antología de la Divulgación de la Ciencia en México* (pp. 47-63). México, México: DGDC-UNAM.
- Guerrero, R. (2003). Internet como medio de divulgación: De Eolo a Pandora?. *Quark*, (28), 15.
- Haro, S. (2010). ¿Comunicación o promoción de la ciencia? *El Cerebro Digital*, (15), 21-23.
- Internet World Stats. (2016). *World Internet users and population stats*. Bogota, Colombia: Miniwatts Marketing Group. Recuperado de <http://www.internetworldstats.com/>
- Jianmin, L. (2005). Estudio sobre la popularización de la ciencia en las ciudades modernas. *Quark*, (37), 72-82.
- Meadows, A. J. (1999). *A comunicação científica*. Brasilia, Brasil: Lemos.
- Moirand, S. (2000). Formas discursivas da divisão de saberes na mídia. *RUA: Núcleo de Desenvolvimento da Criatividade*, (6), 9-24.
- Orlandi, E. (1999). *Análise do Discurso: princípios e procedimentos*. São Paulo, Brasil: Pontes.

Pontes, A. (2005). Aplicaciones de las tecnologías de la información y de la comunicación en la educación científica. *Eureka*, 2(1), 330- 343.

Sanz, E. (2002). Ciencia digit@l: Ciencia para todos en Internet. *Mediatika*, (10), 184-197.

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. (2000). *Informe mundial sobre la comunicación y la información 1999-2000*. Recuperado de <http://unesdoc.unesco.org/images/0011/001159/115908sb.pdf>

A PRESERVAÇÃO DA MEMÓRIA DOS GEOPARQUES NO BRASIL: O CENTRO DE MEMÓRIA DO QUADRILÁTERO FERRÍFERO EM MINAS GERAIS

Cátia Rodrigues Barbosa¹, Mônica Elisque Carmo², Leonardo Barreto Oliveira³

¹*Ufmg-Eci, 0000-0002-2684-504X, catiarbp1@hotmail.com*

²*Ufmg-Eci /Iphan-MG, 0000-0001-7641-5075, m_elisque@yahoo.com.br*

³*Ufmg-Eci /Iphan, 0000-0003-3357-9111, barretoleonardo6@gmail.com*

RESUMO Este artigo apresenta a fase inicial da pesquisa sobre a preservação da memória organizacional do Geoparque no Brasil proposto para o Quadrilátero Ferrífero, no Estado de Minas Gerais. Trata-se de uma pesquisa sobre a organização e acesso ao patrimônio cultural do Geoparque. Como metodologia foi feito um mapeamento das informações sobre o tema, no portal Capes e em sites de órgãos públicos, empresas privadas e organizações não governamentais. Portanto é uma pesquisa de natureza documental, exploratória e descritiva, com abordagem qualitativa, com a finalidade de conhecer os estudos e pesquisas que têm sido desenvolvidos referentes à preservação da memória dos Geoparques no Brasil. Embora o conceito de Geoparque seja de fácil compreensão e está consolidado, a implementação e inclusão do geoparque do Quadrilátero Ferrífero, no Estado de Minas Gerais, na Rede Global de Geoparques requer o desenvolvimento de técnicas e metodologias apropriadas. Nesta primeira fase da pesquisa, se busca melhor entendimento sobre o tema e o contexto no qual se insere. Neste momento, pode-se constatar que a publicação sobre a preservação da memória organizacional dos Geoparques, no Brasil, é praticamente inexistente. Portanto, os resultados deste estudo, subsidiarão as demais fases da pesquisa sobre a criação do Centro de Memória do Geoparque Quadrilátero Ferrífero, em Minas Gerais, que demanda uma pesquisa devido à escassez de literatura. Visto que, as pesquisas sobre a implantação de Geoparques no Brasil encontra-se em estágio inicial. Entendemos, assim, ser este o momento para a discussão e formulação de propostas visando à preservação da memória organizacional do Geoparque Quadrilátero Ferrífero, Minas Gerais.

PALAVRAS-CHAVE *geoparque, patrimônio cultural, centro de memória, memória organizacional, organização do conhecimento.*

ABSTRACT This article presents the initial phase of the research on the preservation of the organizational memory of the Geopark in Brazil, and has as boundary the Iron Quadrangle, in the State of Minas Gerais. It is a research on the organization and access to the cultural heritage of the Geopark. As a methodology, a mapping of the information about the theme was carried out on the Capes portal and on sites of public agencies, private companies and non-governmental organizations. Therefore, it is a documental, exploratory and descriptive research, with a qualitative approach, with the purpose of knowing the studies and researches that have been developed regarding the preservation of the memory of the Geoparks in Brazil. Although the concept of Geopark is easy to understand and already consolidated, its implementation and inclusion in the Global Geopark Network requires the development of appropriate techniques and methodologies for its consolidation. In this first phase of the research, it constitutes a better understanding about the theme and the context in which it is inserted. At this moment, it can be verified that the publication on the preservation of the organizational memory of the Geoparks in Brazil is non-existent. Therefore, the results of this study will support the other phases of the research on the creation of the Quadrilátero Ferrífero

Geopark Memory Center in Minas Gerais, which demands extensive research due to the scarcity of literature, since research and initiatives on the implementation of Geoparks In Brazil they are in an early stage. Therefore, it is the ideal moment for the discussion and formulation of proposals aiming at the preservation of the organizational memory of the Iron Quadrangle Geopark, Minas Gerais.

KEY WORDS *geopark, cultural patrimony, memory center, organizational memory, organization of knowledge*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Este artigo traz como problema de pesquisa questões sobre a organização e o acesso ao patrimônio cultural dos Geoparques no Brasil, sobretudo o questionamento sobre a preservação da memória organizacional. Na delimitação da pesquisa foi selecionado o Quadrilátero Ferrífero que submeterá à apreciação, novamente, um dossiê à Divisão de Ciências Ecológicas e da Terra da UNESCO, se candidatando ao título de Geoparque. Pretende, com este pleito fazer parte da Rede Global de Geoparques.

A memória das organizações tem sido objeto de estudos e reflexões, evidenciando a sua importância para a educação, pesquisa e tomada de decisões. Este levantamento inicial da pesquisa consiste em verificar se existe alguma iniciativa sobre a criação de Centros de Memória dos Geoparques no Brasil, se existe, quais são e como funcionam?

Este estudo tem como objetivo apresentar o estado da arte sobre a preservação da memória dos Geoparques no Brasil. Como metodologia buscou a realização de um mapeamento iniciado com o levantamento bibliográfico sobre o tema preservação da memória dos Geoparques no Brasil. A documentação foi coletada na base de periódicos nacionais e internacionais da Capes, bem como em *sites* de órgãos públicos, empresas privadas e organizações não governamentais que se referem ao tema.

Por meio deste procedimento investigativo buscou-se conhecer e identificar a existência de publicações referentes a estudos e pesquisas que tratam da gestão da informação com a finalidade da preservação da memória dos Geoparques.

O motivo da escolha do tema preservação da Memória do Geoparques, especificamente do Quadrilátero Ferrífero decorre da observação que as iniciativas para a proteção desses acervos, apesar de atual e relevante, tem sido pouco fundamentada documentalmente demandando o estabelecimento de critérios, metodologia e técnicas apropriadas.

As instituições públicas, privadas e do terceiro setor, produzem, acumulam, armazenam e organizam documentação em formatos e suportes diversificados, fruto da produção da atividade intelectual do seu corpo técnico no cumprimento de suas atribuições. Esse acervo preservado expressa a missão, o objetivo e as atividades da instituição e mais importante que isto, reflete sua imagem. Por esse motivo, muitas instituições têm investido nas Unidades de Documentação, como lugar de preservação da memória organizacional, muitas denominadas Centros de Memória.

A partir dos anos 1980 começam a surgir, no Brasil, os centros de memória, ligados a organizações públicas e privadas e também àquelas sem finalidade lucrativa, chamadas de terceiro setor. [...] Embora não se disponha de dados mais precisos sobre o fenômeno, possível atrelar o crescimento numérico dos centros de memória a certos elementos conjunturais que trazem à tona a necessidade de preservação da cultura e identidade das organizações. [...] (CAMARGO & GOULART, 2015, p.63)

A constituição do Centro de Memória do Geoparque Quadrilátero Ferrífero deve ser pautada em uma política sólida, pré-definida e fundamentada na missão do Geoparque.

GEOPARQUE

A partir deste breve prólogo sobre do termo memória organizacional defendemos a importância de se definir estratégias para a preservação dos registros produzidos pelo Quadrilátero Ferrífero, como fonte de informação fundamental para o seu desenvolvimento sustentável, econômico, educacional e para a pesquisa científica. De suma importância a compreensão do conceito de Geoparque, bem como a produção de registros desse patrimônio natural.

Os Geoparques são áreas geográficas únicas com limites claramente definidos, cujas paisagens de importância geológica internacional são administrados a partir de um conceito holístico de proteção, educação e desenvolvimento sustentável. Abrangem um notável patrimônio geológico, de incomum importância científica, raridade e beleza, revestidos de valores educacionais e /ou econômico. Tem como missão proteger e difundir bens de natureza arqueológica, ecológica, histórica e culturais da região. O patrimônio geológico de um Geoparque Global utiliza seu patrimônio em conexão com todos os demais aspectos do patrimônio natural e cultural da sua área, com a finalidade de intensificar a conscientização e compreensão dos principais problemas da sociedade. Tem como fundamento central o uso sustentável dos recursos da Terra, mitigação dos efeitos da mudanças climáticas e redução dos riscos relacionados aos desastres naturais.

METODOLOGIA

Trata-se de pesquisa de natureza documental, exploratória, sob uma abordagem qualitativa, que tem como finalidade a realização de uma revisão de literatura, em um processo de busca, análise e descrição de um corpo do conhecimento objetivando responder a uma pergunta específica. “Literatura que cobre todo o material relevante, que é escrito sobre um tema: livros, artigos de periódicos, relatos governamentais, teses, dissertações e outros tipos” (USP, 2015).

Deste modo, o corpus da pesquisa se constitui em um análise qualitativa das publicações recuperadas a partir das consultas que foram feitas do seguinte modo:

Na primeira etapa, fase inicial da pesquisa, a primeira página na internet a ser examinada foi da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) e realizada a leitura da publicação Geoparques do Brasil: propostas – Vol. 1 (SCHOBENHAUS; SILVA. 2012).

Dando sequência ao levantamento das informações sobre o tema, foram consultados os sites da: Companhia de Pesquisa e Recursos Minerais/Serviço Geológico do Brasil (CPRM/SGB); Instituto do

Patrimônio Histórico e Artístico Nacional (IPHAN); Geopark Araripe; Universidade Regional do Cariri (URCA)/Museu de Paleontologia; Universidade Federal de Minas Gerais/Instituto de Geociências da (UFMG/IGC); Geopark Quadrilátero Ferrífero.

Na segunda etapa da pesquisa, foi por meio do portal Capes, com a busca pelo termo “Geoparque”, com expansão de resultados, sem refinamento dos resultados. Essa investigação tão ampla, deve-se ao fato do conceito Geoparque ser muito recente. Por esse motivo, a expectativa de um número escasso de recuperação de informações sobre o tema, era pressentido. Na terceira etapa foi realizada a busca por assunto nos mesmos moldes da segunda etapa. As estratégias utilizadas foram: busca avança, sem data limite, qualquer tipo de material, qualquer idioma. No entanto, foram combinados os termos: “Geoparque e Centro de Memória”; “Geoparque e Organização da Informação”; “Geoparque e Gestão da Informação”; “Geoparque e Museu”; Geoparque e Biblioteca”; “Geoparque e Arquivo”. Conforme demonstrado na Figura 1.

TERMO	RECURSO (QUANTIDADE)	DATAS- LIMITE	OBSERVAÇÕES
GEOPARQUE	Artigos (21); Atas de Congresso (15); Dissertações e Teses (14); Livros (2)	2005-2016	Apenas uma publicação na Ciência da Informação
GEOPARQUE E CENTRO DE MEMÓRIA			Nenhum resultado encontrado
GEOPARQUE E ORGANIZAÇÃO DA INFORMAÇÃO	Livro (1) Artigo (1)	2014-2015	Os temas abordam educação, turismo e geoconservação
GEOPARQUE E GESTÃO DA INFORMAÇÃO	Dissertação (3)	2012-2013	Abordam conforto térmico, desempenho cognitivo. Base de dados com inventário, informação geográfica.
GEOPARQUE E MUSEU	Artigo (2)	2014-2015	Processo aprendizagem nos museus; centro de pesquisa e geoturismo
GEOPARQUE E BIBLIOTECA	Artigo (2)	2010-2012	Uso turístico; geossítio
GEOPARQUE E ARQUIVO			Nenhum resultado encontrado

Figura 1: Segunda e terceira etapa do levantamento bibliográfico no Portal Capes.

Após esse levantamento bibliográfico inicial, pode-se constatar claramente que não há estudos e pesquisas referentes a iniciativas voltadas a preocupação ou a intenção da preservação da memória organizacional dos Geoparques, sobretudo no Brasil. No entanto, fazem-se necessários, levantamentos complementares da bibliografia sobre o tema, em bases especificamente voltadas ao campo da Ciência da Informação. Visando ampliação do levantamento, que será realizado na próxima etapa da pesquisa.

A primeira etapa da pesquisa constituiu-se em uma melhor apreensão sobre o tema, como meio de obter dados e informações preliminares sobre o tema, o contexto no qual se insere, quais instituições estão

envolvidas e qual a missão de um Geoparque e quais seria a relevância da pesquisa para a contribuição para evolução do conhecimento humano e científico.

A segunda etapa e terceira etapa da pesquisa buscou-se o aprofundamento sobre as discussões que estão sendo realizadas no meio acadêmico sobre o tema, bem como suas reflexões e considerações. Qual está sendo o direcionamento dos estudos e pesquisas científicas

No próximo estágio a pesquisa se constituirá em investigar o mapeamento de revisão de literatura sobre as iniciativas de implantação de Centros de Memória do Geoparques no Mundo, especificamente no âmbito da Ciência da Informação, buscando estudos e pesquisas sobre as iniciativas da preservação da memória dos Geoparques, por meio da gestão e organização do conhecimento, possibilitando por meio das tecnologias da informação e comunicação.

Esses estudos subsidiarão o projeto sobre a criação do Centro de Memória do Geoparque Quadrilátero Ferrífero, em Minas Gerais.

RESULTADOS

O presente trabalho elegeu o Quadrilátero Ferrífero como modelo para tratar da organização do conhecimento para o Centro de Memória do Geoparque, pelo fato de ter sido previsto inicialmente na sua Estrutura de Governança, como Projetos e Programas Estruturantes Prioritários, a Casa do Geoparque e Centro de Informações.

Com a finalidade promover a preservação da memória organizacional para fins de pesquisa e tomada de decisão, durante a elaboração do dossiê de candidatura do Geoparque Quadrilátero Ferrífero, apresentada à UNESCO durante a Conferência Internacional de Geoparques Globais realizada no Canadá em setembro de 2014, foi apresentado um plano preliminar de ação, que consta “uma lista de temas estruturantes” contemplando “Programas e Projetos de Ensino e Pesquisa”, descritos no quadro abaixo. (Geoparque Quadrilátero Ferrífero).

Como se pode observar na Figura 2, a Estrutura de Governança, já havia sido previsto uma unidade informacional, denominada Casa do Geoparque e Centro de Informação.



Figura 2: Estrutura de Governança. Casa do Geopark e Centro de Informações

Fonte: Geopark Quadrilátero Ferrífero. (2016). Geoparque. Recuperado em 11 setembro, 2016 de <http://www.geoparkquadrilatero.org/>.

Deste modo, os Programas e Projetos apresentados a UNESCO pelo Geoparque Quadrilátero Ferrífero, tem o mesmo entendimento no que diz respeito a preservação da memória organizacional do Geoparque, bem como a disponibilização das informações para fins de estudos, pesquisas e tomada de decisões.

Localizado na região centro-sudeste do Estado de Minas Gerais, ao sul da capital Belo Horizonte e abrangendo uma área em torno de 7.000 km², “internacionalmente reconhecido como um importante terreno pré-cambriano com significativos recursos minerais, em especial ouro e ferro” (AZEVEDO, et al. 2012, p. 185).

[...] A designação “Quadrilátero” é em função do arranjo geométrico de sua morfoestrutura e foi utilizada em 1933 pelo geólogo Luiz Flores de Moraes Rego, para definir a área onde estão concentradas “As jazidas de ferro do centro de Minas Gerais”, em artigo assim intitulado (MACHADO, 2009). Desde as publicações produzidas pelo grupo de geólogos do DNPM e do U.S. Geological Survey formado em 1945, para mapeamento geológico da área e avaliação de suas reservas minerais, a região passou a ser conhecida como Quadrilátero Ferrífero. [...].(AZEVEDO et al. 2012, p. 186)

Na proposta de elaboração do Geoparque Quadrilátero Ferrífero foram identificados cinquenta e cinco sítios de interesse natural e cultural, emblemáticos da história geológica e da mineração, da ecologia e da cultura. Em 1997 foi instituída uma comissão com a finalidade de identificar, descrever e divulgar os sítios do patrimônio geológico brasileiro. (AZEVEDO, et al. 2012, p. 195).

A criação dos Geoparques, proposto pela UNESCO, está intrinsicamente ligados à conservação geológica, a sustentabilidade, a educação e a pesquisa científica. A diversidade dos sítios acima elencados, naturalmente reflete o volume de informações produzidas, em diversos suportes e formatos. Apesar de existirem algumas unidades informacionais constituídas, na maioria dos casos, museus. Percebe-se a ausência de um tratamento dos “registros dessas memórias organizacionais”, com a finalidade da gestão da informação que possibilite reunir, organizar, divulgar e tornar acessível o conhecimento cultural, científico e tecnológico produzido pelo Geoparque.

CONCLUSÕES

Este momento de implantação de geoparques no Brasil, apresenta-se uma conjuntura favorável a consolidação de uma unidade de informação híbrida – Centro de Memória do Geoparque do Quadrilátero Ferrífero, com acervos diversificados, cuja principal finalidade seja a gestão da informação, independentemente de seu suporte e formato, com a finalidade de preservação da memória, para fins de educação, pesquisa e tomada de decisão.

Nesta pesquisa realizada não foram encontrados nenhum resultado de iniciativas voltadas a preservação da memória organizacional, invariavelmente os estudos (artigos, capítulos de livros, monografias, teses, dissertações) de maior relevância sobre o tema abordam estudos e pesquisas sobre aos geosítios, envolvendo geodiversidade, ecoturismo, negócios.

Este artigo teve como finalidade a realização de um levantamento preliminar sobre o tema Centro de Memória do Geoparque Quadrilátero Ferrífero, que demanda uma extensa pesquisa à escassa literatura sobre o tema, e conclui que devidamente fundamentado e baseado em uma investigação minuciosa, possibilitará sua aplicabilidade visando a construção do conhecimento e contribuição para a sociedade.

Os resultados demonstram que no Brasil, as pesquisas e iniciativas sobre Geoparques encontram-se em estágio embrionário. Por conseguinte, os procedimentos adotados em relação a gestão da informação referente a memória dos Geoparques, demanda iniciativa das instituições para sua formulação e consolidação.

REFERÊNCIAS BIBLIOGRÁFICAS

Alves-Mazzotti, A. J. (2002). A revisão bibliográfica em teses e dissertações: Meus tipos inesquecíveis – o retorno. In Bianchetti, L. & Machado, A. M. N. (Org.). *A bússola do escrever: desafios e estratégias na orientação de teses e dissertações* (p.p. 25-44). São Paulo: Cortez.

Azevedo, U. R., Machado, M.M.M., Castro, P. T. A., Renger, F. E., Treviso, A., & Beato, D. A. C. (2012). Quadrilátero Ferrífero (MG). In Schobbenhaus, C., & Silva, C.R. (Orgs.), *Geoparques do Brasil: propostas* (p.p. 183-220). Rio de Janeiro: CPRM.

Brilha, J. (2012). A rede global de geoparques nacionais: Um instrumento para a promoção internacional da geoconservação. In Schobbenhaus, C., & Silva, C. R. *Geoparques do Brasil: propostas* (Vol. 1, p.p.29-38). Rio de Janeiro: CPRM.

Camargo, A. M., & Goulart, S. (2005). *Centros de memória: Uma proposta de definição* (Coleção Sesc Culturas). São Paulo: Edições Sesc São Paulo.

Castells, M. (1999). *A sociedade em rede: V.1. A era da informação: Economia, sociedade e cultura* (6ª ed.). São Paulo: Paz e Terra.

Cervo, A. L., & Bervian, P. A. (2002). *Metodologia da pesquisa científica*. (5ª ed.). São Paulo: Prent Hall.

Companhia de Pesquisa de Recursos. (2016). *Geoparque*. Recuperado em 02 agosto, 2016 de <http://www.cprm.gov.br/>.

Davenport, T. H. (1998). *Ecologia da informação: Por que só a tecnologia não basta para o sucesso na era da informação*. São Paulo: Futura.

Geopark Quadrilátero Ferrífero. (2016). *Geoparque*. Recuperado em 11 setembro, 2016 de <http://www.geoparkquadrilatero.org/>.

Instituto do Patrimônio Histórico e Artístico Nacional. (2016). *Declaração dos direitos à memória da terra*. Recuperado em 12 setembro, 2016 de http://portal.iphan.gov.br/uploads/temp/Declaracao_Internacional_dos_Direitos_a_Memoria_da_Terra.pdf

Pereira, C. de S. (2013, setembro). Memória Organizacional: Conceitos e práticas em construção. *Anais do Encontro da Associação Nacional de Pós Graduação e Pesquisa em Administração*, Rio de Janeiro, Rio de Janeiro, Brasil, 37. Recuperado em 10 setembro, 2016 de http://www.anpad.org.br/admin/pdf/2013_EnANPAD_ADI471.pdf

Organização das Nações Unidas para a Educação, a Ciência e a Cultura. (2016). *Geopark*. Recuperado em 12 agosto, 2016 de <http://www.unesco.org/new/en/natural-sciences/environment/earth-sciences/unesco-global-geoparks/>

Schobbenhaus, C., & Silva, C. R. (2012). O papel do serviço geológico no Brasil na criação de geoparques e na conservação do patrimônio geológico. In: Schobbenhaus, C., & Silva, C. R. (Orgs.) *Geoparques do Brasil: Propostas* (Vol. 1, p.p. 11-28). Rio de Janeiro: CPRM.

Schobbenhaus, C., & Silva, C. R. (Orgs.). (2012). *Geoparques do Brasil: Propostas*. Rio de Janeiro.

Universidade de São Paulo. Instituto de Psicologia. Biblioteca. (2015). *Tipos de revisão de literatura*. Recuperado 01 outubro, 2016 de <http://www.fca.unesp.br/Home/Biblioteca/tipos-de-revisao-de-literatura.pdf>

O PROCESSO DE REFERÊNCIA NAS BIBLIOTECAS UNIVERSITÁRIAS BRASILEIRAS COM USO DAS REDES SOCIAIS

Ana Roberta Sousa Mota

Universidade Federal da Paraíba, 0000-0001-6663-504, anarobertamota@gmail.com

RESUMO Com a criação da internet os serviços das bibliotecas se modificaram e a proporção da explosão bibliográfica cresceu à medida que as tecnologias de comunicação se desenvolveram a grande rede mundial ampliou as possibilidades de comunicação e investigação. A cultura de convergências dos medias e o uso das Tecnologias Digitais nas Bibliotecas Universitárias para disseminar informações revela o objetivo do artigo, em analisar a utilização das redes sociais como ferramenta de comunicação junto aos utilizadores. Aborda conceitos biblioteca universitária e redes sociais. Enfatiza a importância das redes sociais e o uso das novas tecnologias de comunicação e informação para disseminação de informações relativos a produtos e serviços de uma biblioteca universitária. A metodologia empregada baseou-se na aplicação de questionário online com bibliotecários de referência e levantamento em sítios na internet de bibliotecas universitárias brasileiras que utilizam redes sociais. No estudo verificou-se que com a evolução das bibliotecas e advento de TIC, as bibliotecas universitárias têm se inserindo paulatinamente nos novos paradigmas da informação e comunicação em plataformas digitais de forma proativa utilizando as redes sociais para o acesso e disseminação de informações da própria biblioteca. O processo de comunicação para a transação de referência ocorre predominantemente com a utilização de e-mail e telefone fixo, as redes sociais são utilizadas pela maioria das BUB, sendo o Facebook e o Twitter as mais empregadas, principalmente na divulgação de produtos e serviços, ou seja, para marketing e estratégias de promoção/publicidade. Contudo, para além de ferramenta de marketing digital dessas bibliotecas universitárias, vemos que os serviços de referência também estão sendo utilizados para atender à necessidade informacional dos usuários.

PALAVRAS-CHAVE *Biblioteca Universitária, Redes Sociais, Serviço de Referência, Tecnologia da Informação e Comunicação.*

ABSTRACT With the creation of the internet, libraries' services have changed and the proportion of the bibliographical explosion has increased as communication technologies have developed the great world wide network has expanded the possibilities of communication and research. The culture of convergence of media and the use of Digital Technologies in University Libraries to disseminate information reveals the purpose of the article in analyzing the use of social networks as a tool for communication with users. It addresses university library concepts and social networks. It emphasizes the importance of social networks and the use of new communication and information technologies to disseminate information about products and services of a university library. The methodology used was based on the application of an online questionnaire with reference librarians and survey on internet sites of Brazilian university libraries that use social networks. In the study it was verified that with the evolution of the libraries and the advent of TIC, university libraries have gradually inserted themselves in the new paradigms of information and communication in digital platforms in a proactive way using the social networks for the access and dissemination of information of the own library . The communication process for the referral transaction occurs predominantly with the use of e-mail and landline telephones, social networks are used by most BUBs,

with Facebook and Twitter being the most employed, mainly in the dissemination of products and services, That is, for marketing and promotion / advertising strategies. However, in addition to the digital marketing tool of these university libraries, we can see that reference services are also being used to meet users' informational needs.

KEYWORDS *University Library, Social Networks, Reference Service, Information Technology and Communication.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Com a criação da internet os serviços das bibliotecas se modificaram e a proporção da explosão bibliográfica cresceu à medida que as tecnologias de comunicação se desenvolveram. A grande rede mundial ampliou as possibilidades de comunicação e investigação, Guimarães (2005, p. 159) ressalta que “com a introdução World Wide Web, a internet se tornou um dos principais recursos de comunicação no mundo atual.” Enquanto Souto (2003, p. 74) afirma que “com o uso do computador a literatura científica aumentou nos últimos anos. As facilidades de acesso permitidas pela internet contribuíram para o aumento dessa produção científica”. Fatores como estes contribuíram para a proliferação de: textos, artigos, livros digitais, resenhas, relatórios, pré-prints, teses, dissertações, chats, e-mails, videoconferências e sites, além dos multimédias, definidos pela APDSI (2007, p. 70) como “tecnologias da informação que permitem a utilização simultânea de vários tipos de dados digitais (textuais, visuais e sonoros) no interior de uma mesma aplicação ou de um mesmo suporte” o mesmo tempo em que a internet facilitou o acesso à informação e a comunicação tornou também à medida e a organização deste conteúdo um desafio para sociedade.

De acordo com Castells (1999, p. 46) uma característica importante da sociedade informacional, ainda que não esgote todo o seu significado é, “a lógica de sua estrutura básica em redes, o que explica o uso do conceito de 'sociedade em rede’”. Ao surgir a sociedade em rede torna-se possível o desenvolvimento das novas tecnologias da informação que, “agruparam-se em torno de redes de empresas, organizações e instituições para formar um novo paradigma sociotécnico” (Castells, 1999, p.77) cujos aspectos centrais, representam a base material da sociedade da informação.

A apresentação deste contexto fez com que muitos questionassem sobre a permanência das bibliotecas, devido às facilidades que a tecnologia proporciona, porém, comungamos com Cunha (2000, p. 75) ao afirmar que “as bibliotecas sempre acompanharam e venceram os novos paradigmas tecnológicos”, o papel das bibliotecas de armazenar, disponibilizar e facilitar o acesso à informação ampliou-se para o ambiente virtual, fazendo com que as mesmas se adaptem as mudanças e aperfeiçoem os serviços oferecidos.

A relevância das informações para o desenvolvimento social e organizacional é notória por toda sua ênfase na sociedade contemporânea. A utilização adequada de informações que circulam nas instituições configura-se em um contexto representando uma vantagem competitiva e desenvolvimento estratégico. Inovações tecnológicas e diversas possibilidades de comunicação têm contribuído para a interação entre ambientes, indivíduos e informações, permitindo maior coesão e alcance de resultados desejados.

A biblioteca universitária definida como “um sistema orgânico de atividades que envolvem a produção e registros de conhecimentos, recursos materiais e humanos necessários para servir de suporte às funções básicas em instituições de nível superior” (Mangue, 2007, p. 26) é parte do ambiente universitário e estão aderindo cada vez mais as TIC como ferramenta de disseminação do conhecimento e dos serviços oferecidos à comunidade acadêmica. Nessa perspectiva, as redes sociais, definidas por Marteleto e Tomaél (2005, p.75) como “um conjunto de pessoas (ou organizações ou outras entidades sociais) conectadas por relacionamentos sociais, motivados pela amizade e por relações de trabalho ou compartilhamento de informações e, por meio dessas ligações, vão construindo e reconstruindo a estrutura social”, disponíveis na web 2.0, a exemplo do o Twitter, Facebook, LinkedIn e Slideshare tem se configurado de grande relevância na disseminação de informações e serviços oferecidos para a comunidade acadêmica.

Um novo comportamento é exigido aos bibliotecários, o de mediador entre serviços em plataformas digitais e os utilizadores, Silva (2010, p.25) afirma que a mediação é destinada, em especial, aos

info-incluídos e os born digital ou nativos da internet. Os serviços de informação multiplicaram-se e complexificaram-se até se instalarem na internet e, aqui, a função mediadora de comunicação no espaço social e a função mediadora institucional, com as estratégias comunicacionais específicas dos respectivos actores e agentes, não desapareceram, nem tendem, necessariamente, a desaparecer, mas podem transformar-se e coexistir com um emergente novo tipo de mediação – deslocalizada ou dispersa (na internet/redes conexas), institucional, colectiva, grupal, pessoal e até anónima, interactiva e colaborativa. Possíveis traços caracterizadores, entre os quais importa destacar a interação e os processos colaborativos, sociais, de participação cívica, espontânea e militante.

Diante do contexto apresentado pelas novas tecnologias da informação e comunicação em plataformas digitais, é de extrema importância que as bibliotecas se adaptem a esta cultura de convergência de novas mídias e que seus profissionais se aperfeiçoem para oferecer novos serviços em novas plataformas e ferramentas que utilizam da inteligência coletiva e compartilhamento de informações. Para isto, o objetivo deste artigo é analisar a utilização das redes sociais como ferramenta de comunicação junto aos utilizadores.

METODOLOGIA

A metodologia empregada baseou-se na pesquisa bibliográfica e levantamento em sítios na internet de bibliotecas universitárias brasileiras que utilizam redes sociais. No estudo verificou-se que com a evolução das bibliotecas e advento de TIC, as bibliotecas universitárias têm se inserindo paulatinamente nos novos paradigmas da informação e comunicação em plataformas digitais de forma proativa utilizando as redes sociais para o acesso e disseminação de informações da própria biblioteca. Para perceber a compreensão dos bibliotecários identificar esta afirmativa, perguntou-se em questionário *online* a 40 bibliotecários de bibliotecas universitárias brasileiras (Universidade Federal de Goiás, Universidade Federal de Mato Grosso do Sul, Fundação Universidade Federal da Grande Dourados, Universidade Federal de Mato Grosso, Universidade de Cuiabá, Universidade de Brasília, Universidade Católica Dom Bosco, Universidade Católica de Brasília, Universidade Federal do Rio Grande do Norte, Universidade Federal da Bahia, Universidade Federal da Paraíba, Fundação Universidade Federal do Vale do São Francisco, Universidade Federal de Pernambuco, Universidade Federal Rural de Pernambuco, Universidade Federal Rural do Semi-Árido, Universidade Federal de Campina Grande, Universidade Federal do Pará, Universidade Nilton Lins, Universidade da Amazônia, Universidade

Federal do Amazonas, Universidade Federal do Acre, Universidade Federal Rural da Amazônia, Universidade Federal de Roraima, Fundação Universidade Federal do Tocantins, Universidade Federal de Lavras, Universidade Federal de Minas Gerais, Universidade Federal de Viçosa, Universidade Federal de São Carlos, Fundação Universidade Federal do Abc, Universidade Estadual de Campinas, Universidade Federal de São Paulo, Pontifícia Universidade Católica do Rio de Janeiro, Universidade Federal do Rio Grande do Sul, Universidade Federal de Santa Catarina, Universidade de Santa Cruz do Sul, Universidade Positivo, Universidade Tecnológica Federal do Paraná, Universidade Federal do Paraná, Universidade de Caxias do Sul e Universidade do Vale do Rio dos Sinos): Quais meios utilizados para comunicação com o usuário? Se os bibliotecários utilizam redes sociais? Quais as redes são utilizadas e com que finalidade.

RESULTADOS

A respeito dos meios utilizados para a comunicação com o usuário. A maioria respondeu ser por e-mail (98%), seguido do telefone fixo (93%), rede social (63%), computador (55%), página oficial da biblioteca na internet (53%), telefone inteligente (23%), *chat* (20%), fax (10%) e, tablete (5%). No entanto, percebe-se que o e-mail e o telefone fixo são os principais meios de comunicação com a biblioteca. As redes sociais são utilizadas por mais da metade das bibliotecas, juntamente com os computadores e *website* oficial, enquanto, *smartphones*, *chat*, fax e tablete são pouco utilizados.

O uso de redes sociais pelo serviço de referência, aponta positivamente 55% das bibliotecas inquiridas e negativamente 45% das bibliotecas.

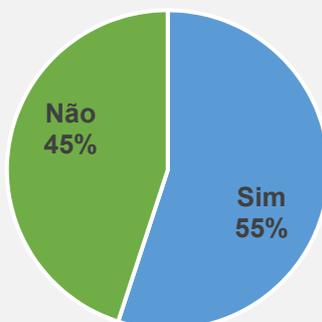


Gráfico 1- Uso de redes sociais pelo serviço de referência N=40

Fonte: dados da pesquisa, 2016

Em caso positivo pediu-se para que o respondente indicasse quais redes sociais eram utilizadas nas bibliotecas em que atuam. As respostas, demonstram que: 24 usam Facebook, 10 Twitter, três Flickr e Google + e, um Academia.edu. Na categoria outros, identificou-se: 2 blogs e 1 whatsapp.

Para compreender a finalidade do uso dessas ferramentas, indagou-se acerca das funções do uso das redes sociais. Sendo assim, o resultado obtido indica que 96% afirmam utilizar as redes sociais para divulgar seus produtos e serviços, 68% disseram ser para interagir com o usuário ou atrair usuários, 64% apontaram que é para se comunicar, 32% para buscar informações e manter-se atualizado e, 4% buscar parcerias. Assim, de acordo com os resultados, revela-se que a maioria dos bibliotecários utilizam as

redes sociais para divulgação de produtos e serviços, ao passo que a minoria busca parcerias através das redes sociais.

Tabela 1 - Finalidade para o uso de redes sociais em bibliotecas universitárias

Funções para utilização de redes sociais pelo serviço de referência.	Frequência absoluta	Frequência relativa (%)
Interagir com o usuário	17	68
Divulgar seus produtos, serviços	24	96
Buscar informações e manter-se atualizado	8	32
Buscar parcerias	1	4
Comunicar-se	16	64
Atrair usuários	17	68

Fonte: dados da pesquisa, 2016

CONCLUSÕES

O processo de comunicação para a transação de referência ocorre predominantemente com a utilização de e-mail e telefone fixo, as redes sociais são utilizadas pela maioria das BUB, sendo o Facebook e o Twitter as mais empregadas, principalmente na divulgação de produtos e serviços, ou seja, para marketing e estratégias de promoção/publicidade. Contudo, para além de ferramenta de marketing digital dessas bibliotecas universitárias, vemos que os serviços de referência se utilizam das redes sociais para atender à necessidade informacional dos usuários.

REFERÊNCIAS BIBLIOGRÁFICAS

- APDSI. (2007). *Glossário da sociedade da informação*. Retrieved from <http://www.apdsi.pt/index.php/news/138/122/Glossario-da-Sociedade-da-Informacao---Versao-2007>
- Castells, M. (1999). *A sociedade em rede*. São Paulo: Paz e Terra.
- Cunha, M. B. (2000). Construindo o futuro : a biblioteca universitária brasileira em 2010. *Ciência Da Informação*, 29(1), 71–89.
- Guimarães, Â. de M. (2005). Internet. In *Introdução às fontes de informação*. Belo Horizonte: Autêntica.
- Mangue, M. V. (2007). *Consolidação do processo de informatização em sistemas de bibliotecas universitárias da África do Sul, Brasil e Moçambique*. Universidade Federal de Minas Gerais, Belo Horizonte. Retrieved from http://www.bibliotecadigital.ufmg.br/dspace/bitstream/handle/1843/VALA-74QHWR/doutorado__manuel_valente_mangue.pdf;jsessionid=8077D5623900D245E1D0893F640FF1EB?sequence=1
- Marteletto, Regina Maria; Tomaél, M. I. (2005). Metodologia de Análise de Redes Sociais (ARS). In *Métodos qualitativos de pesquisa em Ciência da Informação*. São Paulo: Polis.
- Souto, L. F. (2003). Recuperação de informações bases de dados: usos de tesouros. *Transinformação*, 15(1), 73–81.

A ORGANIZAÇÃO DE DOCUMENTOS TÉCNICOS DE AERONAVES À LUZ DA CLASSIFICAÇÃO FACETADA

Ana Patrícia Guimarães¹, Renato Rocha Souza²

¹*Instituto Brasileiro de Informação em Ciência e Tecnologia, 0000-0003-4718-5797, anapat_guimaraes@yahoo.com.br*

²*Fundação Getúlio Vargas, 0000-0002-1895-3905, rsouza.fgv@gmail.com*

RESUMO Apresenta uma proposta de classificação de documentos, com base nos princípios da classificação facetada e sua possível aplicabilidade nos documentos da área de manutenção aeronáutica, no âmbito da Força Aérea Brasileira. Discorre sobre a Classificação Facetada e o contexto de organização dos documentos e as especificidades das publicações da área de manutenção e operação de aeronaves, usadas na Força Aérea Brasileira. Trata-se de uma pesquisa exploratória e de caráter qualitativa, que utilizou como metodologia a análise de documentos, em busca da sistematização das informações sobre a operação e a manutenção de aeronaves. Conclui que o uso da classificação facetadas nesses documentos é viável para representação e organização da informação uma vez que possibilita a estruturação do processo de disseminação seletiva da informação.

PALAVRAS-CHAVE *Classificação Facetada, Organização da Informação, Operação e manutenção de aeronaves, Força Aérea Brasileira.*

ABSTRACT The paper presents a proposal for classifying documents using a faceted approach, and its application to the organization of the documents of the Brazilian Air Force. It presents the theoretical bases derived from the faceted classification in particular. It also presents the context for the creation and management of the documents about some models of the fleet. It involves an exploratory study with a qualitative approach, and it developed, as methodology, the analysis of the documents and the synthesis of the information about the operation and maintenance of the selected aircraft models. We conclude that the use of the facet classification in these documents is viable to representation and organization of information, and helps structuring the information selective dissemination process.

KEYWORDS *Faceted Classification, Information Organization, Operation and maintenance of aircraft, Brazilian Air Force.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Este trabalho problematiza sobre uma proposta de classificação de documentos, com base nos princípios da classificação facetada e sua possível aplicabilidade nos documentos da área de manutenção e operação de aeronaves, usados no âmbito da Força Aérea Brasileira (FAB). O trabalho tem como objetivo a sistematização das informações sobre operação e manutenção aeronáutica, tendo como suporte as teorias da Organização do Conhecimento e da Classificação, especificamente a Classificação

Facetada, aplicando-se métodos e técnicas da Ciência da Informação. Para tal, foi utilizado como metodologia a análise dos documentos utilizados na FAB e a posterior extração dos termos para definição das categorias e facetadas.

As atividades de manutenção e operação das aeronaves demandam uma atualização constante dos profissionais que executam as missões impostas no dia a dia. O volume expressivo de informação gera a necessidade de organizar e classificar os documentos, com vistas à recuperação da informação. Assim, faz-se necessário a criação de mecanismos que atendam as demandas informacionais desses profissionais.

Para que seja efetuada a operação e a manutenção de aeronaves e equipamentos, bem como o suprimento de material aeronáutico, os técnicos necessitam consultar e cumprir as determinações constantes nos documentos, chamados Publicações Técnicas. Entende-se por Publicação Técnica, os documentos cujo conteúdo seja aplicável à operação, manutenção, suprimento, inspeção e modificação do material aeronáutico ou do material bélico.

Dada a importância da manutenção e operação de aeronaves, torna-se imprescindível a existência de um acervo organizado e adequado para o uso dos operadores. Apesar do controle e organização desses documentos, observa-se que ainda não é realizada uma classificação temática dos documentos. Nesse sentido, e sob o olhar da Ciência da Informação, parece propício buscar meios de aprimorar o serviço de organização desse acervo.

Por assim ser, a classificação facetada apresenta-se como uma alternativa para a organização e recuperação da informação contida nesses documentos. Ressalta-se que estudos dessa natureza são de fundamental importância para os campos da Teoria das Classificações e Organização da Informação e do Conhecimento.

CONTEXTUALIZAÇÃO

A classificação é uma atividade que permeia a vida do ser humano e integra seu cotidiano, tornando-se parte constitutiva da sociedade. Dentre os diferentes tipos de classificação bibliográfica, foi escolhido para essa pesquisa o uso da classificação facetada, que segundo Duarte é um tipo de classificação “capaz de identificar características comuns a diversas categorias de um assunto, organizando-o em partes denominadas de facetadas” (DUARTE, 2010, p. 47).

A classificação facetada foi introduzida por Ranganathan, que propôs uma fragmentação de assuntos em cinco categorias fundamentais: Personalidade, Matéria Energia, Espaço e Tempo – o PMEST. Essa classificação trouxe flexibilidade aos sistemas classificatórios, que até então eram lineares.

Na FAB são utilizados inúmeros documentos nos processos de operação e manutenção de aeronaves, denominados Publicações Técnicas que são definidas como “documento cujo conteúdo seja aplicável à operação, manutenção, suprimento, inspeção e modificação do material aeronáutico ou do material bélico”. (MCA 5-2, 2009, p. 15). O uso e o cumprimento das diretrizes descritas nas publicações técnicas é de caráter obrigatório, uma vez que são imprescindíveis para a segurança de voo.

As publicações técnicas são divididas de acordo com sua tipologia. Os manuais técnicos e diretivas técnicas são os mais representativos, sendo esses utilizados na realização dessa pesquisa. São

documentos revisáveis e dinâmicos, que podem sofrer modificações em seu conteúdo, quando são identificadas melhorias, adequações ou mudanças de procedimentos. Possui apresentação padronizada e texto estruturado em função do assunto. Sempre apresentam na página de rosto a data da edição básica e a data e número de revisão, que indica a quantidade de atualizações que o documento já teve.

Uma única aeronave possui diversos volumes de publicações técnicas que compõe o seu acervo, visando atender todos os seus itens e sistemas (hidráulico, elétrico, pneumático, combustível, etc.). Esses documentos são distribuídos a todos as organizações militares que operam as aeronaves, sendo o Parque de Material Aeronáutico dos Afonsos (PAMAAF), responsável pelo gerenciamento, controle e distribuição dos documentos no âmbito da FAB.

METODOLOGIA

Para que se cumprisse o objetivo geral dessa pesquisa foram analisadas as publicações técnicas usadas na operação e manutenção de aeronaves, bem como os processos de gerenciamento e disseminação desses documentos na FAB. Após a identificação das características físicas e temáticas dos documentos, foi possível definir as categorias e facetas, tendo como base os princípios da classificação facetada.

Essa pesquisa pode ser caracterizada como uma pesquisa de abordagem qualitativa, de caráter exploratório, que utiliza como metodologia a análise documental, que consiste no uso de documentos como fonte sistemática de estudo.

A primeira etapa da pesquisa consistiu na análise de documentos técnicos do acervo da FAB. A FAB possui uma frota com diversos tipos de aeronaves e cada tipologia possui um conjunto vasto de documentos que trazem informações específicas sobre aquela aeronave, considerando suas características e peculiaridades.

Diante dessa diversidade na coletânea de publicações técnicas da FAB, fez-se necessário delimitar o corpus do estudo, a começar pela escolha das aeronaves. Nessa pesquisa foram selecionados dois tipos: as aeronaves Bandeirantes C-95 e Brasília C-97, ambas fabricadas pela empresa Embraer, e em seguida, selecionada a coletânea de publicações técnicas que compõe o acervo das respectivas aeronaves.

Dentre as publicações existentes no acervo foram escolhidos para análise os documentos da tipologia ordens técnicas e diretivas técnicas. A escolha deve-se ao fato de que estes documentos são os mais usuais e amplamente utilizados em todas as atividades de manutenção. As ordens técnicas e diretivas técnicas possuem uma padronização na sua apresentação, o que facilita a análise dos documentos em seu conteúdo temático. Os dados foram coletados *in loco* na Biblioteca Técnica do Parque de Material Aeronáutico dos Afonsos, sediado no Rio de Janeiro.

A partir dessa análise, foi possível identificar características físicas e temáticas das publicações técnicas e extrair termos que permitiram a obtenção de um vocabulário. A análise dos documentos se deu através dos principais pontos de acessos. Foram eles: capa, título, folha de rosto, notas, tabela de conteúdo e lista de páginas efetivas.

Outro aspecto importante foi a realização de um estudo acerca dos processos de gerenciamento, recebimento, cadastro e distribuição dos documentos para os usuários de diversas organizações militares do Brasil. Foram mapeadas todas as aeronaves Bandeirantes C-95 e Brasília C-97 da frota da

FAB e as organizações militares que fazem uso dessas aeronaves, bem como a localização geográfica das aeronaves.

A segunda fase da pesquisa corresponde à análise dos princípios da classificação facetada para a elaboração da estrutura classificatória dos documentos. Nessa etapa foram adotadas as diretrizes apresentadas por Barbosa (1972), no qual a autora define as etapas para a organização de um sistema de classificação facetada. As etapas foram adaptadas nesse estudo, considerando-se o objetivo da pesquisa.

Primeiramente foi realizada a definição de termos, provenientes da análise dos principais pontos de acessos dos documentos. A partir da obtenção do vocabulário, foram identificadas e categorizadas as facetas. Para a geração das categorias afetas às publicações técnicas, foram usadas as categorias fundamentais da classificação facetada de Ranganathan, o PMEST. Em seguida, foram estabelecidas as facetas correspondentes a cada uma dessas categorias.

A primeira fase, de identificação dos documentos e mapeamento da aeronaves ofereceram elementos descritivos e temáticos que deram subsídio para a segunda fase da pesquisa, e definição dos termos, identificação e categorização das facetas.

RESULTADOS

Após análise das características que norteiam os processos internos da FAB de gerenciamento e distribuição das publicações técnicas, obteve-se como resultado que as aeronaves Bandeirantes C-95 e Brasília C-97 obtêm apoio logístico e suprimento de materiais através do PAMAAF. Foram encontradas 16 Organizações militares que operam as aeronaves, no território brasileiro, dos tipos Bases Aéreas, Parques de Material e Unidades de Ensino e Pesquisa.

A partir desse resultado é possível observar a complexidade e a importância do processo de distribuição dos documentos. Dessa forma, identificou-se um elemento que é importante na caracterização dos documentos de operação e manutenção de aeronaves: o **Público Alvo**.

Outro aspecto relevante foi a análise da tipologia dos documentos. Conforme mencionado na metodologia, existem diversos tipos de documentos e nessa pesquisa, foram analisados os documentos do tipo Manual Técnico e Diretiva Técnica. A partir das características de cada um desses tipos de documentos, observou-se que outro elemento a ser considerado é **Tipo de Documento**.

Em relação às características das publicações técnicas em si, observou-se que cada documento faz referência a uma determinada aeronave. Assim, existe, por exemplo, um Manual de Voo para a aeronave Bandeirantes C-95 e um Manual de Voo para a aeronave Brasília C-97. Isso acontece para todos os documentos. Foi identificado então o terceiro elemento, que versa sobre o tipo de **Aeronave**. Essa é uma das principais características do documento, que vai direcionar todo o sistema classificatório.

Outra característica importante é que cada documento possui uma finalidade específica, ou seja, possui aplicabilidade na aeronave. O quarto elemento identificado foi **Finalidade** do documento, que versa sobre o uso do documento. E por fim, a última característica diz respeito a atualização

do conteúdo dos documentos, ou seja, em qual versão o documento se encontra. Nesse trabalho, foi adotada a nomenclatura **Status** do documento, que diz respeito à atualização.

Dadas as principais características dos documentos que compõe o acervo da FAB, foi elaborada uma estrutura classificatória, com base no método de categorização, utilizando-se das categorias fundamentais de Ranganathan, conforme ilustrado na Figura.

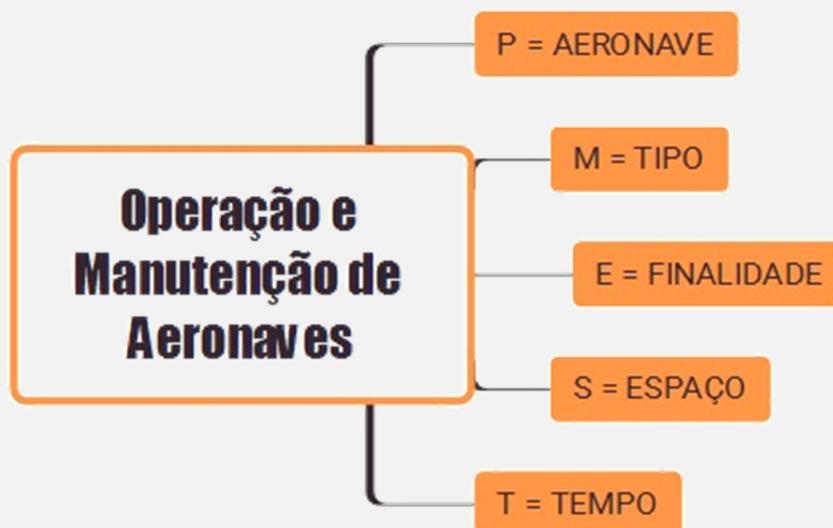


Figura 1: Categorias do sistema classificatório

A análise da Figura 1 nos permite associar as categorias fundamentais de Ranganathan no contexto dos documentos de operação e manutenção de aeronaves da FAB. Assim sendo:

- a) Personalidade = Aeronave – versa sobre o modelo da aeronave
- b) Matéria = Tipo de documento – versa sobre a tipologia dos documentos
- c) Energia = Finalidade do Documento – versa sobre o uso desses documentos
- d) Espaço = Público Alvo – versa sobre as unidades militares da FAB
- e) Tempo = Status do documento – versa sobre a atualização dos documentos

Cada uma das categorias possui suas respectivas facetas. Foram criadas 13 facetas, que compõe as 5 categorias do sistema classificatório, conforme detalhado a seguir.

A categoria Personalidade é a característica que distingue o assunto ou um domínio do conhecimento. Nesse trabalho são abordados dois tipos de aeronaves. Assim a categoria personalidade é representada por duas facetas referentes ao assunto Aeronave: **Bandeirantes C-95** e **Brasília C-97**. Essas duas facetas correspondem ao pilar das atividades de operação e manutenção realizadas pelo PAMAAF.

A Categoria Matéria é composta pelo material e suas características e propriedades. Nessa categoria foi considerada a tipologia dos documentos, sendo definidas as facetas: **Manual Técnico** e **Diretiva Técnica**.

A categoria Energia corresponde às ações e processos. Desse modo, formaram-se cinco facetas baseadas na finalidade do documento: **Manutenção, Reparo, Descritivo, Informativo e Operação**.

A categoria Espaço corresponde aos locais para onde são designados a documentos e quais as unidades fazem uso desses documentos. Trata-se da categoria correspondente ao público alvo, que tem como facetas: **Oficinas e Unidades Aéreas**.

A categoria Tempo refere-se a períodos. Nesse trabalho, tempo diz respeito à versão dos documentos e seu status de atualização, característica essa intrínseca das publicações técnicas. As facetas são: **Edição Básica e Revisão**.



Figura 2: Categorias e facetas do sistema classificatório

A figura 2 ilustra todas as facetas do sistema classificatório nas suas respectivas categorias. A partir dessa estrutura, é possível constatar que existem subsídios para a construção de um sistema classificatório facetado para os documentos de operação e manutenção de aeronaves da FAB.

CONCLUSÃO

O presente trabalho problematizou sobre uma proposta de classificação de documentos, com base nos princípios da classificação facetada, visando otimizar a recuperação da informação no âmbito da FAB. A partir da análise dos documentos de manutenção e operação das aeronaves Bandeirantes C-95 e

Brasília C-97 e do estudo acerca do processo de gerenciamento desses documentos na FAB, foi possível obter subsídios para a categorização e identificação das facetas

Frente à diversidade e especificidades dos diferentes tipos de aeronaves e de publicações técnicas relacionadas, foi necessário selecionar tanto os tipos de aeronave, como os tipos de documentos. Nessa pesquisa levou-se em consideração o âmbito e público interno da FAB, os operadores, mecânicos, inspetores e documentos utilizados pelas equipes de manutenção das aeronaves mencionadas. Consequentemente, para outros modelos de aeronaves, pode-se obter resultados diferentes.

A partir da análise das publicações técnicas da FAB em seus elementos descritivos e temáticos, foi possível concluir que cada documento é direcionado para um tipo de aeronave e por esse motivo, a distribuição e o acesso aos documentos deve ser adequada ao público a que se destina.

A classificação facetada foi a que apresentou elementos básicos para a sistematização das informações contidas nas publicações técnicas das aeronaves. Das etapas executadas para a estruturação do esquema classificatório as mais representativas para subsidiar a classificação facetada foram a definição de termos e assuntos, a categorização das facetas.

Conclui que o uso da classificação facetadas nos documentos do acervo da FAB é viável para representar e organizar a informação. E também é possível vislumbrar a possibilidade de estruturação do processo de disseminação seletiva da informação aos usuários do trabalho.

Cabe destacar que essa pesquisa se deteve nas etapas de análise dos documentos para definição das categorias e facetas. Não houve até o momento o detalhamento das facetas e não foram inseridos os termos correspondentes as facetas, sendo essa uma atividade futura, sugerida para continuidade e aprimoramento da pesquisa em questão.

REFERÊNCIAS BIBLIOGRÁFICAS

BARBOSA, Alice P. (1972). Classificações Facetadas. *Ciência da Informação*, 1(2), 73-81.

BRASIL. Diretoria de Material Aeronáutico e Bélico (2009). *Manual do suprimento de publicações do SISMA e do SISMA B*: MCA 5-2. Rio de Janeiro: Comando da Aeronáutica.

DUARTE, Elisabeth Andrade. (2010) Classificação facetada: um olhar sobre a construção de estruturas semânticas. *Revista Digital de Biblioteconomia e Ciência da Informação*, 7(2), 46-58.

RANGANATHAN, S. R. (1967) *Prolegomena to Library Classification*. Bombay: Asia Publ. House.

III CONGRESSO ISKO ESPANHA-PORTUGAL

XIII CONGRESSO ISKO ESPANHA

Universidade de Coimbra

ISBN 978-972-8627-75-1