



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



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



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.

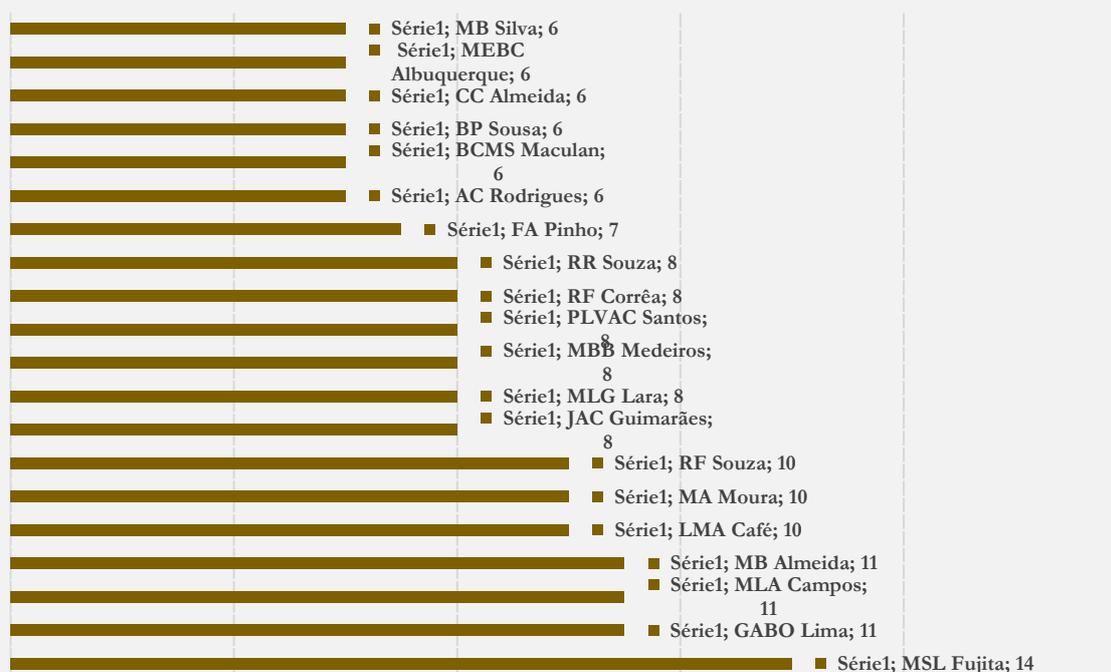


Gráfico 1: Distribuição das Autorias dos Artigos sobre ORC no Brasil (2011-2014)

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.

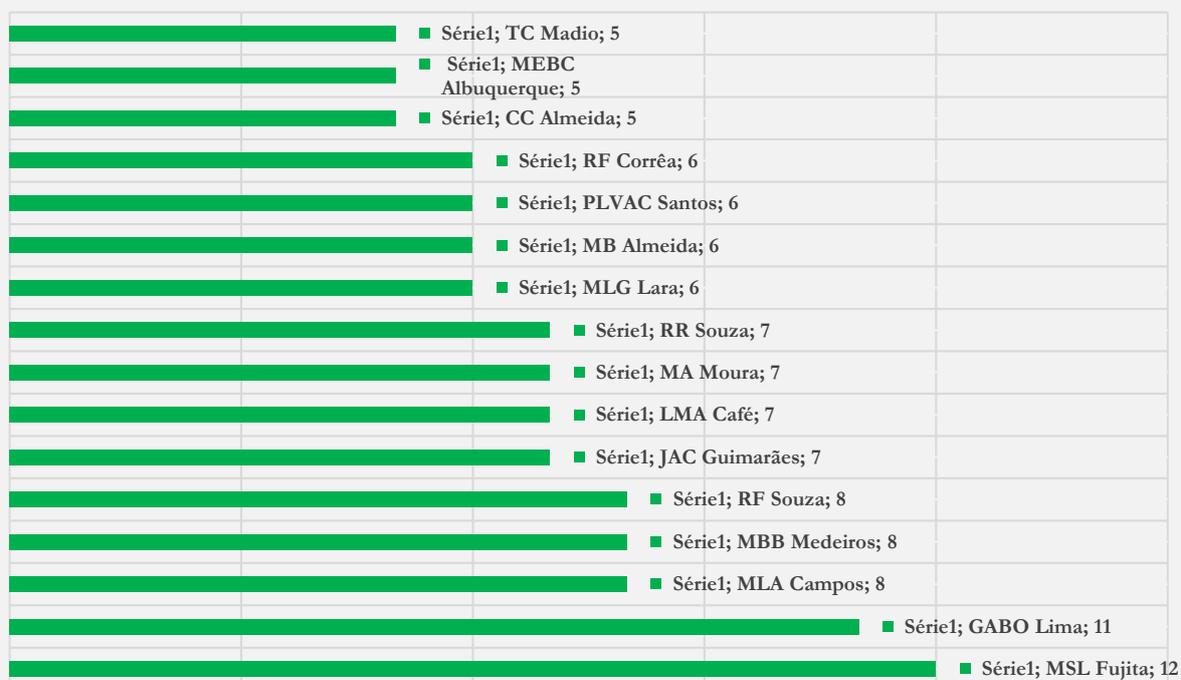


Gráfico 2: Distribuição dos Coautores dos Artigos sobre ORC no Brasil (2011-2014)

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.