



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



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*.