

TESAUROS CONCEITUAIS E ONTOLOGIAS DE FUNDAMENTAÇÃO: ASPECTOS INTERDISCIPLINARES

Jackson da Silva Medeiros¹, Maria Luiza de Almeida Campos²

¹Bibliotecário – Mestrando em Ciência da Informação, Universidade Federal Fluminense, Niterói, RJ

²Doutora em Ciência da Informação, Professora Adjunta da Universidade Federal Fluminense/Programa de Pós-Graduação em Ciência da Informação/UFF, Niterói, RJ

Resumo

Este trabalho tem por objetivo apresentar características existentes nos modelos conceituais de tesauros conceituais e ontologias de fundamentação como aspectos interdisciplinares para representação de domínios de conhecimento dentro da Ciência da Informação e da Ciência da Computação, respectivamente. Estes instrumentos, os quais estão pautados em modelos conceituais construídos a partir de teorias independentes de domínio, são modelos representativos de domínios e atuam como sistemas de organização e representação do conhecimento com a finalidade de recuperação de informação.

Palavras-Chave: Tesauros conceituais; Ontologias de fundamentação; Representação de domínios de conhecimento; Organização do conhecimento.

Abstract

This paper aims to presents existing characteristics in conceptual models of conceptual thesauri and foundational ontologies as interdisciplinary aspects for representation of domains of knowledge within the Information Science and Computer Science, respectively. These instruments, which are guided by conceptual models constructed from independent field theories, are models representative of domains and act as knowledge and representation systems for the purpose of information retrieval.

Keywords: Conceptual thesauri; Foundational ontologies; Representation of domains of knowledge; Knowledge organization.

1 Introdução

Este trabalho apresenta uma discussão oriunda da revisão de literatura realizada em pesquisa de Mestrado, ainda em andamento, junto ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal Fluminense. A abordagem objetiva tratar a representação de domínios de conhecimento, a partir da

observação da construção de modelos conceituais de tesouros conceituais e ontologias de fundamentação, na Ciência da Informação e na Ciência da Computação, respectivamente.

As disciplinas que trabalham com modelos de representação de conhecimento, em geral, o fazem a partir de modelos próprios para o fim desejado, utilizando modelos criados a partir de um problema específico, ocasionando dúvida quando a adequabilidade destes modelos em outras situações.

Por este motivo existe a necessidade da criação de modelos baseados em teorias independentes de domínio, ou seja, que tomem por base teorias e métodos capazes de representar domínios a partir de abordagens que não são construídas para resolver determinada situação, compreendendo questões teóricas e metodológicas que possam ser utilizadas na construção de modelos independentemente de problemas específicos.

Neste sentido, existem estudos que investigam a representação de um domínio a partir da sistematização de teorias que possibilitam a construção de modelos conceituais e sistemas de organização do conhecimento. Com efeito, o processo de modelização de domínios de conhecimento permite a utilização, a partir de uma dada realidade, de sistematizações teórico-metodológicas baseadas em princípios subtendidos ao ato de modelizar, como o método de raciocínio, o objeto de representação, as relações entre os objetos e as formas de representação (CAMPOS, 2004).

Outrossim, buscamos por elementos que revelem aspectos interdisciplinares entre Ciência da Informação e Ciência da Computação no que tange a utilização de instrumentos de representação de domínios, tomando como base a criação de modelos conceituais de tesouros conceituais e ontologias de fundamentação. De fato, os aspectos interdisciplinares que envolvem tesouros conceituais e ontologias de fundamentação estão inseridos na complexidade de representação de domínios de conhecimento, sendo importante notar o que está subjacente aos instrumentos,

ou seja, a questão teórico-metodológica observada em suas construções.

Para tal fim, deve ser observado como se dá a construção de modelos conceituais de tesouros conceituais e ontologias de fundamentação para elaboração dos referidos instrumentos, uma vez que sua constituição parte da criação de modelos conceituais de representação de domínios de conhecimento, com abstração de seus elementos constituintes. Com isto, apontamos que a construção desses tipos de sistemas deve ser pautada em modelos conceituais e nas relações existentes entre os objetos, já que somente o tratamento linguístico acarreta problemas de precisão na recuperação de informação.

2 Tesouro e tesouro conceitual

A palavra tesouro tem seu significado originário do latim e do grego, significando tesouro. Em 1852, Peter Mark Roget publica o *Thesaurus of English Word and Phrases*, dando popularidade à palavra. Já em 1950, Hans Peter Luhn utiliza o termo *Thesaurus* para dar nome a um sistema de palavras autorizadas utilizado na IBM Research Center. Logo que constroi seu instrumento, Luhn percebe que havia necessidade dos termos constituintes explicitarem relações entre si, melhorando seu sistema de recuperação de informação (CAMPOS; GOMES, 2006).

No que se refere às definições estabelecidas, a UNESCO, em 1973, através do programa UNISIST (*UNESCO's World Scientific Information Programme*), define tesouro para a área de Ciência da Informação a partir de sua estrutura e de sua função. Em termos de sua função, tesouro é um dispositivo de controle terminológico usado para traduzir a linguagem natural dos documentos, dos indexadores ou dos usuários numa mais restrita linguagem do sistema, bem como para traduzir a linguagem do sistema de volta em linguagem natural. No que tange a sua estrutura, o tesouro é um vocabulário controlado e dinâmico de termos semântica e genericamente relacionados que cobrem um domínio específico do conhecimento.

O cenário sobre construção de tesouros começa a evoluir a partir do final dos

anos de 1970 com a Teoria do Conceito, apresentada por Ingetraut Dahlberg (DAHLBERG, 1978a), com a importância dos estudos recaindo sobre a unidade de representação em um tesouro. Dahlberg discute, então, qual seria a unidade de representação mais adequada para um tesouro, isto é, o que a autora define como conceito. Isto permitiu tratar o termo, verificando-se sua caracterização como representação de um conceito, e suas características (CAMPOS, 2001).

A captura e trabalho com palavras coletadas a partir de linguagem natural ocasiona problemas na constituição de um tesouro, uma vez que afeta o discurso em que está inserido. Observando esse fato, o tesouro conceitual parte do princípio do tratamento do conceito, ou seja, “o contexto do tesouro é o domínio e não do discurso” (CAMPOS; GOMES, 2006, p. 349). Neste sentido, cabe abstrair a formação do conceito a partir de suas classes gerais, permitindo observar a estrutura conceitual de um domínio (CAMPOS; GOMES, 2006). Isso faz com que o conteúdo do conceito seja explicitado e revele, ainda, seus relacionamentos.

Após análise do objeto/referente a que as palavras ou expressões se referem, tem-se o termo, com sua forma final estabelecida, bem como as relações entre ele e outros conceitos, revelados na análise. Por sua natureza verbal, a denominação do objeto/referente mascara, por vezes, sua verdadeira natureza, dificultando sua posição na estrutura do tesouro e é a análise conceitual que irá revelá-la. (CAMPOS; GOMES, 2006, p. 349).

O tesouro conceitual, como modelo de organização e representação do conhecimento, utiliza critérios das linguagens documentárias verbais aliadas a um sistema de classificação facetada. Com isso podemos dizer que estes são frutos de um método próprio para sua construção, o qual leva em conta a Teoria da Classificação Facetada e a Teoria do Conceito. Isto posto, entendemos um tesouro conceitual como um tipo de tesouro que está baseado no conceito, como unidade representacional, e na categorização, como norteadora da organização do conceito em um sistema de conceitos (CAMPOS; GOMES, 2006).

Considera-se, assim, conceito como uma “unidade de conhecimento”

(DAHLBERG, 1978a). É a partir dele que o tesouro conceitual tem sua constituição. A Teoria do Conceito, apresentada por Ingetraut Dahlberg em 1978, permite maior compreensão do conceito, fornecendo “bases seguras tanto para o estabelecimento de relações, como para a sua realização no plano verbal, ou seja, a determinação do que se denomina termo” (CAMPOS, 2001, p. 87). Em conjunto com a Teoria da Classificação Facetada, a Teoria do Conceito contribui “para a elaboração de tesouros conceituais porque estabelecem bases para identificação dos conceitos, dos termos e das relações entre eles, e, ainda, para sua ordenação sistemática” (CAMPOS; GOMES, 2006, p. 349).

Em síntese, os elementos do tesouro conceitual, que tem como bases teórico-metodológicas a Teoria do Conceito e a Teoria da Classificação Facetada, são:

- os conceitos, representados pelos termos;
- categorias e classes, que estabelecem a ordenação lógica e hierárquica dos conceitos;
- definições, que, bem constituídas, permitem posicionar um conceito em um sistema de conceitos;
- relações entre os conceitos, ou seja, sua mútua ligação.

3 Ontologia e ontologia de fundamentação

O termo ontologia tem origem no início século XVII, mais especificamente nos anos de 1606 e 1613, quando Jacob Lorhard escreve o *Ogdoas Scholastica* e Rudolf Göckel o *Lexicon philosophicum*. No entanto, ganha popularidade no âmbito da Filosofia através de Christian Wolff, com a publicação de *Philosophia prima sive Ontologia* (GUIZZARDI, 2005). Em Inteligência Artificial este termo é utilizado por empréstimo da Filosofia, sendo que estes estudos tiveram início nos anos 1990. A ideia era que ontologia fosse utilizada como possibilidade de representar a existência dos objetos e eventos, bem como suas relações. Neste sentido, ontologia era um artefato tecnológico que permitiria um modo de trabalhar com integração e recuperação da informação, gestão do conhecimento, sistemas multiagentes, etc. As

ontologias pretendem que a comunicação possa se dar entre humanos e máquinas, diferentemente dos tesauros. Outro aspecto também muito ressaltado pela Inteligência Artificial foi o uso de ontologias para representar conhecimento em dado domínio, como relatado por Guarino (1998).

Uma ontologia é, segundo Guarino (1998),

uma teoria lógica correspondente ao significado pretendido de um vocabulário formal, ou seja, seu compromisso ontológico com uma conceitualização particular de mundo. Os modelos pretendidos de uma linguagem lógica usada como vocabulário são restritos pelo seu compromisso ontológico. Uma ontologia indiretamente reflete esse compromisso (e a conceitualização subjacente) pela aproximação desses modelos pretendidos. (GUARINO, 1998).

Assim, baseados em Guarino (1998), podemos compreender uma ontologia como um artefato tecnológico que contém um conjunto de regras que delimitam o significado intensional de um vocabulário formal, permitindo que, a partir de um acordo ontológico, conhecimento possa ser compartilhado e inferido.

No que se refere à sua constituição, segundo Ramalho (2010), uma ontologia é composta por:

- Classes e subclasses: agrupam um conjunto de elementos, *coisas*, do *mundo real*, que são representadas e categorizadas;
- Propriedades: descrevem as características e/ou qualidades das classes;
- Relacionamentos: ligações entre classes, descrevendo e rotulando os tipos de relações existentes no domínio representado;
- Regras e axiomas: enunciados lógicos que possibilitam impor condições e possibilitam a realização de inferências automáticas a partir de informações que não necessariamente foram explicitadas no domínio;
- Instâncias: indicam os valores das classes e subclasses, constituindo uma representação de objetos ou indivíduos pertencentes ao domínio modelado, de acordo com as características das classes, relacionamentos e restrições definidas;



- Valores: atribuem valores concretos às propriedades, indicando os formatos e tipos de valores aceitos em cada classe.

A sua organização, como consta acima, permite que determinado domínio de conhecimento seja representado e que, por se constituir em um artefato tecnológico que realiza inferências com base em seu conteúdo, informações sejam recuperadas de forma mais eficaz.

Como os estudos sobre ontologia são realizados por várias áreas, ocorre a criação de instrumentos variados. É possível identificar, mesmo com a apresentação de propriedades distintas, uma tipologia bem definida de ontologias, já que existem características e componentes básicos que estão presentes em grande parte delas (ALMEIDA; BAX, 2009). Neste sentido, podemos, dentre outras tipologias, analisá-las quanto ao grau de formalismo e o aspecto estrutural. Enquanto a primeira diz que uma ontologia pode ser (a) altamente informal, (b) semi-informal, (c) semi-formal e (d) rigorosamente formal (USCHOLD; GRUNINGER, 1996), a segunda pode ser classificada em (a) ontologia de alto nível, (b) ontologia de domínio, (c) ontologia de tarefa e (d) ontologia de aplicação (GUARINO, 1998).

A partir da necessidade de compor a representação de dado domínio, nosso interesse incide sobre a ontologia de domínio, a qual busca descrever formalmente classes de conceitos e os relacionamentos de determinada área com o objetivo de compartilhar um consenso terminológico, e a ontologia formal, a qual representa, formaliza o conhecimento existente, permitindo que seja acessado e compartilhado através de conceitos e categorias que satisfaçam a compreensão de um domínio.

Deste modo, entendemos que uma ontologia formal está baseada em elementos como categorias e axiomas, demonstrando relações e propriedades dos conceitos e permitindo que a interpretação sobre dado conceito seja restrita, tomada através de um vocabulário controlado definido, delimitando o significado intensional de um vocabulário e permitindo maior controle sobre o domínio que está sendo mapeado, assegurando qualidade às inferências realizadas.

Baseada na ontologia formal, a ontologia de fundamentação, por sua vez, tem por objetivo identificar categorias gerais de certos aspectos da realidade que não



são específicos a um campo científico, descrevendo conhecimento independentemente de linguagem, de um estado particular das coisas ou ainda do estado de agentes (GUIZZARDI, 2005).

Borgo e Masolo (2009) definem ontologias de fundamentação a partir de quatro itens: (i) têm grande alcance; (ii) podem ser altamente reutilizáveis em cenários de modelagem diferentes; (iii) são filosóficas e conceitualmente bem fundamentadas; e (iv) são semanticamente transparentes e, portanto, ricamente axiomatizadas.

Com base nas definições apresentadas, podemos dizer que uma ontologia de fundamentação é uma ontologia formal que permite a elaboração de modelos de domínios, sendo altamente reutilizável. É também caracterizada por ser filosoficamente bem fundamentada, recebendo aportes da linguística, ontologia formal e ciências cognitivas, permitindo a explicitação de uma visão da realidade, ou seja, do acordo ontológico estabelecido, com determinação de regras de restrição, bem como conceitos, categorias e meta-propriedades.

Com efeito, nota-se a importância das ontologias de fundamentação no que tange a seus princípios filosóficos bem fundamentados para modelagem domínios. Esta importância é explicitada por Guizzardi, Falbo e Guizzardi (2008) ao relatarem que a ontologia de fundamentação permite a explicitação dos compromissos ontológicos da ontologia que representa um domínio, agregando fidelidade, consistência e clareza na representação.

Percebemos, apoiados no discurso de Guizzardi (2005), a principal contribuição das ontologias de fundamentação como sendo a busca pela caracterização fiel de uma conceitualização, indo ao encontro do seu compromisso ontológico, permitindo que haja uma representação do domínio em que a ontologia é empregada. Assim, fornecendo semântica baseada no mundo real e restringindo interpretações sobre seus conceitos com base em um vocabulário próprio. Isto nos permite interpretar as ontologias de fundamentação como fornecedoras de uma estrutura ontológica, as quais criam um *corpus* capaz de testar e validar o modelo conceitual gerado.

Uma vez que a ontologia de fundamentação é uma ontologia que pretende fornecer maior nível semântico do mundo para a modelagem conceitual de dado domínio de conhecimento, aborda questões como, por exemplo, (a) noções de tipos e suas instâncias; (b) objetos, e suas propriedades intrínsecas; (c) a relação entre identidade e classificação; (d) distinções entre tipos e suas relações; (e) relações parte-todo (GUIZZARDI, 2005), além de possuir elementos como classes, propriedades, relacionamentos, valores e regras na constituição do artefato.

5 Considerações finais

A partir do que foi exposto neste trabalho, pôde-se perceber que os elementos constituintes de tesouros conceituais e ontologias de fundamentação apresentam grande semelhança, como classes, conceitos, categorias, definições e relações, embora esta última em tesouros seja um modelo diádico, revelando as categorias as quais os conceitos pertencem, e em ontologias se apresente como um modelo triádico, buscando representar conhecimento de forma mais rica e, assim, manipulável por máquina a partir de uma proposição, explicitando a relação propriamente dita (SALES; CAMPOS; GOMES, 2008). Também cabe ressaltar que, por serem processáveis por máquina, as ontologias possibilitam a realização de inferências, utilizando-se de axiomas e regras de restrição.

A Ciência da Informação se utiliza de bases teórico-metodológicas para subsidiar a construção de sistemas de organização do conhecimento, fornecendo teorias e metodologias para que seja realizada a modelagem conceitual e, conseqüentemente, a criação de modelos representativos de domínios, como tesouros conceituais. A Ciência da Computação tem adotado, nos últimos anos, abordagens que permitem modelar parte de um domínio tendo como base teorias independentes de domínio, como as ontologias de fundamentação.

As bases teóricas utilizadas pela Ciência da Informação advêm da Teoria do Conceito, a qual fornece subsídios para pensar o conceito na forma de seu

conteúdo, e da Teoria da Classificação Facetada, onde se estabelece o uso de categorias para a organização de conceitos, permitindo seu posicionamento em um sistema de conceitos. A Ciência da Computação, com a utilização de ontologias de fundamentação, revela aportes da Filosofia, ciências cognitivas e ontologia formal, fornecendo uma proposta para representação de domínios.

Notamos, por fim, a possibilidade destes instrumentos – tesouros conceituais e ontologias de fundamentação – servirem como base para organizar, representar e recuperar informação de forma eficaz, uma vez que seus modelos conceituais estão baseados em teorias independentes de domínio, permitindo que representações de domínios sejam realizadas sob do enfoque do estabelecimento de conceitos e suas relações, criando um sistema conceitual.

6 Referências

ALMEIDA, M. B.; BAX, M. P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da Informação**, Brasília, v. 32, n. 3, p. 7-20, set./dez. 2003. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/download/17/12>>. Acesso em: 25 jul. 2009.

BORGIO, S. MASOLO, C. Foundational choices in DOLCE. In: STAAB, S.; STUDER, R. (Ed.). **Handbook on Ontologies**. 2. ed. Springer Verlag, 2009. p. 361-382. Disponível em: <<http://www.loa-cnr.it/Papers/IOShandDOLCEv16.pdf>>. Acesso em: 23 nov. 2009.

CAMPOS, M. L. A. **Linguagem documentária**: teorias que fundamentam sua elaboração. Niterói: EdUFF, 2001. 133 p.

CAMPOS, M. L. A. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. **Ciência da Informação**, Brasília, v. 33, n. 1, p. 22-32, jan./abr. 2004. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/download/56/53>>. Acesso em: 12 ago. 2009.

CAMPOS, M. L.; GOMES, H. E. Metodologia de elaboração de tesouro conceitual: a categorização como princípio norteador. **Perspectivas em Ciência da Informação**, v. 11, n. 3, p. 348-359, set./dez. 2006.



DAHLBERG, I. A referent-oriented analytical concept theory of interconcept.
International Classification, Frankfurt, v. 5, n. 3, p. 142-150, 1978a.

DAHLBERG, I. **Ontical structures and universal classification**. Bangalore: Sarada Ranganathan Endowment, 1978b. 64 p.

GUIZZARDI, G. **Ontological foundations for structural conceptual models**. 416 f. Tese (PhD em Computer Science) – Twente University of Technology, Twente, Holanda, 2005.

GUIZZARDI, G.; FALBO, R. A.; GUIZZARDI, R. S. S. A importância de Ontologias de Fundamentação para a Engenharia de Ontologias de Domínio: o caso do domínio de Processos de Software. **IEEE LATIN AMERICA TRANSACTIONS**, vol. 6, n. 3, jul. 2008.

GUARINO, N. Formal Ontology in Information Systems. In: GUARINO, N. (Ed.). **Formal Ontology in Information Systems**. Amsterdam: IOS Press, 1998. p. 3-15. Disponível em: <<http://www.loa-cnr.it/Papers/FOIS98.pdf>>. Acesso em: 12 jul. 2009.

RAMALHO, R. A. S. **Desenvolvimento e utilização de ontologias em bibliotecas digitais**: uma proposta de aplicação. 145 f. Tese (Doutorado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP, Marília, 2010.

SALES, L. F.; CAMPOS, M. L. A.; GOMES, H. E. Ontologias de domínio: um estudo das relações conceituais. **Perspectivas em Ciência da Informação**, v. 13, n. 2, p. 62-76, maio/ago. 2008. Disponível em: <<http://www.eci.ufmg.br/pcionline/index.php/pci/article/download/219/474>>. Acesso em: 12 nov. 2009.

UNESCO. **Guidelines for the establishment and development of monolingual thesauri**. Paris: UNESCO, 1973. 37 p.

USCHOLD, M.; GRUNINGER, M. Ontologies: principles, methods and applications. **The Knowledge Engineering Review**, v. 11, n. 12, p. 93-136, 1996.

