

Elaboração de modelos de domínio em ontologias: a abordagem onomasiológica e a função da definição

Maria Luiza de Almeida Campos

Pós-Doutora pela Universidade de São Paulo (USP), São Paulo, Brasil.

Pós-Doutora pela Fundação Instituto Oswaldo Cruz (Fiocruz), Rio de Janeiro, Brasil.

Doutora em Ciência da Informação, Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT - UFRJ), Rio de Janeiro, Brasil.

Professora da Universidade Federal Fluminense. Departamento de Ciência da Informação e do Programa de

Pós-Graduação em Ciência da Informação (UFF – PPGCI), Niterói, Brasil.

E-mail: maria.almeida@pq.cnpq.br

Submetido em: 10/07/2017. Aprovado em: 05/09/2017. Publicado em: 28/12/2017.

RESUMO

A captura do conhecimento para a elaboração de modelos de domínios pode adotar uma abordagem semasiológica de perspectiva linguística (partindo do termo para o conceito) ou a abordagem onomasiológica, de perspectiva extralinguística (partindo do conceito para o termo). No que tange às ontologias, a adoção da abordagem semasiológica, na maioria das vezes, no que concerne à atividade de captura de conhecimento, reflete um procedimento de análise de textos eletrônicos, partindo da linguagem natural e do texto escrito em seus diversos formatos. Já a abordagem onomasiológica considera primeiramente que se chega ao termo através do extralinguístico, ou seja, a partir do entendimento do referente no âmbito de um domínio de especialidade; o termo seria o término de uma atividade de conotação sobre o referente. Este artigo aborda a metodologia apoiada na abordagem onomasiológica elaborada por I. Dahlberg, e demonstram-se os papéis que as definições possuem nesta atividade. A definição, neste sentido, é entendida como produto de um acordo estabelecido entre o modelizador/ontologista e o especialista do domínio; ela reflete o modelo pretendido de representação do mundo modelado.

Palavras-chaves: Modelo de domínio. Ontologia. Abordagem onomasiológica. Abordagem semasiológica. Captura do conhecimento.

Domain models development in ontologies: the onomasiological approach and the function of the definition

ABSTRACT

The knowledge capture to elaborate domain modeling can have a semasiological approach from a linguistic perspective (from the term to the concept) or the onomasiological approach of an extralinguistic perspective (from the concept to the term). Regarding the ontologies, the use of the semasiological approach in the majority of the times, what involves the activity of knowledge capture reflects a procedure of electronic texts analysis starting with natural language and written text in its various formats. The onomasiological approach, on the other hand, considers at first that the term is reached through the extralinguistic, that is, from the understanding of the referent within a specialty domain; the term would be the end of an activity of connotation about the referent. In this paper we expose the methodology supported by the onomasiological approach elaborated by I. Dahlberg, and it demonstrates the roles that the definitions have in this activity. The definition in this case is understood as the product of an agreement established between the modeler/ontologist and the domain expert; It reflects the intended model of representation of the modeled world.

Keywords: Domain model. Ontology. Onomasiological approach. Semasiological approach. knowledge capture.

Elaboración de modelos de dominio en ontologías: el enfoque onomasiológico y la función de la definición

RESUMEN

La captura del conocimiento para la elaboración de modelos de dominios puede adoptar el enfoque semasiológico de perspectiva lingüística (partiendo del término para el concepto) o el enfoque onomasiológico, de perspectiva extralingüística (partiendo del concepto para el término). En lo que atañe a las ontologías, la adopción del enfoque semasiológico, la mayoría de las veces, en lo que refiere a la actividad de captura de conocimiento, refleja un procedimiento de análisis de textos electrónicos, partiendo del lenguaje natural y del texto escrito en sus diversos formatos. Ya el enfoque onomasiológico considera primero que se llega al término a través del extralingüístico, o sea, a partir del entendimiento del referente en el ámbito de un dominio de especialidad; el término sería la terminación de una actividad de connotación sobre lo referente. En este artículo, abordamos la metodología apoyada en el enfoque onomasiológico elaborado por I. Dahlberg y demostramos los papeles que las definiciones poseen en esta actividad. Se entiende la definición, en este sentido, como producto de un acuerdo establecido entre el ontólogo y el especialista del dominio; ella refleja el modelo pretendido de representación del mundo modelado.

Palabras clave: Modelo de dominio. Ontología. Enfoque onomasiológico. Enfoque semasiológico. Captura del conocimiento.

INTRODUÇÃO

Ao longo dos últimos anos, as ontologias informacionais vêm sendo consideradas como sistemas formais para representar conhecimento em determinado domínio, atuando assim como um modelo formal para sistematizar a expressão do conhecimento que manipulam. Desde a década de 1990, foram conceituadas como uma especificação explícita de uma conceituação (GRUBER, 1993a). As ontologias são, portanto, consideradas como um corpo de conhecimento sistematizado em dado domínio, no qual entidades e relações que se manifestam são representadas através de dada formalização.

Sua finalidade, como muito bem apresentada por Bachimont (2006), atende a todos os que necessitam sistematizar a expressão do conhecimento que manipulam. Formalmente, uma ontologia é uma declaração de uma teoria lógica definindo um conjunto de primitivos de representação com os quais modelamos o domínio do conhecimento; tais primitivos são classes, atributos e relações. As definições explicitam este entendimento, enquanto os axiomas formais restringem a interpretação. A ontologia fornece uma linguagem de comunicação com o agente (SANTOS, 2010).

Os modelos de domínios resultantes dessas operações de conceituação em ontologias informacionais podem ser definidos como modelos que representam o conhecimento de uma área de aplicação (área de assunto, de missão ou de problema) e que devem servir como uma representação de consenso (ou modelo de referência) de uma conceituação compartilhada por determinada comunidade.

No âmbito da inteligência artificial, na década de 1990, Davis; Shrobe e Szolovits, (1993) iniciam uma reflexão sobre a variedade de representações possíveis que podem ser produzidas a partir de uma interpretação sobre o mundo, acentuando a importância de uma visão compartilhada sobre o conhecimento que se quer representar. Para Davis, apesar de a representação do conhecimento ser um dos conceitos mais centrais e familiares em inteligência artificial, a questão fundamental consiste em defini-lo, reforçando assim a posição de Newell (1982) sobre a imperfeição na representação do conhecimento (CAMPOS, 2010).

Nesse sentido, consideramos que determinar uma representação significa optar por um conjunto de compromissos ontológicos que sejam pragmaticamente possíveis de serem representados e, no caso de ontologias informacionais, possíveis

de serem representados computacionalmente. Logo, a representação do modelo de domínio em uma ontologia informacional requer uma escolha compartilhada de conceituações como uma forma de pensar sobre um mundo fechado e computacionalmente representável.

Nessa perspectiva, este artigo propõe discutir teoricamente um dos primeiros procedimentos para a elaboração de ontologias de domínio – o levantamento de *corpus* terminológico consistente que descreva este domínio. De modo geral, as propostas metodológicas pressupõem a existência ou mesmo a captura de uma estrutura de entidades **já existentes** sugerindo, a partir dessa estrutura, critérios para verificar a explicitação de propriedades e relações entre elas (FERNANDEZ; GOMEZ-PERZ; JURISTO, 1997; DING & FOO, 2002; FERNANDEZ-LÓPEZ, 2001; GANGEMI; PISANELLI; STEVE, 1999; GRUNINGER; FOX, 1995; GUARINO; WELTY, 2002; METAXIOTIS; PSARRAS; ASKNOUNIS, 2001; MIZOGUCHI, 2014; PASLARU-BONTAS, 2007; SURE; STAAB; STUDER, 2002; USCHOLD; KING, 1995). Entretanto, temos observado a necessidade de estudos que venham a discutir procedimentos metodológicos para a formação de *corpus* terminológico em ontologia de domínio. Esse *corpus*, consistentemente constituído, agregando as relações e axiomas, permite que se criem condições para a implementação de raciocínio automático (CAMPOS; ROCHA, 2017).

Em domínios de especialidade, as definições 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 que tais conteúdos presentes nas definições complementam a descoberta deste *corpus* terminológico.

A determinação de definições como um ponto de partida para a sistematização do *corpus* terminológico em ontologia de domínios se sustenta, neste estudo, por meio de uma abordagem onomasiológica. Esta abordagem é uma das orientações metodológicas para a elaboração de terminologias e tem sido recuperada

pela investigação na área da inteligência artificial e pelo desenvolvimento das ontologias. No entanto, as questões que envolvem a elaboração e a sistematização/classificação de domínios de conhecimento pertencem também à ciência da informação, através das teorias e metodologias visando à organização do conhecimento (VICKERY, 1997).

No campo da ciência da informação e dos estudos em linguagens documentárias, temos Dahlberg como uma das representantes da abordagem onomasiológica no campo da atividade tesaurográfica (DAHLBERG, 1980). Em contraposição à abordagem semasiológica, método também utilizado para a elaboração de terminologias de perspectiva linguística (partindo do termo para o conceito), a abordagem onomasiológica possui uma perspectiva extralinguística, partindo do conceito para o termo. A abordagem onomasiológica se caracteriza 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, nos quais esta “versão do referente” está consolidada nas definições (DAHLBERG, 1978a, 1978b, 1983; RIGGS, 1979, 1989a, b, 1996). Assim, o foco da análise se concentra no referente, no objeto (podendo ser abstrato ou concreto) e é a partir dele que as definições são elaboradas. Na verdade, a definição passa a ser um enunciado sobre o entendimento daquele referente em um dado domínio.

Consideramos assim discutir sob este enfoque as questões que envolvem a elaboração de modelos de domínios evidenciando o *modus operandi* para o levantamento do *corpus* terminológico.

Nesse sentido, apresentamos os papéis que as definições possuem neste processo: elas são ao mesmo tempo elementos que permitem a captura do conhecimento sobre o domínio, mas também são produtos de um acordo estabelecido entre o modelizador/ontologista e o especialista do domínio para a elaboração do modelo pretendido de representação do mundo modelado.

Este artigo está assim organizado: além desta introdução, apresentamos os tipos de definição em ontologias, considerando aqui as definições operativas como as mais apropriadas, ressaltando que, neste contexto, as definições são o resultado de uma prática de autoria, do especialista e do ontologista. Após, vamos discutir os procedimentos de captura do conhecimento sobre o domínio no âmbito da elaboração de modelos de domínios em ontologias. Investigamos os procedimentos semasiológicos e onomasiológicos, para finalizar com a apresentação do método onomasiológico de I. Dahlberg em sua teoria do conceito voltada para o referente. Logo depois, as considerações finais deste trabalho.

TIPOS DE DEFINIÇÃO EM ONTOLOGIAS

Podemos considerar que a função de uma definição, de forma geral, é principalmente ser capaz de explicitar, através do *definiens* (conjunto de enunciados sobre o que se quer definir) o seu respectivo *definiendum* (símbolo que se quer definir). Assim, a sua elaboração pressupõe, em ontologias, determinar condições necessárias e suficientes que permitam distinguir os objetos entre si em um dado mundo possível.

Nesse sentido, segundo Marconi (1997), o significado de um termo (o valor semântico de um sinal) surge quando falantes competentes – por exemplo, os especialistas em um determinado domínio – usam consistentemente uma série de letras ou símbolos para se comunicar sobre algo no mundo. Isso é o que é o referente do termo – por exemplo, objetos, processos ou atributos relevantes para a atividade em que os especialistas estão envolvidos. O significado é, então, algo como uma representação mental (ou conceito) que os especialistas têm em mente, quando usam um termo para se referir a um tipo de coisa no mundo onde está delimitado seu domínio de atuação. Nesse caso, a representação mental expressa por um termo é composta de certas características que são (pelo menos em parte) também apresentadas em uma definição. Logo, o termo, a representação mental e o enunciado definitório têm

todos, o mesmo referente. E é sobre ele, o referente, que as definições são elaboradas. Assim, uma definição é um enunciado encontrado como um pequeno texto no campo de definição de uma entrada de dicionário especializado e nos campos de definição ou comentário de uma ontologia (SEPPÄLÄ, 2015).

Considerando que o enunciado definitório, fruto da compreensão sobre o referente e da representação mental produzida pelo especialista, deve evidenciar o entendimento acordado em um dado domínio, a explicitação desse enunciado se dá de duas formas em ontologias: como definições textuais e como definições lógicas. Tal classificação é apresentada por Seppälä (2014), que afirma que as ontologias têm, por um lado, axiomas que fazem parte de uma definição lógica de termos e, por outro, definições em linguagem natural, as definições textuais. Em uma ontologia, uma definição textual é, idealmente, uma frase que designa o objeto e suas propriedades a partir de determinado comprometimento ontológico e que tem por função especificar o significado pretendido dos termos da ontologia a fim de evitar ambiguidades e erros de interpretação pelo homem. Esse tipo de definição da linguagem natural também é encontrado em dicionários terminológicos especializados. Dessa maneira, uma definição textual normalmente indica o tipo de classe a que as instâncias (o referente) do termo definido pertencem, e distingue essas instâncias de outras que se enquadram no mesmo tipo, listando uma ou mais das características dessas instâncias, sendo que em ontologia o primeiro elemento também deve ser o gênero a que a instância pertence. As definições lógicas possuem axiomas que restringem o significado pretendido de um termo afirmando as condições necessárias para a sua utilização. Eles assim funcionam de forma análoga às condições necessárias discutidas nas definições textuais e sua função de desambiguação é também análoga. Além disso, os axiomas possibilitam a formação das estruturas taxonômicas em ontologias e permitem que a máquina realize funções que permitirão estabelecer inferências sobre o domínio de forma automática.

As definições são o resultado de uma prática de autoria, do especialista e do modelizador/ontologista. Tais definições, lógicas ou textuais, necessitam, além de conter elementos que identificam o gênero próximo e a diferença específica, de outros que também possam identificar componentes, partes, função, finalidade, causa, etc. As definições que são dessa forma elaboradas foram denominadas por Hegenberg (1974) definições operacionais – aquelas em que os elementos da definição não identificam somente o gênero próximo e a diferença específica –, mas também elementos que procuram definir certas operações em que o conceito seria aplicado. Esta forma de definir faz parte de um movimento denominado operacionismo, que na década de 1960 do século passado vem se contrapor a procedimentos metodológicos de caráter mais conceitual e filosófico. Dentro dessa perspectiva, o operacionismo defende uma posição de que só é possível conhecer significados de um conceito a menos que possamos especificar as operações que foram realizadas, por nós ou por outrem, ao aplicar o conceito a qualquer situação concreta. Portanto, as definições operativas revelam a maneira de aplicação dos conceitos dentro de um dado contexto (CAMPOS, 2010).

No campo do trabalho terminológico, especificamente relacionado à elaboração de definições, diversas questões são apresentadas relacionadas à forma de construção de definições, ou seja, se estas estão de acordo com uma perspectiva linguística ou terminológica; se é possível definir algo a partir da existência ou não do referente no mundo (definições reais ou léxicas); se há diferenças entre os conceitos de descrição e definição; se a definição está relacionada a uma função comunicativa ou referencial da linguagem, entre outras tantas questões de fundamental importância relacionada ao ato de definir. No que tange ao objetivo do estudo que ora estamos desenvolvendo em ontologias, vamos nos concentrar na primeira questão que pretende colocar em discussão as perspectivas linguísticas e terminológicas para a captura do conhecimento sobre o domínio, e que

consequentemente está relacionada à dada forma semasiológica ou ontológica, respectivamente, de elaboração de definições.

ONTOLOGIAS E OS PROCEDIMENTOS DE CAPTURA DO DOMÍNIO: A ELABORAÇÃO DE MODELOS DE DOMÍNIOS

Em 1986, B.C. Vickery, em seu artigo “Knowledge Representation: a brief review” (VICKERY, 1986), já nos alertava sobre a importância da representação do conhecimento para sistemas computacionais e para as áreas de conhecimento para as quais esta questão era cara, como a ciência da informação, a inteligência artificial, a psicologia e a linguística. Nesse artigo, o enfoque não estava voltado para relacionar a representação do conhecimento às possibilidades tecnológicas da época: apresentava temáticas que se referiam à estrutura semântica de frases e enunciados, categorias e relações na análise de domínios e a representação do conhecimento por argumentos e axiomas; temáticas atualmente caras para aqueles que se dedicam ao estudo da representação e recuperação no ambiente da Web Semântica e de sistemas inteligentes. E, nesse mesmo trabalho, ele nos alerta que, já naquela época, a representação do conhecimento era relevante, além dos documentos e índices, e deveria ser discutida no âmbito da estrutura de registros e arquivos em bases de dados; da estrutura de dados nos programas de computador; da estrutura sintática e semântica da linguagem natural; da representação do conhecimento em inteligência artificial; dos modelos de memória humana – em todos esses campos nos quais é necessário decidir como o conhecimento pode ser representado de forma que as representações possam ser manipuladas.

Como se observa, a questão da representação do conhecimento vem preocupando o mundo da ciência da informação desde sua gênese. Atualmente, com o advento da Web Semântica, outras temáticas, na maioria das vezes não tão novas, vêm se incorporar a este mundo.

Modelos de domínios são obtidos a partir de processos que têm por objetivo uma representação do conhecimento de alta qualidade de fenômenos selecionados em algum domínio. Os modelos resultantes devem facilitar, dependendo do objetivo para o qual o modelo foi elaborado, a representação e a recuperação de informações/conhecimento; nesse caso, em sistemas de recuperação de informações. Para que um modelo possa ser uma descrição adequada da realidade do domínio, ele deve apresentar informações precisas e claras, não permitindo a ocorrência de ambiguidades sobre os aspectos que devem ser modelados.

Na elaboração desses modelos, procedimentos metodológicos são utilizados para a modelagem do domínio, ou seja, aqueles relacionados à modelagem conceitual e os relacionados à modelagem ontológica.

Modelagem conceitual consiste na identificação, análise e descrição dos elementos e restrições de um domínio do mundo real, normalmente utilizando uma linguagem de modelagem, a fim de que tais elementos sejam incorporados em um sistema de informação. Já modelagem ontológica pode ser definida como a criação de uma ontologia específica de domínio, através da captura das entidades relevantes do domínio e incorporação das mesmas em um conjunto de categorias que revelam sua natureza, por meio de uma linguagem de especificação de ontologia. Como pode-se perceber, existe uma forte semelhança entre esses dois tipos de modelagem, uma vez que ambas objetivam capturar e modelar elementos do mundo real. Porém, enquanto na modelagem conceitual busca-se estabelecer as relações entre os conceitos abstraídos de um domínio, na modelagem ontológica o objetivo é identificar os objetos e entender sua natureza por meio da descrição de suas propriedades. [Segundo Poli (2001), a primeira possui um caráter epistemológico, em oposição ao caráter ontológico da segunda]. Dessa forma, pode-se deduzir que as duas modelagens são possivelmente complementares, ou seja, a modelagem ontológica pode constituir-se numa base para modelagem conceitual, no sentido de prover ao projetista, de forma clara e sem ambiguidades, o conhecimento necessário sobre o domínio a ser modelado. (VILLELA, 2004, p. 243)

Consideramos, como Villela (2004), que ambas as abordagens se completam. Entretanto, nossa experiência tem evidenciado que, no âmbito dos modelos de domínio em

ontologias, a modelagem ontológica, que parte do extralinguístico, nos parece a mais apropriada, como defenderemos adiante.

Como Gruber (1993b) afirma as ontologias não pretendem representar todo o conhecimento, elas refletem determinada visão do mundo, em consenso dentro de uma comunidade. Ontologias informacionais podem ser consideradas, assim, como o estudo das características básicas de uma dada realidade. Visando explicitar a forma pela qual essas características básicas podem ser apropriadas pelo ontologista, é que apresentamos dois procedimentos que vêm sendo utilizados para a captura do conhecimento sobre um dado domínio, a perspectiva linguística e a extralinguística.

Alguns autores apresentam (CABRÉ, 2003; GOMES; CAMPOS; GUIMARÃES, 2010; LARA, 2010; SANTOS, 2010; HERMANS, 1989; ROCHE, 2006; TÁLAMO; LARA, 2009) a perspectiva linguística como um procedimento que parte do discurso verbal na forma de texto enquanto a perspectiva extralinguística parte do especialista enquanto detentor do conhecimento de um campo de especialidade. Temos, assim, um procedimento que parte do texto de especialidade e outro que parte do discurso do especialista para a elaboração de conceituações sobre o domínio e, no caso das ontologias de domínio, o modelo que o representa. Denomina-se a perspectiva linguística como aquela que se apoia em uma abordagem semasiológica, e a perspectiva extralinguística como aquela que se apoia em uma abordagem onomasiológica.

A abordagem semasiológica considera o termo como uma unidade lexical empregada no âmbito de um domínio de especialidade. A terminologia denominada moderna (CABRÉ, 2003) estuda os termos em contexto, ou seja, no interior do texto, sendo que uma palavra pode adquirir o estatuto de termo em função do seu papel no texto. Assim, a abordagem semasiológica parte do texto e da possibilidade de extração desses termos a partir de textos, assim como os seguidores dessa abordagem consideram o levantamento do corpus terminológico

para a elaboração de modelos como atividade que pressupõe a análise textual para a determinação desse corpus. No que tange às ontologias, a adoção dessa abordagem, na maioria das vezes, no que concerne à captura do conhecimento, reflete um procedimento de análise de textos eletrônicos, partindo da língua natural e do texto escrito em seus diversos formatos (SANTOS, 2010).

Por outro lado, a abordagem onomasiológica considera primeiramente que se chega ao termo através do extralinguístico, ou seja, a partir do entendimento do referente no âmbito de um domínio de especialidade, o termo seria o término de uma atividade de conotação sobre o referente. O termo denota o conceito. Nessa perspectiva, o conceito enquanto representação mental precede a forma linguística ou qualquer outro símbolo, sendo que as relações entre conceitos são estabelecidas em função de parâmetros extralinguísticos, a partir da identificação das propriedades do referente. No campo das ontologias informacionais, tal abordagem tem sido empregada na etapa de levantamento do conhecimento sobre o domínio, através de procedimentos de entrevistas com especialistas do domínio. Esse procedimento na elaboração de ontologias tem sido denominado captura do conhecimento sobre o domínio (CEUSTERS; SMITH, 2010; FERNANDES; GUIZZARDI; GUIZZARDI, 2011; GUIZZARD; MORALES-RAMIREZ; PERINI, 2014; TAIRAS; MERNIK; GRAY, 2008).

A construção das bases de conhecimento nas ontologias guarda alguma semelhança com sistema de recuperação de informação, entretanto ela tem suas especificidades. Assim, nesse sistema

o percurso da enunciação envolve posições relativas tanto ao enunciador (emissor) quanto ao enunciatário (receptor) que correspondem aos percursos onomasiológicos e semasiológicos, respectivamente. Nesse sentido, do conceito ao dito, tem-se a onomasiologia, e do dito à sua interpretação, a semasiologia. Logo, a orientação onomasiológica é da categoria do enunciador e a semasiológica é da categoria da recepção. (POTTIER, 1992 citado por LARA, 2010)

No caso das ontologias, em que a questão que se coloca é a descrição de forma consistente sobre um domínio que se deseja representar, a questão da recepção não se coloca, pois a comunicação se dá entre pares.

Para a elaboração de modelos de domínio em ontologias, consideramos defender a abordagem onomasiológica. Entretanto, vamos apresentar uma forma de procedimento de se chegar ao extralinguístico que não parte inicialmente do especialista, mas do estabelecimento de indagações sobre o referente no interior de um domínio. Tais indagações colocam o modelizador em uma posição de investigador sobre o domínio e não como um receptor do conhecimento do especialista ou mesmo de suas produções textuais, como tentaremos demonstrar.

Actualmente, o papel da terminologia na construção de ontologias não passa exclusivamente pelo texto. E essa tendência encontra-se, curiosamente, nos princípios metodológicos estabelecidos por Eugen Wüster [...]. Do nosso ponto de vista, embora uma perspectiva semasiológica não seja adequada para, sozinha, formalizar conhecimento num sistema computacional, pode contribuir com informação sobre os conceitos e as suas denominações, abarcando, por isso, o linguístico e o extralinguístico, uma vez que o significado dos termos contém também uma significação ligada ao acto comunicativo e cultural da sua produção, ligada, portanto, a uma parte conceptual e extralinguística. A terminologia que parte de uma abordagem linguística privilegia a verbalização do conhecimento. A terminologia que parte de uma abordagem extralinguística centra-se na conceptualização do conhecimento. (SANTOS, 2010, p. 75 e 89)

Para a realização de uma conceituação, vamos nos reportar a Dalhberg por meio de sua teoria do conceito voltada para o referente para apresentarmos o procedimento de como capturar o conhecimento sobre o domínio a partir de uma abordagem extralinguística.

A TEORIA DO CONCEITO E A CAPTURA DO CONHECIMENTO SOBRE O DOMÍNIO: A ABORDAGEM ONOMASIOLÓGICA

Dahlberg (1981) afirma que a compreensão sobre algo em dado domínio de especialidade pressupõe o conhecimento de três elementos: a) o seu referente, que são os objetos concretos e abstratos e que se encontram em um espaço extralinguístico; b) as características e propriedades desses referentes, que são proposições adequadas enunciadas no âmbito de um domínio sobre este referente; e c) a expressão verbal utilizada e que promove a comunicação entre um certo grupo de falantes. Isso se configura em seu triângulo conceitual. E esclarece: a fim de acumular conhecimento sobre o referente, deve-se coletar enunciados sobre o referente A, de modo a identificar suas características, em que dada característica nos revela a categoria à qual o referente pertence. As informações sobre B, suas características, ajudam a identificar o gênero próximo e seus demais relacionamentos como outros conceitos em dado domínio. O conhecimento de C determina a expressão verbal apropriada a ser usada.

A Teoria do Conceito (DAHLBERG, 1978a) fornece um método para analisarmos um item de referência em dado domínio, propiciando o que denominamos aqui captura de conhecimento sobre o domínio. Esse método se estabelece no âmbito de uma perspectiva onomasiológica, pois parte da análise de um elemento extralinguístico (o referente), a partir do qual são estabelecidas perguntas sobre este elemento, a fim de permitir certa sistematização desse referente em dado domínio. Nesse sentido, a predicação que ocorre sobre o referente não é feita de maneira aleatória; sua teoria nos indica um *modus operandi* de como devem ser realizadas as asserções sobre o referente em dado domínio, como veremos adiante. Aqui ainda é necessário ressaltar o aspecto sistemático de sua teoria, ou seja, possibilita a elaboração de definições que se interconectam formando um todo coeso que, neste caso, representa o conhecimento no âmbito de um domínio de especialidade. Por isso é que em seus trabalhos ela afirma que os conceitos se apresentam uns em relação aos outros, formando um sistema de conceitos.

Seu *modus operandi* parte, assim, da identificação das propriedades do referente. Logo, através da formulação dos enunciados sobre as propriedades dos referentes se obtêm as características dos respectivos conceitos. Para Dahlberg, o ato de definir é uma construção de aspectos selecionados do referente realizado pelo “olhar” de um grupo que aquiesce sobre o seu entendimento em um dado espaço de conhecimento. Portanto, esses elementos que formam o conceito – as características – são propriedades dos referentes (objetos concretos ou abstratos) que, no nível do conceito, passam a ser também características do conceito. As características são, pois, representações que se estabelecem a partir de um “acordo” sobre as propriedades dos referentes. Por isso, ‘conceito’ em Dahlberg representa uma unidade de conhecimento sobre o acordo firmado no interior de um domínio de conhecimento.

As características também são conceitos e são usadas para classificar e definir outros conceitos. Por exemplo, uma das características do conceito ‘Verniz à base de resina’ é ‘resina’ e ‘resina’ também é um conceito. Se não soubermos o que é ‘resina’, não poderemos entender corretamente o que seja um ‘verniz à base de resina’. Nesse sentido, o conceito nunca é estabelecido de forma isolada, mas sempre em relação com outros. Ele se estabelece através da comparação com outros conceitos, quando se identificam semelhanças e diferenças (por meio da análise e comparação das características), o que leva à reunião de conceitos, relacionados de maneira variada. Por exemplo, existem semelhanças entre os conceitos ‘fungicida’, ‘herbicida’ e ‘inseticida’. Eles têm em comum a característica ‘que mata’. O que os diferencia entre si é o objeto que sofre a ação da substância nomeada: o fungo, a erva daninha, o inseto, respectivamente. Se acrescentarmos o conceito ‘biocida’, veremos que ele também tem a característica ‘que mata’ além da característica ‘organismo vivo’ que, num grau de generalização maior, engloba os outros três, estabelecendo uma posição de superordenação (GOMES; CAMPOS, 2002).

A identificação das características dos conceitos é de fundamental importância para o entendimento do conceito em dado domínio, e assim consequentemente possibilita sua classificação ou, dito de outro modo, sua ordenação no interior

deste domínio. As características são elementos da definição e revelam as conexões existentes entre o conceito em questão e os demais através de suas relações. Dahlberg (1978b) classifica as características em essenciais e acidentais. As características essenciais são consideradas como os elementos necessários e suficientes, em um dado espaço de conhecimento, para determinar de que tipo um conceito é, ou seja, determinar sua identidade. Já as características acidentais são opcionais para um determinado tipo de conceito, ou seja, dito de outra maneira, não são necessárias ou constantes. Assim, por exemplo, para ser um coração, algumas características essenciais são: ser órgão, ser feito de músculo e circular sangue; e algumas características acidentais são: ser grande ou apresentar sopro. A característica essencial permite a ordenação do conceito no domínio, tendo por função estruturar modelos de domínios, como as taxonomias elaboradas para a estruturação do conhecimento de um domínio em ontologias. Por isso ela afirma que as características essenciais *definem*; as características acidentais *descrevem* (DAHLBERG, 1978b). A definição para ela é ordenadora, na medida em que possibilita a classificação do objeto no domínio. Por isso o primeiro elemento da definição deve revelar o gênero próximo.

A definição revela, pois, a ordenação dos conceitos a partir de um compromisso ontológico (acordo firmado) estabelecido em um dado domínio; ela é a consequência dessa ordenação (classificação).

Como dito anteriormente, nessa perspectiva, os conceitos se definem uns em relação aos outros. Nesse processo, ficam evidenciadas certas relações que fornecem as bases para sua ordenação. Sempre que houver entre dois conceitos características iguais, opostas, semelhantes ou outras, pode-se considerar que existem relações entre eles. As relações são classificadas em quantitativas e qualitativas (DAHBERG, 2014).

A espécie quantitativa compara dois conceitos do ponto de vista estritamente formal; aqui estão incluídas as seguintes relações: de identidade, de inclusão, de interseção e de disjunção. A espécie qualitativa

considera os aspectos formal e material, podendo ser classificadas como relação formal/categorial; relação material/paradigmática; relação sintagmática funcional. Essa rede de relações constitui o sistema de conceitos. A relação formal – categorial depende completamente da espécie de item de referência que se está analisando, no caso, fenômenos, objetos gerais, objetos materiais, quantidades, qualidades, comparações, operações, estados, processos, períodos, posições, lugares no espaço, por exemplo. Esta relação reúne os conceitos dentro de uma mesma categoria. Isso ocorre porque tais conceitos são todos de mesma natureza, ou seja, a partir da análise do conceito, as características essenciais levam à mesma categoria. A relação material-paradigmática permite organizar conceitos de mesma natureza, ou seja, aqueles que estão no interior de uma categoria. Pode ser hierárquica, partitiva e de oposição. A relação hierárquica baseia-se na relação lógica de implicação e é de dois tipos: relação de abstração (relação gênero-espécie) e relação lateral (relação dos conceitos no renque). A relação partitiva existe entre o todo e suas partes, e as partes também podem estar relacionadas entre si. A relação de oposição pode ser de três espécies: relação de oposição contraditória, relação de oposição contrária, relação positiva-indiferente-negativa. A relação funcional-sintagmática se dá entre conceitos de diferentes categorias. A partir de um conceito que denote um processo ou operação, leva a conceitos que suplementam essas ações, como na seguinte sequência: produção – produto – produtor – comprador, etc. (CAMPOS, p. 104, 2001).

Nos procedimentos de elaboração de modelos de domínios, as relações categoriais são fundamentais, pois permitem uma primeira estruturação do conceito no interior de um domínio.

Podemos ver que as categorias têm uma capacidade de estrutura: não apenas estruturam, de fato, todos os nossos elementos de conhecimento e unidades do conhecimento; elas fornecem, ao mesmo tempo, por este meio, o esqueleto, os ossos e tendões para estruturar todo o nosso conhecimento. Com seu uso consciencioso, então, o corpo do nosso conhecimento pode se manter unido, pode se mover, pode se manter flexível – e pode crescer organicamente. (DAHLBERG, 1978a, p. 34)

As categorias apresentadas por Dahlberg são metaconceitos que possibilitam que o modelizador chegue a qualquer domínio e consiga estabelecer uma primeira ordenação. Elas possibilitam a identificação dos referentes no interior de um domínio. No nível de uma representação, são classes de grande abrangência em um domínio, acima das quais não se pode colocar qualquer conceito. Em seu trabalho recente, Dahlberg (2014) apresenta a seguinte classificação para as categorias, comparando as categorias de Aristóteles e de Ranganthan: Entidades (princípios, objetos materiais, objetos imateriais) – em Aristóteles seriam as substâncias e, em Ranganthan, Personalidade; Qualidades (qualidades, quantidades, relação) – em Aristóteles se apresenta distinguida pelas categorias qualidade, quantidade e relação e, em Ranganthan, é representada pela categoria matéria; Atividades – que são as ações, processos e estados em Aristóteles e, em Ranganthan, denomina-se energia; por último, Dimensão – que em Aristóteles é considerada como as categorias lugar, tempo, condição e, em Ranganthan, é identificada pelas categorias espaço e tempo.

As categorias são utilizadas para a ordenação dos conceitos, mas, no nível do referente, do extralinguístico, elas também são utilizadas para se fazer perguntas sobre o domínio: em determinado domínio existem princípios, objetos materiais, imateriais, qualidades, processos, condição, entre outros?

A abordagem onomasiológica não inicia com o levantamento de *um corpus terminológico*, mas com a identificação dos referentes que podem popular as categorias – que funcionam aqui como metarrepresentações do domínio. Essa forma de adentrar no domínio nos leva, no caso das ontologias em que desejamos que o conhecimento possível possa ser explicitado, a poder realizar a captura do conhecimento auxiliado não somente pelo especialista, mas também pela consulta a glossários ou outros textos de especialidade sobre o domínio, para encontrar conceitos que respondam positivamente ou não a essas categorias. Esse é um procedimento de idas

e vindas, de análise e síntese. E, assim, quando se dá a identificação a qual categoria o referente pertence, é possível determinar as relações (lógicas, de parte – todo, funcionais) no interior dessas categorias, de forma a se construir um modelo de domínio para ontologias. A classificação desses conceitos no interior do domínio nos permitirá então determinar a sua enunciação por meio de uma definição que espelhe o acordo estabelecido.

Nesse sentido, mesmo que sejam utilizadas em um primeiro momento definições como um texto de especialidade, elas sofrerão alterações, para que seja possível evidenciar uma representação que expresse o compromisso ontológico assumido, no interior do mundo possível a ser representado. As definições são fruto desse acordo estabelecido no interior do domínio.

CONCLUSÕES

Neste estudo apresentamos discussões que envolvem as abordagens semasiológicas e onomasiológicas para a captura do conhecimento sobre domínios de especialidade. No caso de ontologias, o modelo de domínio elaborado deve ser expresso por meio de uma taxonomia que espelhe as conceituações e também de axiomas que descrevem as operações do referente no interior desse domínio. As definições em ontologias devem atender à leitura de humanos e máquinas e, assim, são a soma das descrições expressas na taxonomia e nos axiomas. Portanto, as definições em ontologias, além de serem apresentadas em forma de texto para serem lidas por humanos, devem também ser apresentadas por meio de axiomas que descrevem o objeto e suas relações no interior de um dado domínio. Consideramos que a abordagem onomasiológica, pautada em uma forma de captura de conhecimento, adentra o domínio através do entendimento dos objetos de conhecimento desse domínio, ou seja, que parte do extralinguístico apresenta um potencial metodológico bastante adequado à elaboração de modelos de domínio em ontologias, pois permite uma compreensão consensual de como os referentes se relacionam em um mundo fechado.

Consideramos, assim, que o papel das definições em uma ontologia é, portanto, precisar os atributos do referente e suas relações de forma consensual, assim assegurando uma herança através de um tipo de hierarquia, além da descrição de outros elementos que estão em relação no interior do domínio. A coerência nas definições e, por conseguinte, na classificação, exige uma perspectiva unificadora requerida pelo domínio a ser especificado, para que se possa atender a uma representação consistente e que atenda a determinado propósito.

AGRADECIMENTOS

Agradecimentos ao apoio da Agência de fomento CNPq/Brasil, pelo auxílio através de bolsa de produtividade pesquisa que possibilitou o desenvolvimento dos estudos.

REFERÊNCIAS

BACHIMONT, B. *Qu'est-ce c'est une ontologie?*. *Technolangue.net*, 3 juil. 2006. Disponível em: <http://www.technolangue.net/imprimer.php?id_article=280>. Acesso em: 15 ago. 2017.

CABRÉ, T. Theories of terminology their description, prescription and explanation. *Terminology*, v. 9, n. 2, p. 163-199, 2003.

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

_____. O papel das definições na pesquisa em ontologia. *Perspectivas em Ciência da Informação*, v. 15, n. 1, p. 220-238, 2010. Disponível em: <<http://basessibi.c3sl.ufpr.br/brapci/v/a/8769>>. Acesso em: 15 ago. 2017

_____. ROCHA, R. O papel das definições na formação de corpus terminológico para elaboração de ontologias de domínios: uma experiência a partir de extração automática e manual. In: CONGRESSO ISKO ESPANHA-PORTUGAL, 3., 2017, Lisboa. *Anais...* Lisboa: Faculdade de Letras da Universidade de Coimbra, 2017. No prelo.

CEUSTERS, W.; SMITH, B. Foundations for a realist ontology of mental disease. *Journal of Biomedical Semantics*, v. 1, n. 10, p. 1-23, Apr. 2010.

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

_____. Conceptual definitions for interconcept. *International Classification*, v. 1, n. 8, p.16-22, 1981.

_____. *Die pilotstudie DB-Thesaurus*. Frankfurt: Deutsche Bibliothek, 1980.

_____. *Ontical structures and universal classification*. Bangalore: Sarada Ranganathan Endowment for Library Science, 1978a.

_____. Terminological definitions: characteristics and demands. In: PROBLÈMES de la définition et de la synonymie en terminologie. Québec: Girsterm, 1983. p. 13-51.

_____. *Wissensorganisation: Entwicklung, Aufgabe, Anwendung, Zukunft*. W✓rzburg: Ergon, 2014. p. 35-51

DAVIS, R.; SHROBE, H.; SZOLOVITS, P. What is knowledge representation?. *AI Magazine*, n. 14, v. 1, p.17-33, 1993.

DING, Y.; FOO, S. Ontology research and development part 2: a review of ontology generation. *Journal of Information Science*, v. 5, n. 28, p. 375-388, 2002.

FERNANDES, P. C. B.; GUIZZARDI, R. S. S.; GUIZZARDI, G. Using goal modeling to capture competency questions in ontology-based systems. *Journal of Information and Data Management (JIDM)*, v. 2, n. 3, p. 527-540, 2011.

FERNANDEZ-LÓPEZ, M. Overview of methodologies for building ontologies. In: WORKSHOP ON ONTOLOGIES AND PROBLEM-SOLVING, METHODS: LESSONS LEARNED AND FUTURE TRENDS, 1999, Stockholm. *Proceedings...* Stockholm: Richard Benjamins, 2001. p. 26-34.

_____. GÓMEZ-PÉREZ, A.; JURISTO, N. Methontology: from ontological art towards ontological engineering. In: SPRING SYMPOSIUM ON ONTOLOGICAL ENGINEERING, 1997, Palo Alto, CA. *Proceedings...* Menlo Park, CA: The AAAI Press, 1997. p. 33-40.

GANGEMI, A.; PISANELLI, D. M.; STEVE, G. An overview of the ONIONS Project: applying ontologies to the integration of medical Terminologies. *Data Knowledge Engineering*, v. 2, n. 31, p. 183-220, 1999.

GOMES, H.; CAMPOS, M. L. A. Aspectos teóricos aplicados às linguagens documentárias: módulo 2. 2002. Material didático para curso à distância no Sesc.

_____. GUIMARÃES, L. S. Organização da informação e terminologia: a abordagem onomasiológica. *DataGramZero*, v. 11, n. 5, out. 2010. Disponível em: <<http://www.brapci.ufpr.br/brapci/v/a/9489>>. Acesso em: 15 ago. 2017

GRUBER, T. A translation approach to portable ontology specifications. *Knowledge Acquisition*, London, v. 5, n. 2, p. 199-220, 1993b.

_____. Towards principles for the design of ontologies used for knowledge sharing. *International Journal Human Computer Studies*, v. 43, n. 5/6, p. 907-928, 1993a.

- GRUNINGER, M.; FOX, M. S. Methodology for the design and evaluation of ontologies. In: WORKSHOP ON BASIC ONTOLOGICAL ISSUES IN KNOWLEDGE SHARING, AT THE INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 14., 1995, Montreal. *Proceedings...* Montreal: IJCAI, 1995.
- GUARINO, N.; WELTY, C. Evaluating ontological decisions with OntoClean. *Communications of the ACM*, v. 45, n. 2, p. 61-65, Feb. 2002.
- GUIZZARDI, R. S. S.; MORALES-RAMIREZ, I.; PERINI, A. A goal-oriented analysis to guide the development of a user feedback ontology. In: WORKSHOP CO-LOCATED WITH CHAISE. CEUR WORKSHOP, 7., 2014, Thessaloniki. *Proceedings...* Thessaloniki: CEUR-WS, 2014. p. 16-17. v. 1157.
- HEGENBERG, L. *Definitions: termos teóricos e significado*. São Paulo: Cultrix, 1974.
- HERMANS, A. La définition des termes scientifiques. *Meta*, v. 34, n. 3, p. 529-532, 1989.
- LARA, M. L. G. Terminologia e Ciência da Informação. In: ALVES, I. M. (Org.). *Os estudos lexicais em diferentes perspectivas*. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas, 2010. p. 8-16. v. 2.
- MARCONI, D. *Lexical competence: language, speech and communication series, a Bradford book*. Cambridge: The MIT Press, 1997.**
- METAXIOTIS, K.; PSARRAS, J. E.; ASKOUNIS, D. Building ontologies for production scheduling systems: towards a unified methodology. *Information Management & Computer Security*, v. 9, n. 1, p. 44-51, 2001.
- MIZOGUCHI, R. What ontological engineering can do for solving real-world problems. *Knowledge and Systems Engineering Advances in Intelligent Systems and Computing*, v. 244, p. 3-3, 2014.
- NEWELL, A. The Knowledge level. *Artificial Intelligence*, v. 18, p. 87-127, 1982. Disponível em: <<http://repository.cmu.edu/cgi/viewcontent.cgi?article=2616&context=compsci>>. Acesso em: 15 ago. 2017.
- PASLARU-BONTAS, E. *A contextual approach to ontology reuse: methodology, methods and tools for the semantic web*. 2007. Dissertação (Mestrado em Matemática)- Freien Universitat Berlin, Berlin, 2007.
- POLI, R. *Framing ontology*. 2001. Disponível em: <[http://www.formalontology.it/Framing first.htm](http://www.formalontology.it/Framing%20first.htm)>. Acesso em: 15 ago. 2017.
- POTTIER, B. *Sémantique générale*. Paris: PUF, 1992.
- RIGGS, F. W. Information Science and social science: the need for onomantics. *International Forum for Information and Documentation*, v. 1, n. 14, p.12-21, 1989a.
- _____. Onomantics and terminology. *Knowledge Organization*, n. 23, p. 25-3, 1996.
- _____. Terminology and lexicography: their complementarity. *Terminology and lexicography: their complementarity*, v. 2, n. 2, p. 90-110, 1989b.
- _____. Terminology for the Social Sciences. In: INTERNATIONAL SYMPOSIUM THEORETICAL AND METHODOLOGICAL PROBLEMS OR TERMINOLOGY, 1979, Moscow. *Proceedings...* Moscow, 1979. p. 591-606.
- ROCHE, C. Lexical and conceptual structures in ontology. In: INTERNATIONAL CONFERENCE ON INDUSTRIAL, ENGINEERING AND OTHER APPLICATIONS OF APPLIED INTELLIGENT SYSTEMS, 2006, Annecy. *Proceedings...* Heidelberg: Springer Berlin / Heidelberg, 2006. p. 1034-1041.
- SANTOS, C. S. A. *Terminologia e ontologias: metodologias para representação do conhecimento*. 2010. 245 f. Tese (Doutorado em Linguística) – Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa, Departamento de Línguas e Culturas, Universidade de Aveiro, Lisboa, 2010. Disponível em: <<https://ria.ua.pt/bitstream/10773/2876/1/2010000702.pdf>>. Acesso em: 15 ago. 2017.
- SEPPÄLÄ, S. An ontological framework for modeling the contents of definitions. *Terminology*, v. 21, n. 1, p. 23-50, June 2015.
- _____. SCHREIBER, Y.; RUTTENBERG, A. Textual and logical definitions in ontologies. In: INTERNATIONAL WORKSHOP ON DEFINITIONS IN ONTOLOGIES, 2., 2014, Houston. *Proceedings...* Buffalo: University at Buffalo, 2014. Disponível em: <<http://ceur-ws.org/Vol-1309/paper3.pdf>>. Acesso em 15 ago. 2017.
- SURE, Y.; STAAB, A.; STUDER, R. Methodology for development and employment of ontology based knowledge management applications. *SIGMOD Record*, v. 31, n. 4, p.18-23, 2002.
- TAIRAS, R.; MERNIK, M.; GRAY, J. Using ontologies in the domain analysis of domain-specific languages. In: INTERNATIONAL WORKSHOP ON TRANSFORMING AND WEAVING ONTOLOGIES IN MODEL DRIVEN ENGINEERING, 1., 2008, Toulouse. *Proceedings...* Toulouse: Springer, 2008. Disponível em: <<http://ceur-ws.org/Vol-395/paper02.pdf>>. Acesso em: 15 ago. 2017.
- TÁLAMO, M. F. T. G. A. M.; LARA, M. L. G. Interfaces entre linguística, terminologia e documentação. *Brazilian Journal of Information Science*, v. 3, n. 2, p. 54-70, 2009. Disponível em: <<http://www.brapci.ufpr.br/brapci/v/a/9185>>. Acesso em: 15 ago. 2017.

USCHOULD, M.; KING, M. Towards a methodology for building ontologies. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, WORKSHOP ON BASIC ONTOLOGICAL ISSUES IN KNOWLEDGE, 14., 1995, Montreal. *Proceedings...* Montréal: DBLP, 1995. Disponível em: <<http://www.aiai.ed.ac.uk/project/oplan/documents/1995/95-ont-ijcai95-ont-method.pdf>>. Acesso em: 15 ago. 2017.

VICKERY, B. C. Knowledge representation: a brief review. *Journal of documentation*, v. 42, n. 3, p. 145-59, Sept. 1986.

_____. Ontologies. *Journal of Information Science*, v. 23, n. 4, p. 227-286, 1997.

VILLELA, M. L. B.; OLIVEIRA, A. P.; BRAGA, J. L. Modelagem ontológica no apoio à modelagem conceitual. In: SIMPÓSIO BRASILEIRO DE ENGENHARIA DE SOFTWARE, 18., 2004, Brasília, DF. *Anais...* Brasília, DF: SBES, 2004. p. 241-256.