

Princípios para modelagem de domínio: a posição de Barry Smith e de Ingetraut Dahlberg

Maria Luiza de Almeida Campos

Pós-Doutorado pela Universidade de São Paulo (USP) – São Paulo, SP - Brasil. Pós-Doutorado pela Fundação Oswaldo Cruz (Fiocruz) – Rio de Janeiro, RJ - Brasil. Doutora em Ciência da Informação pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e a Universidade Federal do Rio de Janeiro (UFRJ) – Rio de Janeiro, RJ – Brasil. Professora da Universidade Federal Fluminense (UFF) - Niterói, RJ – Brasil.

<http://lattes.cnpq.br/9545682339961651>

E-mail: marialuizalmeida@gmail.com

Hagar Espanha Gomes

Livre-docência pela Universidade Federal Fluminense (UFF) - Niterói, RJ – Brasil. Mestre em Ciência da Informação pelo Instituto Brasileiro de Bibliografia e Documentação (IBBD/IBICT) – Rio de Janeiro, RJ – Brasil.

<http://lattes.cnpq.br/9943296027300097>

E-mail: hagarespanhagomes@gmail.com

Recebido em: 10/07/2014. Aprovado em: 09/07/2015. Publicado em: 15/01/2016.

Resumo

Na elaboração de modelos de domínios, é fundamental a determinação da unidade de representação que irá compor a estrutura de classes deste domínio - neste caso a discussão do que se pode determinar como conceito. Para tal, considera-se necessário se ter claro em que corrente filosófica se está apoiando, quando se refere a conceitos em diversos contextos. Apoiar-se em dois autores clássicos da área de ciência da informação e da ciência da computação, respectivamente Ingetraut Dahlberg e Barry Smith, para introduzir-se uma discussão sobre o conceito de “conceito”. Além disso, discutem-se questões relacionadas à estruturação/classificação de um domínio, apontando em última análise convergências e divergências na postura teórica de cada um desses autores, propiciando maior entendimento para aqueles que estão preocupados com a modelagem de domínios de conhecimento/atividades. Por último, abordam-se questões relativas às relações semânticas e definições no intuito de evidenciar pontos fundamentais para estruturação de domínios de conhecimento.

Palavras-chave: Modelagem de domínio. Ontologias. Sistema de conceitos.

Guiding principles for domain modelling: Barry Smith and Ingetraut Dahlberg standpoints

Abstract

Representational unit determination is fundamental for structuring classes when modelling domains – in this case, the discussion on what can be determined as a concept. Therefore, it is the philosophical current that supports what is meant by concept in several contexts has to be made clear. Two classical theorists of Information Science and Computer Science, respectively Ingetraut Dahlberg and Barry Smith, provide the support for introducing a discussion about ‘concept’. Furthermore, issues relating to domain structuring/classification are discussed, identifying at the end of the analysis convergences and divergences in their theoretical points of view, offering a better understanding to researchers dealing with knowledge /activities domain modelling. Finally, issues about semantic relations and definitions are addressed, in order to highlight key points for knowledge domain structuring.

Keywords: Domain modelling. Ontologies. Concept systems.

Principios para modelaje de dominio: la posición de Barry Smith y de Ingetraut Dahlberg

Resumen

En la elaboración de modelos de dominios, es fundamental determinar la unidad de representación que compondrá la estructura de clases de este dominio – en este caso la discusión de lo que se puede determinar cómo concepto. Para tanto, se considera necesario estar claro en que corriente filosófica está el apoyo, cuando se refiere a conceptos en contextos diversos. Dos autores clásicos del área de ciencia de la información y de la ciencia de la computación, respectivamente Ingetraut Dahlberg y Barry Smith, son el apoyo para introducir una discusión sobre el concepto de “concepto”. Además, son discutidas cuestiones relacionadas a la estructuración/clasificación de un dominio, señalando en último análisis convergencias y divergencias en la postura teórica de cada autor, ofreciendo mayor comprensión para los que se preocupan con el modelaje de dominios de conocimiento/actividades. Por último, son abordadas cuestiones relativas a las relaciones semánticas y definiciones, resaltando puntos fundamentales para estructurar dominios del conocimiento.

Palabras clave: Modelaje de dominio. Ontologías. Sistema de conceptos.

INTRODUÇÃO

A busca por informação ágil e precisa tem sido um problema cada vez mais atual, apesar de não ser um problema novo. Existe uma diversidade de pontos de vista a serem considerados, cada qual com seu conjunto de conceitos relevantes e formas de definir e nomear esses conceitos. Além disso, existem diferentes modos de representar esses conceitos e suas relações, ainda que de maneira limitada, em um modelo que procura expressar a semântica pretendida pelos que irão utilizá-lo.

A necessidade da informação ágil e a crescente utilização de recursos computacionais têm mudado um pouco o papel do profissional de informação, ante a uma nova realidade que se impõe cada vez mais e com mais facetas.

Se em um primeiro momento, por exemplo, no âmbito das bibliotecas; o profissional de informação atuava como intermediário para as demandas de usuários em suas buscas; hoje, este papel é desempenhado parcialmente por sistemas informatizados. Para isso, os sistemas necessitam explorar com mais *inteligência* as informações que podem estar disponíveis à busca do usuário.

No segundo momento, o uso de sistemas informatizados apoiados por tesouros e metadados vem deslocar o papel do profissional de informação

para interagir com esses sistemas em apoio às buscas dos usuários e para ajudar na elaboração de novas linguagens que subsidiam essas buscas¹.

Mas, como se sabe, a busca apoiada por linguagens documentárias como os tesouros e vocabulários controlados ainda possui uma semântica muito restrita do ponto de vista da máquina. Esses instrumentos controlam o significado processado apenas por humanos e não por máquinas.

Cabe destacar que a máquina não possui inteligência, estando limitada a processar instruções, programadas por pessoas, que vão atuar sobre dados de entrada. Esses dados, se não forem devidamente caracterizados, não permitirão uma recuperação precisa, sem ambiguidades. No caso de programas de busca, quanto mais precisas forem as descrições utilizadas para as buscas, menos ambígua será a recuperação.

Nesse cenário, novos instrumentos de representação da informação têm sido necessários, instrumentos estes que possuem um grau de formalismo que torne possível o processamento por máquinas, para resolver problemas e

¹ Cabe observar que os padrões de metadados muitas vezes têm sido usados para descrever não o conteúdo, mas as características do documento em si, como, por exemplo, seu autor e título. Enquanto o uso de uma terminologia padrão, como um vocabulário controlado, apóia a busca por palavras-chaves que podem descrever o conteúdo do documento.

questões apresentadas pelos usuários em suas necessidades de busca. Porém, para isso, os novos instrumentos e seu formalismo, por si só, não vão ser suficientes. Existe uma lacuna, a ser ocupada pelo profissional da informação, que vem justamente reinventar o seu papel de intermediário; dessa vez não como a pessoa que se impõe entre o usuário e um acervo informacional, mas como a pessoa que se posiciona entre a questão a ser atendida e a máquina, de acordo com a perspectiva do usuário.

Ou seja, esse novo intermediário deve traduzir para os formalismos que a máquina entende os conceitos que possam descrever o problema, tal como o usuário o reconhece. Esse sem dúvida é um papel mais complexo. É preciso lidar com os diferentes pontos de vista dos usuários e seus jargões, mas também com os jargões da máquina, em seu formalismo, hoje ainda de semântica limitada, para comportar o complexo no simples, com a menor perda possível de expressividade. Porém, nem sempre o profissional responsável pela tarefa de modelar o domínio possui um conhecimento claro sobre os objetos do domínio e termos que o denotam, como é o caso, na maioria das vezes, do profissional de informação, o que pode resultar em modelos conceituais inconsistentes. No entanto, isto pode ser evitado desde que se possuam as necessárias bases teóricas e metodológicas.

É neste contexto que se coloca em discussão a necessidade de estudos mais aprofundados de teorias e métodos relativos à elaboração de estruturas classificatórias para a representação e recuperação de conhecimento/informação. Esses estudos passam pela questão de desenvolver estratégias e abordagens para a construção de modelos conceituais de domínios e atividades mais consistentes; - e que consigam atingir uma representação de alta qualidade. Estes modelos se apresentam atualmente em instrumentos semânticos como taxonomias, tesauros e ontologias.

Um dos fatores fundamentais na elaboração de modelos conceituais de domínios é a determinação da unidade de representação que irá compor a estrutura de classes deste domínio, ou seja, de que natureza são os elementos de representação que serão utilizados na taxonomia de um domínio. As diversas Teorias de Representação utilizadas

no âmbito da elaboração de modelos de domínios estão apoiadas em correntes filosóficas, que nem sempre são explicitadas pelos autores que as adotam, e isto acarreta uma série de consequências no ato de modelar um domínio, como por exemplo: o entendimento do que pode ser nomeado como instância, classe e mesmo conceitos, entre outros elementos que estão intrinsecamente relacionados com o conceito de “conceito”.

Neste sentido consideramos problematizar, neste artigo, alguns aspectos da discussão sobre as unidades de representação de um domínio, neste caso a discussão do que podemos determinar como sendo conceito. Para tal consideramos necessário se ter claro em qual corrente filosófica estamos nos apoiando, quando nos referimos a conceitos em diversos contextos, para não adotarmos uma postura ingênua diante das decisões tomadas. Assim, apresentaremos uma discussão realizada por Barry Smith sobre o conceito de “conceito” a partir de correntes filosóficas diversas e como isto atinge, de certa forma, nossas decisões na elaboração de modelos de domínios, no caso em ontologias.

A partir de tal discussão pretendemos nos deter em um estudo sobre as posições de Barry Smith e de Ingetraut Dahlberg no que diz respeito à estruturação/classificação de um domínio, apontando em última análise, as convergências e divergências na postura teórica de cada um para possibilitar maior entendimento aqueles que estão preocupados com a modelagem de domínios de conhecimento/atividades.

Pretendemos, também, a partir de tais discussões, introduzir as questões relativas às relações semânticas, pois consideramos que na representação de um domínio dois itens são essenciais para seu adequado entendimento: nós e arestas. Arestas (relações) ligam os nós (conceitos/termos) produzindo grafos. Além disso, ressaltamos o papel das definições em ontologias de domínios. Assim, acreditamos estar abordando os diversos aspectos de Modelagem de Domínios, ou seja, sua unidade de representação, as relações entre tais unidades e a explicitação do compromisso ontológico através de suas definições.

O QUE SE PODE DENOMINAR CONCEITO NAS DIVERSAS CORRENTES FILOSÓFICAS

No intuito de esclarecer o emprego do termo “conceito”, Barry Smith (SMITH, 2004; GRENON; SMITH, 2011) apresenta um conjunto de definições encontrados na literatura para o termo. Para tanto, introduz o problema dos Universais para evidenciar a necessidade de se ter claro em qual corrente filosófica estamos nos apoiando, quando nos referimos a conceitos em diversos contextos; e ele esclarece que contradições estão surgindo porque não está consistentemente evidenciada na literatura a posição dos autores para tais distinções.

O problema dos Universais é uma temática que preocupa filósofos há muito na história da Humanidade. No campo da Representação do Conhecimento e de áreas relacionadas, inclusive aquelas envolvidas com o desenvolvimento de ontologias, tal temática deve ser discutida visando à consistências dos modelos representados. Neste sentido, Barry Smith (2004) defende a tese que as ontologias desenvolvidas como apoio às pesquisas em Ciências Naturais deveriam adotar como tema não os conceitos, mas antes os universais e particulares que existem na realidade e são apreendidos por leis científicas. E, dependendo de como estes Universais são considerados do ponto de vista do modelador, diz que adota em tais modelos uma posição realista, conceitualista ou nominalista. (SMITH, 2004).

Realismo é o nome que se dá a uma posição adotada sobre o modo como conhecemos a realidade na Teoria do Conhecimento ou da Metafísica. Em ambos os casos, o realismo não se opõe ao nominalismo, mas ao conceitualismo. Podemos afirmar *grosso modo* que as diversas correntes realistas possuem duas preocupações fundamentais: como conhecemos a realidade – realismo gnoseológico, e como se dá o modo de ser do real – realismo metafísico (MORA, 2001). Neste artigo, nos interessam questões que envolvem o realismo gnoseológico, que pode ser classificado como realismo ingênuo, em que se pressupõe

que o conhecimento é uma reprodução exata (uma cópia fotográfica) da realidade, ou realismo científico, também denominado empírico ou crítico, no qual se postula que não se pode simplesmente equiparar o percebido com o verdadeiramente conhecido e que é preciso submeter o dado ao exame e ver - para depois levar em conta quando se formulam juízos definitivos - o que há no conhecimento que não é mera reprodução. Os realistas defendem a tese que há Universais e Particulares; - os primeiros podem ser definidos como padrões invariantes, também denominados características ou essências, existindo junto a entidades da realidade. Os Particulares são instâncias desses Universais que existem no mundo real do espaço e tempo. Cada Universal pode ser exemplificado por uma pluralidade de particulares. Assim, “o termo ‘universal’ significa aquilo que as instâncias correspondentes – por exemplo, todas as baleias, todas as enzimas – têm em comum. Sendo que os “universais e suas instâncias gozam de um relacionamento simbiótico: um não pode existir sem o outro” (SMITH, 2004). De acordo com a visão realista, estas relações de similaridade (aquilo que existe em comum entre instâncias) poderão existir mesmo se não houvesse pessoas na posição de observá-las. A relação de similaridade entre as instâncias existe porque o mesmo Universal é instanciado em cada uma das formas individuais dadas. Universais são organizados em árvores de gênero e espécie – são conectados através de relações de subtipo (é um). Assim, no interior da abordagem realista, o conceito pode ser entendido como um Universal.

Contrariamente aos realistas, os conceitualistas defende que não existem Universais junto à realidade, mas somente em nossas mentes. Diferentes pessoas podem compartilhar o mesmo conceito geral, que são neste sentido múltiplos exemplos em diferentes mentes. Entretanto, conceitualistas não acreditam em geral que um conceito que está na mente de uma pessoa possa corresponder a Universais ou invariantes junto à realidade, mas que todo

conceito relaciona-se com suas instâncias em somente um caminho *ad hoc*, ou seja, coisas que são construídas pelo sujeito em sua mente. Mais precisamente, os conceitos em geral são tratados como se fossem todos conceitos *ad hoc*. Na abordagem conceitualista, o conceito de “conceito” pode ser entendido como ideias gerais.

Os nominalistas defendem que não existem universais na realidade e nem em nossa mente, mas somente existem termos gerais. Os nominalistas então negam a existência de conceitos gerais que podem ser compartilhados por uma pluralidade de pessoas. Termos gerais são meras etiquetas de coleções *ad hoc* de coisas particulares ou eventos.

A posição Nominalista surge na Idade Média (MORA, 2001) e foi adotada porque se supunha que admitir Universais (ideias) na mente de Deus era limitar de algum modo a onipotência divina, e admitir Universais (ideias formas) nas coisas significaria supor que as coisas têm, ou podem ter, ideias ou modelos próprios, com o que também se limita a onipotência divina. O nominalismo consistiu em afirmar que um Universal (como uma espécie ou um gênero) não é nenhuma entidade real, nem tampouco está nas entidades reais: é um som vocal, *flatus vocis*. Os universais não se encontram *ante rem*, não estão antes da coisa, ou são anteriores à coisa, como sustenta o realismo (platônico). Não estão tampouco, *in re* – na coisa –, como sustenta o conceitualismo, o realismo moderado, ou o aristotelismo. Os Universais são simplesmente *nomina*, nomes, vocábulos, ou *termini*. O nominalismo sustenta que só têm existência real os indivíduos ou as entidades particulares. Assim, no interior da abordagem nominalista, o conceito de “conceito” pode ser nomeado como termos gerais.

Traços dessas três posições (realismo, conceitualismo e nominalismo) são comumente encontrados na literatura sobre ontologias e terminologia. Entretanto, segundo Barry Smith et al. (KLEIN; SMITH, 2005), como as três posições anteriormente apresentadas não estão claramente distinguidas, contradições são identificadas na literatura relevante.

A POSIÇÃO DE BARRY SMITH NA CONSTRUÇÃO DE ONTOLOGIAS DE DOMÍNIO

Em sua Teoria de Representação, Barry Smith defende que as ontologias devem se apoiar em uma abordagem realista. Esta abordagem busca prover o entendimento de quais entidades existem na realidade, onde “realidade” é aqui entendida no senso mais amplo possível, incluindo, por exemplo, não somente moléculas e planetas, mas todo trabalho da literatura, leis e épocas históricas. Os objetos da investigação ontológica estão dentro de um dado domínio e são entidades de primeira ordem² deste domínio, mais do que conceitos nas mentes das pessoas (especialistas, em particular) que estudam o domínio, ou termos usados (por especialistas, em particular) para se referir a determinado domínio e seus componentes.

Conceitos e termos podem, contudo, perfeitamente ser matéria de ontologias de domínios psicológicos e linguístico, mas, no entanto, como entidades de primeira ordem (como indivíduos). No domínio da filosofia, por exemplo, dizem Smith e seus seguidores, muitas das entidades que a ontologia destina a categorizar são conceitos (GRENON; SMITH, 2011).

Em ontologias de domínios científicos, suas unidades são representações das entidades como elas existem na realidade, e estas unidades devem ser designadas como Universais, aos quais correspondem termos gerais usados em afirmações científicas, sendo os particulares instâncias de tais Universais que existem no mundo real do espaço e tempo.

² A lógica de primeira ordem trata de noções apenas para linguagens formais de primeira ordem. Uma linguagem formal é de primeira ordem se, do ponto de vista de sua semântica lógica, os domínios de suas possíveis interpretações são domínios aos quais apenas pertencem indivíduos (por oposição à classes de indivíduos, a classe de classe de indivíduos etc.) e, se do ponto de vista de sua sintaxe lógica, os quantificadores se ligam apenas a variáveis individuais, por exemplo, por oposição às variáveis de predicado. (BRANQUINHO, et al., 2006).

Neste sentido, encontramos em seus estudos uma distinção entre o que Smith chama de ontologia baseada em uma postura realista ou outra. Na primeira, podemos definir como uma ontologia construída por termos que se destinam a referir exclusivamente a tipos e que correspondem àquela parte do construído de uma teoria científica que é capturada pelos seus termos gerais constituintes e suas inter-relações³. A segunda inclui representações artificiais compreendendo uma taxonomia como parte própria, cuja unidade representacional designa algumas combinações de tipos, classes e certas relações entre eles. Consequentemente, uma ontologia baseada no realismo se refere a um sistema de tipos assim como um sistema de conceitos se refere a conceitos (KLEIN; SMITH, 2005).

Considerando que a terminologia da ontologia não é ainda estabelecida em um caminho consistente no ciclo da informática e da terminologia, Barry Smith et al. (KLEIN; SMITH, 2005; SMITH, 2004) apresentam todo um estudo de princípios fundamentais para distinguir sua proposta teórica das demais propostas atualmente existentes⁴. Para a compreensão desses princípios é necessário primeiramente se ter claro que, segundo Smith, sistemas de conceitos são sistemas de significado, enquanto os sistemas de representação em ontologias são relativos a entidades do mundo real;

³ Sua proposta teórica está baseada em princípios como: a noção de entidade, que são de duas espécies, tipos e instâncias. Estas podem ser classificadas de acordo com sua persistência no tempo e no espaço, ou seja: ocorrentes, também denominados perdurantes, que podem ser exemplificados como processos (estendido no tempo) e eventos (momentâneo, transitório); continuantes, também denominados endurantes, que são de dois tipos: independentes e dependentes. Continuantes independentes são coisas (objetos do nosso dia a dia) nos quais os continuantes dependentes (qualidades, funções, papéis e disposições) são inerentes. Os continuantes dependentes possuem uma relação de dependência existencial. Eles só existem porque são qualidades, funções, papéis e disposição de algum continuante independente. Para maior entendimento de sua teoria, onde outras noções são apresentadas, como, por exemplo, a noção de função, papel e disposição consultar o artigo "Function, Role and Disposition in Basic Form Ontology" (ARP; SMITH, 2008).

⁴ Podemos citar como exemplos de outras Teorias de Representação, no âmbito das ontologias, aquelas que embasam as ontologias "Descriptive Ontology for Linguistic and Cognitive Engineering" (DOLCE) e; a "Unified Foundational Ontology" (UFO).

tanto os investigados pela ciência natural (por exemplo, células ou elétrons) quanto os existentes em domínios administrativos.

Sua abordagem filosófica trata de questões que venham a ressaltar o papel das definições, de significado e de referência para a elaboração de ontologias de domínio. Esta abordagem permite, por meio de exame cuidadoso e análise lógica, identificarmos ambiguidades e erros na estruturação dos elementos de uma ontologia, o que permite a representação consistente do domínio visando a possibilitar o raciocínio lógico. Barry Smith et al. (ARP; SMITH, 2008) consideram que as ontologias não devem ser elaboradas a partir de mecanismos automáticos, mas sim serem elaboradas por humanos, pois não existem atualmente mecanismos automáticos com a consistência necessária para processamento de linguagens naturais. Sua experiência mostrou que, dados os problemas que apresenta, a revisão do produto do processamento de linguagem automatizada é uma tarefa que raramente leva a resultados estruturalmente sólidos. Técnicas automatizadas produzem redes de termos "associados", que são chamados de "redes lexicais". Estas redes não são as mais apropriadas para representar a estrutura de um domínio. A informação terminológica e lexical com base em vínculos de coocorrência é útil para determinados fins de recuperação, mas ela não fornece uma representação confiável de um domínio.

Outro problema para as limitações das abordagens lexicais é a falta de interoperabilidade que promovem. Isto porque, até quando o mesmo termo aparece em pluralidade de tais sistemas, não há qualquer garantia que esteja sendo tratado de forma semelhante. Um critério importante na qualidade das ontologias é que elas devem, tanto quanto possível, abraçar um princípio de modularidade, ou seja, convergir em uma ontologia única para cada domínio, onde ontologias de domínios vizinhos devem ser trabalhadas juntas.

Como apontado, a abordagem defendida baseia-se em uma visão de ontologias como consistindo

em representações de entidades no domínio da realidade para a qual a ontologia se refere. Somente na base de representações deste tipo ele acredita ser possível fazer progressos coerentes na ligação entre diferentes sistemas de terminologia (por exemplo, em diferentes línguas). Neste sentido, dois princípios são fundamentais para a elaboração de ontologias consistentes: **o princípio da relevância e o princípio da modularidade**. Através do princípio da relevância é possível identificar entidades e características de entidades que pertencem exclusivamente ao domínio selecionado, por exemplo, se estivermos interessados na Filosofia de Bertrand Russel, selecionamos entidades deste domínio e não de sua biografia como ativista político (GRENON; SMITH, 2011). E através da observância ao princípio da modularidade é possível a inclusão de novos elementos de forma consistente a um domínio selecionado.

Quanto ao conceito de “conceito”, defende Smith (2004) que ao utilizarmos esta expressão deveremos estar nos referindo exclusivamente ao significado de um termo geral, em que este significado deve ser único e acordado pelas pessoas responsáveis por determinado campo disciplinar. Neste sentido consideramos que a Teoria da Representação apresentada por Barry Smith se apoia na corrente do Realismo Científico, pois um dos aspectos considerados ao sentido dado aos Universais passa pelo exame de pessoas responsáveis em um campo disciplinar.

Ainda sob o aspecto do que consideramos conceito como unidades representacionais em domínios de conhecimento, Smith apresenta uma discussão que vai de encontro à posição de E. Wüester (1981) em sua Teoria Geral da Terminologia, que consideramos importante ressaltar visando posteriores esclarecimentos. Wüester, segundo Smith et al. (2005), apresentou uma teoria que estava fadada ao fracasso quanto ao conceito de conceito, ressaltando inclusive que esta posição vem provocando confusão com sérias consequências em projetos de informática médica. Argumenta que os conceitos não são claramente definidos e que a definição é frequentemente vista como o resultado significativo

de um processo cognitivo, por exemplo, na forma do conhecimento na mente de um especialista ou na entrada em um sistema: “problemas então surgem, porque simultaneamente são feitas tentativas para usar o mesmo termo para fins de raciocínio entre relações no meio de entidades diferentes no mundo real” (KLEIN; SMITH, 2005, p. 433).

Evidencia que na abordagem de Wüester conceito é entendido como “conceito mental”, ideia, pensamento (*noesis* em Aristóteles), que diz respeito a certo estado da mente de alguns indivíduos, estado que pode ser invocado pelo uso de um termo geral correspondente. Considera que esta abordagem se apoia em uma visão psicológica, em que os conceitos são entidades mentais, analogamente relacionadas a ideias e crenças. Esta visão não pode ser considerada apropriada para uma proposta de padronização do domínio de sistemas terminológicos, pois considera que há alto grau de arbitrariedade e variabilidade quando conceitos são construídos na mente das pessoas. E considera que os pesquisadores que adotam tal postura precisam ainda encontrar um meio de explicar por que conceitos e suas características ora são “criaturas da mente” e ora são “propriedades de objetos no mundo”. Assim, não fica claro para as pessoas envolvidas com a tarefa de elaboração de uma terminologia se o que a constitui é a representação de ideias que estão na cabeça das pessoas, o significados das palavras, o conhecimento existente a partir do consenso de especialistas em uma disciplina ou os tipos de entidades no mundo (KLEIN; SMITH, 2005).

Também não considera apropriado o uso da expressão “unidade de conhecimento” adotado pela norma ISO 1087 para nomear conceito de conceito. Para ele, a melhor expressão seria significado consensual para evitar conotações psicológicas do mesmo. A defesa está pautada no entendimento de que podemos ter um **significado acordado de um termo** sem que este tenha correspondência com uma unidade de conhecimento. Por exemplo, unicórnio não corresponde a uma entidade da realidade sobre a qual o conhecimento possa ser obtido.

Assim, interpreta a expressão “conhecimento” como aquela que pode indicar existência na realidade. Entretanto, como veremos adiante, a expressão “unidade de conhecimento” não está relacionada com a questão da existência na realidade, e sim com um acordo firmado por um grupo de especialistas competentes em dado domínio de conhecimento; pelo menos é o que verificamos em Dahlberg (1978).

A POSIÇÃO DE INGETRAUT DAHLBERG NA ELABORAÇÃO DE SISTEMAS DE CONCEITOS

Em sua Teoria do Conceito, Dahlberg (1978) defende “conhecimento” como “conhecimento científico”, em que existe um acordo firmado sobre o sentido de uma unidade em um universo de discurso por um grupo de especialistas. Em seus estudos, Dahlberg é quem propõe esta expressão para fazer frente à expressão “unidade de pensamento” adotada por Wüester; para ela, tal expressão leva também a uma visão psicológica, que combate. A norma ISO 1087 adota a expressão “unidade de conhecimento” para a definição de conceito, após Dahlberg apresentar tais discussões em sua Teoria do Conceito.

Dahlberg (1978a), em seu trabalho *Estruturas Ônticas*, aponta para uma postura teórica na elaboração de sistema de conceitos que privilegia também uma abordagem realista e se apóia em teóricos como Vollmer (1975 apud DAHLBERG, 1978a), Campbell e Popper (1974 apud DAHLBERG, 1978a) para esta tomada de posição. Defende a tese que “todo reconhecimento da realidade é de natureza hipotética, e existe um mundo que é independente de nossa consciência, que é estruturado e autocontido. Este mundo é em parte reconhecível e compreensível por percepção, pensamento e ciência intersubjetiva” (DAHLBERG, 1978a, p.28). Afirma, ainda, que para os problemas de organização do conhecimento “é essencial que este argumento nos permita vencer a abordagem idealista existente de acordo com a qual a formação sobre o mundo ocorre exclusivamente através de formas de percepção de espaço e tempo dadas *a priori* e através das formas de pensar e de argumentar/raciocar.” (DAHLBERG, 1978a, p.28)

Como Smith, Dahlberg considera que a posição de Wüester quanto a definir o conceito como “unidade de pensamento” é inapropriada, pois pressupõe que adotemos uma postura psicológica e que o conhecimento permanece com o sujeito que o possui, como conceitos subjetivos. Segundo Vollmer (1975 apud DAHLBERG, 1978a, p.10), existe uma distinção entre três estágios de cognição, a saber: percepção (estrutura subscientes), cognição pré-científica e cognição científica. Esta distinção nos ajuda também a entender em que medida Dahlberg condiciona o seu entendimento de conceito como unidade de conhecimento.

Dahlberg defende que o domínio da linguagem comum ocorre nos dois primeiros estágios, ou seja, o da percepção e da cognição pré-científica. Os conceitos, como nomeados e usados por palavras de uma linguagem comum, devem ser conduzidos como propriedade mental e intelectual de cada ser humano. Neste estágio, “conceitos” são entendidos como “unidades de pensamento”. Pensar é um processo cognitivo, dependendo em sua totalidade das percepções e experiências de uma pessoa, suas reflexões sobre o mundo e verificações das aplicações corretas de seus “conceitos”. (DAHLBERG, 1978a, p.11)

Entretanto, quando falamos de conhecimento objetivo - aquele referente ao domínio da ciência, no qual a cognição científica se instaura -, falamos de um conhecimento sobre a realidade, onde se requer que a cognição científica esteja envolvida com conceitos que possam tornar-se explícitos pela linguagem e os quais possam ser apresentados de maneira formal. Estes “conceitos formais”, que diferem daqueles que uma criança adquire informalmente, são aqueles que devem ser considerados “conceitos científicos”, isto é, aqueles conceitos que são os elementos ou unidades de cognição e conhecimento objetivo/científico.

Uma questão que ainda se coloca é: “Como alguém adquire o conhecimento objetivo?” Dahlberg (1978a) em seu trabalho *Ontical Structures*, aponta esta questão:

Há certamente um longo caminho do conhecimento sobre o visível para conhecimento sobre o invisível e imaterial, mas tudo começa por relacionar pensamento

a objetos existentes na realidade. De predicções sobre seu meio-ambiente natural aos experimentos e mensurações, à contagem e inferências, o homem tem procedido para formular as leis da natureza e para aplicar essas leis em seu campo de atividade. Qualquer uma de suas afirmativas relacionadas com a realidade experimentada ou mensurada cria um “elemento de conhecimento”, o qual pode também ser considerado como um conceito básico ou primitivo. E a coleção de tais afirmativas ou elementos de conhecimentos referidos a um objeto da realidade ou a um objeto verificável de pensamento devem ser denominados unidade de conhecimento. Os conceitos científicos são tais unidades de conhecimento como deve ser descrito ou definido usando a linguagem natural (DAHLBERG, 1978a, p.11)

Assim, o conceito em Dahlberg é determinado a partir das seguintes ações: 1. selecionar um item de referência de uma realidade experimentada e mensurada; 2. formular afirmativas verificáveis, e neste sentido predicções verdadeiras sobre este item; 3. atribuir um nome/termo. O conceito assim pode ser representado por um triângulo tendo como **ápice** o *referente*, que é o representante de objetos gerais e individuais existentes em uma realidade científica/verificável. Além do referente, os outros vértices constituem as *características* que se podem inferir deste referente, a partir do contexto que ele está inserido, e uma *denominação* que é a expressão verbal utilizada para a comunicação. É necessário ressaltar que o que Dahlberg considera como características de conceitos não deve ser confundido com propriedades dos conceitos, tais como “bem definidos”, “difícil de definir”, “ser muito específico”, “pertencentes a outra categorias” etc., mas características como **propriedades declaradas dos objetos**, de itens de referência; somente no nível dos conceitos desses objetos é que elas se tornam as características dos conceitos. Existem tantas características quanto há afirmações sobre o conceito em um dado contexto; parece não haver limites para suas relações formais e /ou materiais.

Ela resalta em seus estudos que, das possibilidades de tipos de características existentes, duas são fundamentais: as características necessárias/essenciais e as acidentais. As primeiras são fundamentais para a

estruturação de sistemas de conceitos, e as acidentais podem especificar as diferenças que a caracterizam. (DAHLBERG, 1978a).

As características do conceito são os elementos, segundo Dahlberg, que permitem a elaboração de definições conceituais. Tais definições não seriam consideradas somente como produto da fala de um grupo de falantes, mas fundamentalmente como uma construção conceitual de um acordo firmado por este grupo, a qual permite a produção de um sistema de conceitos, trazendo como elemento agregador deste sistema as categorias/facetetas. Neste sentido, consideramos que Dahlberg, avança no uso das definições como um elemento de descrição do conceito que permite, além da fixação do conceito, seu posicionamento no próprio sistema de conceitos. Podemos observar que mesmo as ontologias de base realista como as defendidas por Smith et al., ainda tratam a definição como “especificação de um conceito, isto é o significado acordado de um termo, por meio de uma declaração descritiva ou uma expressão formal que serve para diferenciá-lo dos outros conceitos.” (KLEIN; SMITH, 2005). Entretanto, apesar de em seus trabalhos ressaltar a importância das definições, ele não aborda de que modo se apropriar das definições para a elaboração de ontologias de domínio. Dito de outra maneira, como elaborar as taxonomias que são as estruturas fundamentais de qualquer ontologia de domínio? Como vimos, Smith resalta que as “redes lexicais” não são as mais apropriadas para representar a estrutura de um domínio, mas não evidencia qual metodologia utilizar.

Neste sentido, afirmamos que Dahlberg avança no uso de definições e apoiada também em Riggs (1989a; 1989b; 1996) com sua “Onomântica”. Ela elabora uma forma alternativa de apresentação de estruturas de domínio, em que princípios de classificação estão presentes em seu estudo, porque se apóiam na sistematização dos conceitos (e não em sua listagem) através das características do conceito como elemento formador de tais estruturas.

A partir de nossos estudos, podemos afirmar que a concepção adotada por Dahlberg no uso de definições reflete uma concepção consensual da realidade sobre um objeto social, denominada “abordagem onomasiológica”. Nesta, tais estruturas de expressões refletem a forma como o referente é, conceitualmente; construído; vale dizer, pela seleção, realce e iluminação de alguns aspectos do ser e, por conseguinte, atenuação, ou mesmo ocultamento, de outros aspectos. 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 - no caso de Dahlberg, de especialistas -, interagem socialmente com o referente. E isto permite um movimento de visão sistêmica sobre o referente, em que as Categorias são o elemento agregador.

RELAÇÕES NA MODELAGEM DE DOMÍNIOS EM SISTEMAS DE CONCEITOS E ONTOLOGIAS

Relação é um aspecto que merece especial atenção dos desenvolvedores de ontologias. Dentre os vários tipos de relações, as mais familiares são as hierárquica e de partição, porém inúmeras outras estão presentes em um domínio. Se para os realistas alguns termos não precisam ser definidos (termos primitivos) em um contexto, isto não significa que não haja o recurso da linguagem natural para descrever os conceitos representados pelos termos, pois o objetivo é obter um padrão definitório para apoiar o processamento automático e a utilização pelas diversas bases de dados. E os diversos tipos de relação são explicitados na descrição (definição). Podemos classificar as relações em dois grupos: **lógicas e ônticas**⁵. As primeiras são abstratas, ocorrem entre conceitos. As segundas ocorrem entre objetos cuja contiguidade se dá no tempo ou no espaço. A descrição mais exata dos elementos de representação revela vários outros tipos de relação. Estas se associam a diferentes tipos de entidades incluídas entre os ‘endurants’.

⁵ Esta classificação se apóia na Teoria Geral da Terminologia desenvolvida por Eugene Wüester (WÜESTER, 1981). Se por um lado a noção de conceito em Wüester nos parece inconsistente, no que tange à classificação e especificação das relações conceituais consideramos que ele propõe uma abordagem bastante consistente.

Na biomedicina,

[i]ndivíduos, universais e coleções têm diferentes propriedades temporais. Indivíduos podem ganhar e perder partes (Por exemplo, organismos ganham e perdem células). Universais podem ganhar e perder instâncias (por exemplo, o universal *ser humano* ganha e perde instâncias cada vez que uma pessoa nasce ou morre). Coleções, por sua vez, são identificadas por seus membros e assim não podem ser membros diferentes em tempos diferentes. (BITTNER; DONNELLY; SMITH, 2004, p. 38)

Tal detalhamento visa tratamento mais acurado dos dados. Identificar os tipos de relação em um domínio específico da biomedicina pode não apresentar dificuldade, pois os termos empregados têm significados dos quais participam os pesquisadores da área. Mas Smith reconhece que um sistema de terminologia requer definições buscando assegurar uniformidade no uso dos termos e nas bases de dados biológicas visando, entre outros, compartilhamento de dados e interoperabilidade. Assim, sistemas de terminologia têm estreita ligação com ontologias. Embora a ênfase nesta discussão se limite à área biomédica – em que Smith atua –, os argumentos anteriores podem ser generalizados.

Terminologia parece, então, ser fundamental para os desenvolvedores de ontologias. Um dos movimentos mais atuantes é o do grupo Gene Ontology Consortium (**GO.consortium 2014**), que reúne diversas ontologias, sendo a Gene Ontology bastante popular devido a seu âmbito.

A Gene Ontology compreende um vocabulário controlado e estruturado no domínio genômico e é considerada como uma ferramenta para representar e buscar informação relacionada a genes e produtos relacionados a todas as espécies, destinada a apoiar a anotação dos pesquisadores.

Apesar de listar os termos em ordem alfabética, a GO apresenta grafos para os três aspectos considerados, a saber, componente celular, função molecular e processo biológico. Tais grafos contornam a limitação da ordem alfabética e auxiliam os pesquisadores a localizar os aspectos

e atributos, as relações entre eles e encontrar os termos correspondentes no vocabulário da GO; tal estrutura permite, ainda, localizar falhas ou ausências nas hierarquias.

O projeto Gene Ontology - GO busca descrições (definições) consistentes dos produtos “genes” nas bases de dados que englobam o Gene Ontology Consortium (**GO.consortium, 2014**); definições que têm merecido cuidadosas análises visando a eliminar inconsistências. Parece que um dos motivos para tais inconsistências se deve à falta de orientação segura para elaborar as definições. A orientação para isso propõe que sejam uma ‘descrição textual do que o termo representa, mais as referências à fonte da informação’ – o que não é suficiente para uma descrição acurada. O tipo de definição estabelecido é a definição lógica, no formato *genus proximum/differentia specifica*⁶.

Outro aspecto que motiva tais inconsistências pode ser a limitação às relações tipo-de e parte-de; esta última, principalmente, mais problemática em relação a sua aplicação, pois nem sempre são usadas de maneira consistente. Na GO além de *part-of*, acrescentam-se as relações *has-part*, *regulates* e *regulated by*, que ‘nem sempre [são] suficientes para definições precisas’ (BITTNER; DONNELLY; SMITH, 2004; SMITH; KUMAR, 2004). As inconsistências semânticas observadas nas definições dificultam, desse modo, sua utilização em vários aspectos. É estranho, ainda, verificar que na orientação para uma ‘definição padrão’, definições iniciem com verbos no gerúndio ou no presente.

Smith e Kumar (2004) analisam a estrutura da GO e, para o que nos interessa aqui, fica a necessidade de critérios que devem ser satisfeitos se uma linguagem tiver de ser estruturada de modo que o conteúdo de informação expresso por seu meio possa ser extraído via procedimentos automáticos que possam apoiar ferramentas de raciocínio lógico.

⁶ <http://geneontology.org/page/ontologo-structure>

Nos estudos que desenvolvemos sobre integração de ontologias de domínio na área de bioinformática⁷, pudemos observar a existência de várias inconsistências semânticas nas Ontologias OBO Foundry – The Open Biological and Biomedical Ontologies –, que foi criada como um experimento colaborativo que envolve ontologias de base científica. A GO faz parte deste consórcio de ontologias.

Tais problemas, encontrados nas ontologias que analisamos (CAMPOS, 2007; CAMPOS et al., 2013) e muitos dos quais reportados na literatura (SMITH; WILLIAMS; SCHULZE-KREMER, 2003; OGREN, et al., 2004; SMITH, et al., 2005; KOHLER et al., 2006), apresentam aspectos de naturezas distintas porém relacionados, tais como:

(i) problemas na hierarquia: deficiência na forma como os conceitos são estruturados; omissão de termos importantes na hierarquia;

(ii) problemas nas relações: poucas relações disponíveis para descrição dos relacionamentos entre os termos e para expressar o conhecimento do domínio das ontologias; relações empregadas de forma equivocada; omissão de relações aparentemente óbvias;

(iii) problemas na definição dos termos: informações implícitas inconsistentes com o conhecimento explicitado na ontologia; falta de padronização para criação de definições;

(iv) problemas contextuais: falta de regras para definir se um dado conceito está ou não presente em determinada ontologia, ou seja, falta de descritores para definir que conceitos são expressos nas ontologias; falta de uma padronização mais formal para a construção das ontologias desse consórcio, fazendo com que termos idênticos representem conceitos diferentes

⁷ Estes estudos foram desenvolvidos entre os anos de 2005 a 2008 através da pesquisa “Integração de ontologias: O domínio da Bioinformática e a problemática da compatibilização terminológica” apoiada com bolsa de Produtividade em Pesquisa pelo CNPq e também no âmbito do “Grupo de Pesquisa Ontologia e Taxonomia: aspectos teóricos e metodológicos”. Os resultados desse estudo podem ser verificados através de nossa produção científica.(CAMPOS et al., 2013; CAMPOS, 2014, entre outros)

e termos diferentes representem o mesmo conceito; falta de documentação das ontologias, mostrando seu escopo, objetivo e temática; falta de alinhamento entre as ontologias que algumas vezes possuem informações repetidas em relação a outras ontologias, não se relacionando explicitamente.

No que tange às definições, é importante tornar explícitas as relações que ocorrem em seu interior, que são mais ricas do que a estrutura sistemática atualmente apresentada. É necessário um procedimento metodológico que não encontramos na literatura relacionada a ontologias, mas consideramos que a ciência da informação possui a resposta, como veremos adiante.

Como exemplo pode-se citar a definição associada ao termo **Mitochondria** na Ontologia GO: “*A semiautonomous, self replicating organelle that occurs in varying numbers, shapes, and sizes in the cytoplasm of virtually all eukaryotic cells. It is notably the site of tissue respiration*”. Nessa definição, é possível extrair três relações, as quais não foram encontradas de forma explícita na ontologia: **is_a semiautonomous, self replicating organelle**; **occurs in varying numbers, shapes, and sizes in the cytoplasm of virtually all eukaryotic cells**; **the site of tissue respiration**, que eventualmente podem vir a complementar ou servir para tornar consistentes as relações já expressas explicitamente na estrutura da ontologia.

O estudo e desenvolvimento de relações padronizadas tornam-se, então, objetivos a serem alcançados para a interoperabilidade. Nesse sentido, atualmente, a OBO vem desenvolvendo uma série de princípios para definições acuradas. O sítio OBO-relations contém uma ‘coleção de relações que pretende principalmente favorecer a padronização entre as ontologias da OBO Foundry e da biblioteca OBO’⁸, de caráter geral e, portanto, aplicáveis a diversos domínios.

⁸ OBO-Relations são aplicáveis a domínios diversos (domain neutral relations) e incluem: part of, has part; realized in, realizes; occurs; contain process; inheres in; bearer of; participates in; has participant; is role of, has quality; has role; derives from, derives into; location of, contained in, contains; located in; boundary of; member of, has member. (<https://code.google.com/p/obo-relations>)

Ontobee é outra iniciativa: trata-se de um servidor *linked-data* projetado para ontologias. ‘Visa a facilitar compartilhamento de dados, visualização, buscas, integração e análise. Oferece 382 ‘termos’ do tipo Object-Property que são empregados em 150 ontologias’⁹ (XIANG, et al., 2011).

A preocupação com uma definição capaz de apoiar procedimentos automáticos é então um objetivo a ser alcançado pelos desenvolvedores de ontologias.

Voltamos, então, à Teoria do Conceito de Dahlberg que através do método analítico leva à descrição do conceito. Dahlberg, ao buscar um método científico para produzir o entendimento de conceitos, acaba por possibilitar uma descrição acurada e precisa necessária às ontologias.

Dahlberg oferece um exemplo de um termo de interesse geral, não necessariamente científico, o qual nos dá uma amostra de como analisar o referente. Seja um sino:

é um recipiente côncavo; normalmente feito de metal, às vezes de chifre, de madeira, de vidro ou de barro; quando oscila, é percutido por um badalo suspenso - em seu interior ou por um martelo externo para produzir um som intenso; é utilizado como instrumento sonoro para as vibrações de um material sólido ressonante; tem uma forma característica que depende do uso que se faz no ambiente cultural e do material com o qual foi construído; possui paredes que podem ser retas, convexas, côncavas, hemisféricas, de forma cilíndrica ou de tulipa; tem uma seção transversa que pode ter uma configuração circular, quadrada, tetragonal, elíptica ou poliédrica; tem uma ampla distribuição geográfica; em geral se distingue por um função cultural bem definida; é usado como sinal de chamamento (por exemplo, às funções religiosas) e com finalidades musicais. (DAHLBERG apud GNOLI; MARINO; ROSATI, 2006).

É possível observar que as afirmativas englobam os muitos aspectos de observação do referente. Dentro de um propósito, ou alvo, é possível selecionar os aspectos necessários.

⁹ www.ontobee.org/index.ph

Vejamos, agora, um exemplo de uma descrição acurada e precisa de um termo relacionado à doença mental, proposta na ontologia:

MANIFESTATION OF A MENTAL DISEASE

= **def.** a **BODILY FEATURE** of an **ORGANISM** that is (a) deviation from clinical normality that is the *realization of* a **MENTAL DISEASE** and is (b) observable (CEUSTERS; SMITH, 2010).

Em negrito estão as características, que também são conceitos, e, portanto, objeto de descrição - e as relações em itálico. Podemos observar que as características estão em maiúsculas – o que significa que também estão definidas.

A definição lógica, segundo a classificação de Dahlberg (1978) para os diversos tipos de definição, é a que mais se ajusta às questões discutidas aqui - tal tipo tem caráter estruturante. Seguindo o princípio Aristotélico do *genus proximum et differentia specifica* a relação hierárquica fica evidenciada na terminologia.

Dahlberg preconiza, ainda – no contexto de uma estrutura classificatória -, a necessidade de se trabalhar numa lista de verbos, deixando evidente que as preocupações dos desenvolvedores de ontologias devem estar presentes também nas pesquisas da ciência da informação, em especial no desenvolvimento dos instrumentos semânticos como taxonomias e sistema facetados de terminologia (DAHLBERG, 1978).

Além disso, a proposta de Dahlberg de adotar um método científico para identificação de um conceito vai ao encontro daquilo que postulam os desenvolvedores de ontologias no sentido de obter uma descrição precisa dos termos.

CONCLUSÕES

O propósito dos estudos que apresentamos tem relação com a nossa postura diante das pesquisas que estamos realizando durante esses anos sobre a Representação de Domínios de Conhecimento no âmbito dos estudos informacionais. Consideramos que não existe modelo certo ou errado, e sim o que atende a determinados propósitos, e os propósitos definidos devem estar explícitos, assim como os pressupostos em que estamos nos apoiando. Esta é uma das etapas que deve ser explicitada na construção de modelos de domínios: evidenciarmos os pressupostos teóricos em que o modelo está apoiado, para que possamos elaborá-los a partir de decisões consistentes e explícitas. Assim, consideramos a importância de se ter um conhecimento ampliado das diversas teorias de representação que são apresentadas em áreas que se preocupam com a Representação de Domínios de Conhecimento. Aqui neste artigo, apresentamos uma síntese de duas delas.

Assim, consideramos que a defesa de uma ou outra teoria é sempre circunstancial, e pode ser considerada uma posição que deverá atender a um propósito específico, que nesta medida é sempre contextual.

REFERÊNCIAS

- ARP, R.; SMITH, B. Function, role, and disposition in basic formal ontology. *Nature Proceedings*..., Jun. 2008.
- BRANQUINHO, J. et al. *Enciclopédia de termos lógico-filosóficos*. São Paulo: Martins Fonte, 2006.
- BITTNER, T.; DONNELLY, M.; SMITH, B. Individuals, universals, collections: on the foundational relations of ontology. In: VARZI, A.; VIEU, L. (Ed). *Formal ontology in information systems*. Amsterdam: IOS Press, p. 37-48, 2004. (Proceedings of the Third International Conference)
- CAMPOS, M.L.A. Integração de ontologias: o domínio da bioinformática. *RECIIS - Revista Eletrônica de Comunicação, Informação & Inovação em Saúde*, v. 1, p. 117-121, 2007. Disponível em: <<http://www.reciis.icict.fiocruz.br/index.php/reciis/article/viewFile/32/44>>. Acesso em: 19 set. 2014.
- _____. et al. Information Sciences methodological aspects applied to ontology reuse tools: a study based on genomic annotations in the domain of trypanosomatides. *Knowledge Organization*, v. 40, p. 50-61, 2013.
- CEUSTERS, W.; SMITH, B. Foundations for a realist ontology of mental disease. *Journal of Biomedical Semantics*, v. 1, n. 10, 2010.
- DAHLBERG, I. A Referent-oriented analytical concept theory of interconcept. *International Classification*, Frankfurt, v.5, n.3, p.142-150, 1978.
- _____. *Ontical structures and universal classification*. Bangalore: Sarada Ranganathan Endowment for Library Science, 1978a.
- GNOLI, C.; MARINO, V.; ROSATI, L. *Organizzare la conoscenza: dalle biblioteche all'architettura dell'informazione per il web*. Milano: Tecniche nuove, 2006.
- GO.consortium**. *Gene Ontology Consortium*. Disponível em: <http://geneontology.org/>. Acesso em: 19 set. 2014.
- GRENON, P.; SMITH, B. Foundations of an ontology of philosophy. *Synthese*, v. 182, n.2, p.185-204, 2011.
- KLEIN, G.O.; SMITH, B. Concept Systems and Ontologies: Recommendations for basic terminology. Recommendations based on discussions between realist philosophers and ISO/CEN experts concerning the standards addressing “concepts” and related terms. 2005. Disponível em: <<http://ontology.buffalo.edu/concepts/ConceptsandOntologies.pdf>>. Acesso em 19 set. 2014.
- KOHLER, J. et al. Quality control for terms and definitions in ontologies and taxonomies. *BMC Bioinformatics*, v.7, n. 212, p. 1-12, 2006.
- MORA, J.F. *Dicionário de filosofia*. São Paulo: Martins Fonte, 2001.
- OGREN, P.V. et al. The compositional structure of gene ontology terms. *Pacific Symposium on Biocomputing*, v. 9, p. 214-225, 2004.
- RIGGS, F.W. Information science and social science: the need for onomantics. *International Forum for Information and Documentation*, v. 14, n.1, p. 12-21, 1989a.
- _____. Terminology and lexicography: their complementarity. *International Journal of Lexicography*, v.2 n.2 p. 90-110, 1989b.
- _____. Onomantics and terminology. *Knowledge Organization*, v. 23, p. 25-23, 1996.
- SMITH, B. Beyond concepts: ontology as reality representation. In: FOIS 2004: International Conference on Formal Ontology and Information Systems, 2004, Turing. *Proceedings*...
- _____.; KUMAR, A. On controlled vocabularies in bioinformatics: a case study in gene ontology. *Drug discovery today: BIOSILICO*, v.2, n.6, p. 246-251, 2004. Disponível em: <ontology.buffalo.edu/bio/Compositionality_in_GO.pdf>. Acesso em: 19 set. 2014.
- _____.; CEUSTERS, W.; TEMMERMAN, R. Wüsteria. *Studies in Health Technology and Informatic*, v.116, p. 647-652, 2005.
- _____.; WILLIAMS, J.; SCHULZE-KREMER, S. The ontology of the gene ontology. In: *AMIA Symposium*, 2003. p. 609-613.
- _____. et al. Relations in biomedical ontologies. *Genome Biology*, v. 6, n. 5, p.1-15, 2005.
- XIANG, Z. et al. Ontobee: a linked data server and browser for ontology terms. In: INTERNATIONAL CONFERENCE ON BIOMEDICAL ONTOLOGIES (ICBO), 2., 2011, Buffalo, NY. *Proceedings*... Buffalo, 2011. p. 279-281. Disponível em:<<http://ceur-ws.org/Vol-833/paper48.pdf>>. Acesso em: 19 set. 2014.
- WÜESTER, E. L'Étude scientifique générale de la terminologie, zone frontalière entre la linguistique, la logique, l'ontologie, l'informatique et les sciences des choses. In: RONDEAU, G.; FELBER, F. (Org.). *Textes choisis de terminologie: I. Fondements théoriques de la terminologie*. Québec: GIRSTERM, 1981. p. 57-114.