

Estudo exploratório sobre ontologias aplicadas a modelos de sistemas de informação: perspectivas de pesquisa em Ciência da Informação

An exploratory study of ontologies applied to information system models: research perspectives in Information Science

Maurício Barcellos Almeida

Universidade Federal de Minas Gerais, Escola de Ciência da Informação, Departamento de Teoria e Gestão da Informação
mba@eci.ufmg.br, + 55 (31) 3409-6107, fax: +55 (31) 3409-5200

Renato Rocha Souza

Universidade Federal de Minas Gerais, Escola de Ciência da Informação, Departamento de Organização e Tratamento da Informação
rsouza@eci.ufmg.br, + 55 (31) 3409-5232, fax: +55 (31) 3409-5200

Kátia Cardoso Coelho

Universidade Federal de Minas Gerais, Escola de Ciência da Informação, Núcleo de Informação Gerencial e Tecnológica
katiaccoelho@gmail.com, + 55 (31) 3280-7485, fax: +55 (31) 3281-3242

Palavras-chave: ontologias, modelagem, modelos conceituais, Ciência da Informação

Resumo

Ontologias tem se tornado um assunto de interesse cada vez maior em diversas áreas. Além de suas origens filosóficas, tem sido amplamente utilizadas em campos de pesquisa relacionados ao desenvolvimento de sistemas. No âmbito dos sistemas de informação automatizados, uma ontologia tem duas aplicações principais: pode ser usada como um componente do sistema, ou como uma referência, um padrão de comparação para avaliar a representatividade do modelo subjacente ao sistema. O presente artigo explora o segundo tipo de aplicação, discutindo a utilidade das ontologias em processos de modelagem conceitual. Após contextualizar a pesquisa e alguns problemas da modelagem conceitual de sistemas de informação, apresentam-se ontologias utilizadas como referência. Conclui-se sobre a relevância da formação em Ciência da Informação para abordar questões de modelagem deficiente e de falta de integração entre sistemas. Espera-se contribuir para o diálogo entre diferentes áreas, bem como para sugerir novas perspectivas de atuação para profissionais de Ciência da Informação.

Palavras-chave: ontologias, Ciência da Informação, sistemas de informação, modelos conceituais

Abstract

Ontologies have become the subject of ever increasing interest in several areas. Beyond its philosophical roots, it has been used broadly in research fields related to systems development. Within the scope of automated information systems, the use of ontology has two main goals: it can be used as a component of a system or as a reference. This reference is a kind of comparison standard employed to evaluate the capacity of the model underpinning the system to represent the world. The present article explores the latter application, discussing the usefulness of ontologies in conceptual modeling processes. After the contextualization of the research and of some problems related to conceptual modeling of information systems, we present ontologies used as references. We conclude by highlighting the importance of the Information Science disciplines in addressing issues related to deficient modeling and to lack of systems integration. We hope to contribute to the interchange between different areas, as well as suggest new perspectives for the work of Information Science professionals.

Palavras-chave: *ontologies, Information Science, Informations Systems, conceptual models*

1. Introdução

A despeito de sua origem na Filosofia, o estudo das ontologias tem se caracterizado pela interdisciplinaridade, envolvendo áreas como Ciência da Computação, Linguística, Ciência da Informação, dentre outras. Na década de 60, o termo *ontologia* passou a ser utilizado na pesquisa em Representação do Conhecimento e Inteligência Artificial. Nesse contexto, o termo

diz respeito a uma linguagem formal, que tem utilizações específicas em arquiteturas de sistemas inteligentes. Nos anos 80, surgiram pesquisas em Engenharia de *Software* aplicando ontologias à fase de modelagem conceitual no desenvolvimento de sistemas de informação (SI).

A modelagem conceitual tem por finalidade descrever parte da realidade e representar processos de interesse em um contexto social para a construção de um SI. A dificuldade em obter a inter-operabilidade entre sistemas tem origem na modelagem deficiente, conduzida "caso a caso" e sem correspondência com o mundo real (Smith & Welty, 2001). Uma alternativa para abordar esse problema é confrontar a linguagem de representação do modelo a uma ontologia, a qual funciona como um padrão de comparação. Para tal, consideram-se ontologias *bem fundamentadas*, ou seja, aquelas baseadas em ontologias filosóficas e em ontologias genéricas. Por sua fundamentação em princípios filosóficos, essas ontologias podem colaborar para solução de problemas de integração entre sistemas, ao facilitar o processo de comunicação humana.

O desenvolvimento de SI automatizados é uma atividade constituída por atividades distintas, as quais podem ser divididas em dois grandes grupos: i) atividades relacionadas a criação dos processos algorítmicos, típicos da etapa de implementação do sistema em computadores; ii) atividades relacionadas aos processos de comunicação e de abstração humana, típicos da etapa da criação de modelos conceituais. Advoga-se que, no âmbito do segundo tipo de atividade, existem oportunidades para a pesquisa em Ciência da Informação (CI), uma vez que a formação dos profissionais da CI pode contribuir para solucionar problemas na modelagem de SI.

Inserindo-se nesse contexto, o presente artigo explora o uso de ontologias como referência para a criação de modelos, com a finalidade de obter melhorias em SI. Para isso, contextualiza-se o uso das ontologias em SI, descrevem-se as principais ontologias que tem sido utilizadas em modelagem conceitual e apresentam-se as iniciativas de pesquisa relacionadas. Espera-se contribuir para o diálogo e o intercâmbio entre diferentes áreas de pesquisa.

O restante do presente artigo está organizado conforme segue. A Seção 2 apresenta um breve histórico dos modelos para SI, inclusive modelos baseados em ontologias. A Seção 3 discorre sobre ontologias utilizadas como referência para criação de modelos, descrevendo ontologias filosóficas e ontologias de alto-nível. A Seção 4 apresenta iniciativas de aplicação de ontologias em modelagem conceitual. A Seção 5 discute perspectivas de pesquisa relacionadas no âmbito da CI e sugere contribuições. A Seção 6 conclui sobre os benefícios da participação de profissionais da CI no desenvolvimento de SI.

2. Modelos para sistemas de informação

Modelos são produzidos para que a compreensão humana possa apreender o mundo em partes, visto que não consegue abrangê-lo em sua totalidade. Segundo Frigg e Hartmann (2006), uma das formas de classificar modelos é considerar suas funções de representação. Nesse sentido, os modelos podem ser modelos de fenômenos, modelos teóricos ou modelos de dados. Os *modelos de dados* proliferam nas organizações como forma de representar o que deve ser codificado em SI.

Ao longo do desenvolvimento de SI, a etapa em que são criados modelos, com vistas ao entendimento humano, é conhecida como *modelagem conceitual*. Os *modelos conceituais* são obtidos a partir de abstrações de aspectos da realidade, na perspectiva de uma pessoa ou grupo de pessoas. As abstrações são formas de especificar as entidades e as relações entre as entidades em um domínio do conhecimento de interesse para o sistema em construção.

O restante da presente seção discorre sobre modelos para SI, descrevendo a evolução dos modelos de dados até os modelos conceituais e, finalmente, os modelos baseados em ontologias.

2.1. Modelos de dados e modelos conceituais

As primeiras iniciativas para especificação de modelos de dados datam de final dos anos 50 (Young & Kent, 1958; Bosak et al, 1962). Nesse contexto, a criação de modelos era orientada para estruturas de dados computacionais. Nos anos 60, a pesquisa em bancos de dados deu origem a três tipos principais de modelos de dados: o modelo hierárquico, o modelo em rede e o modelo relacional. Esses modelos são conhecidos como *modelos lógicos*, porque não fazem a referência a aspectos físicos (codificação). Entretanto, os modelos lógicos têm concepções que limitam sua utilização em modelagem conceitual. Por exemplo, no modelo relacional (Codd, 1970), o conceito *relação* é usado para representar tanto *entidades* quanto *relacionamentos entre entidades* (Peckhman & Maryanski, 1988). Esse fato gera dificuldades de entendimento e induz a erros de modelagem.

Nos anos 70 surgiram os primeiros *modelos semânticos* utilizados em modelagem conceitual em sistemas de gerenciamento de bancos de dados. Destacam-se o modelo de dados semânticos (Abrial, 1974), a arquitetura de três esquemas (Jardine, 1976), o modelo Entidade Relacionamento (ER) (Chen, 1976), dentre outros. A principal característica dos modelos semânticos, em relação aos modelos lógicos, é a facilidade de entendimento. O modelo ER, por exemplo, elimina a sobrecarga do constructo ¹ *relação* existente no modelo relacional e fornece termos adicionais para uso como primitivas de modelagem. A modelagem conceitual teve origem nos modelos de dados semânticos para bancos de dados, os quais correspondem a apenas um componente de SI. Entretanto, a iniciativa foi estendida as linguagens de modelagem conceitual para SI.

Nos anos 90, surgiram as propostas de modelagem *orientadas a objeto*, as quais apresentam características adicionais em relação aos modelos de dados, mas também similaridades. Por exemplo, os constructos: *objetos* equivalem *entidades*, *atributos* equivalem *propriedades*, *relacionamentos* equivalem *associações*, dentre outros (Milton, 2000). A *Unified Modeling Language* (UML) correspondeu a uma tentativa de uniformizar as notações da orientação a objeto, a qual reuniu outras iniciativas (Booch, 1993; Rumbaugh et al, 1991; Jacobson et al, 1992).

Ao longo de sua evolução, a atividade de modelagem conceitual tem sido motivada pela busca de melhores formas para representar a realidade em SI. De fato, a modelagem conceitual é “a atividade de formalmente descrever alguns aspectos do mundo físico e social ao nosso redor para propósitos de compreensão e de comunicação” ² (Mylopoulos, 1992, p. 3). Entretanto, os modelos semânticos dispõem de um conjunto limitado de constructos para criar um modelo conceitual. O modelo ER, por exemplo, pressupõe que a realidade de interesse para o sistema pode ser representadas por dois constructos apenas (*entidade e relacionamento*).

Segundo Smith e Welty (2001), a inconsistência nas atividades de modelagem das primeiras iniciativas de modelagem conceitual é responsável pelas dificuldades atuais de interoperabilidade em SI. Uma alternativa para solucionar esse tipo de problema são os modelos baseados em ontologias. Segundo os autores, “[...] só o uso de uma ontologia de referência robusta, comum - uma taxonomia compartilhada de entidades - pode proporcionar vantagens significativas sobre os métodos *ad-hoc* e baseados na análise individual até então utilizados” ³ (Smith & Welty, 2001, p. 4).

Segundo Guarino (1998), existe a necessidade de criação de uma conceitualização ⁴ comum. Segundo o autor, a possibilidade de integração entre sistemas só é possível apenas se as linguagens subjacentes aos modelos possuem conceitualizações que se sobrepõem em

¹ Um constructo de uma linguagem de modelagem descreve um conjunto de objetos que tem propriedades comuns (Wand & Weber, 1993).

² [...] the activity of formally describing some aspects of the physical and social world around us for purposes of understanding and communication.

³ [...] the provision, once and for all, of a common, robust reference ontology – a shared taxonomy of entities – might provide significant advantages over the ad hoc, case-by-case methods previously used.

⁴ Uma conceitualização é uma visão abstrata e simplificada do mundo que se deseja representar.

algum nível. Muitas vezes tal superposição não existe entre as conceitualizações, como pode parecer em uma avaliação que leva em conta apenas as linguagens. As linguagens de sobrepõem, não as conceitualizações. Em última instância, apenas uma conceitualização compartilhada pode proporcionar condições para interoperabilidade entre SI. Tal conceitualização é produto da comunicação humana e do consenso entre um grupo de especialistas.

2.2. Modelos baseados em ontologias

Ontologias têm sido estudadas na Inteligência Artificial desde os anos 70. Nos anos 90, a pesquisa em Web Semântica aumentou a demanda por ontologias para vários tipos de aplicações. Vickery (1997) relata pesquisa conduzida em uma base de dados⁵, a qual retornou mais de 500 ocorrências do termo. A pesquisa em ontologias têm se caracterizado por interdisciplinaridade, sendo conduzida por autores da Filosofia (Smith, 2003), da Computação (Genesereth & Nilsson, 1987; Gruber, 1993; Guarino, 1998; Sowa, 2000; Wand & Weber, 1990), e na Ciência da Informação (Vickery, 1997; Gilchrist, 2003; Særguel, 1997). Apresentar questões teóricas sobre ontologias está além dos objetivos do presente artigo. Ainda assim, no restante dessa seção, apresentam-se considerações relevantes para tais objetivos.

O estudo das ontologias se caracteriza pela coexistência de abordagens variadas, a qual confirmar seu caráter interdisciplinar. Guarino e Giaretta (1995) citam diversas interpretações para o termo ontologia encontradas na literatura: i) uma disciplina filosófica; ii) um sistema conceitual informal; iii) um estudo de semântica formal; iv) uma especificação de uma conceitualização; v) uma representação de um sistema conceitual via teoria lógica; vi) um vocabulário usado por uma teórica lógica; vii) uma especificação (meta-nível) de uma teoria lógica.

Do ponto de vista filosófico só pode existir uma ontologia. Para lidar com a pluralidade de usos do termo, Smith (1998) distingue dois tipos de ontologias: a *ontologia real (R-ontology)*, a abordagem filosófica, que diz respeito como o universo é organizado; e a *ontologia epistemológica (E-ontology)*, relacionada à tarefa de conceitualizar um domínio. A *E-ontology* atende a necessidade de expressar a ontologia como um artefato de software, presente na abordagem da Engenharia de *Software* e da Representação do Conhecimento.

Segundo Guarino (1998), no mundo dos SI, uma ontologia descreve o significado dos símbolos utilizados no sistema, de acordo com uma conceitualização de mundo. O autor classifica as ontologias em duas dimensões principais: a *dimensão temporal*, que corresponde ao uso da ontologia em SI em tempo de desenvolvimento; e a *dimensão estrutural*, que diz respeito ao uso da ontologia como um componente de um aplicativo. Fonseca (2007) distingue ontologias *de* SI e ontologias *para* SI: no primeiro caso, a ontologia é utilizada para modelagem conceitual; no segundo caso, a ontologia é um componente dos SI que descreve o vocabulário de um domínio com o propósito criar esquemas conceituais. O primeiro caso corresponde à visão de Guarino (1998) e é relevante para os propósitos do presente artigo. Exemplos de iniciativas relacionadas são descritas na seção 4.

Uma ontologia *para* SI deve ser avaliada quanto a sua capacidade de desempenhar a função para a qual foi projetada. Em geral, tal avaliação está relacionada aos mecanismos para interação com a ontologia, ao formalismo de representação do conhecimento considerado e a correspondência da documentação ao processo. No caso das ontologias *de* SI, verifica-se a capacidade de representação das linguagens de modelagem, através da comparação entre constructos disponíveis na gramática da linguagem, e constructos da ontologia de referência. Dessa forma é possível avaliar a linguagem de modelagem, o modelo subjacente ao SI, e não apenas um componente do SI, como ocorre na avaliação de ontologias para SI. Esse tipo de avaliação é essencial para pretensões de inter-operabilidade entre sistemas.

O esquema da FIG. 1 apresenta o modelo e o metamodelo da linguagem de modelagem UML, além de um nível adicional que corresponde a ontologia. Os constructos do segundo nível

⁵ Dialog. Retrieved January 20, 2003, from <http://www.dialog.com/>

(metamodelo) definem os objetos que podem ser representados pelo primeiro nível (modelo). A ontologia, bem fundamentada filosoficamente, é a referência para avaliar em que extensão o metamodelo e o modelo são adequados para representar a realidade.

No presente trabalho, denominam-se *ontologias de referência*⁶, aquelas ontologias fundamentadas na Filosofia, as suas variações, além das ontologias de alto nível, todas utilizadas especificamente para fins de modelagem conceitual. Exemplos de ontologias filosóficas são a ontologia de *Bunge* e a ontologia de *Chisholm*. A ontologia de *Bunge-Wand-Weber* (BWW) é um exemplo de variação de ontologia filosófica. Ontologias de alto nível contêm conceitos genéricos, passíveis de utilização em domínios diversos. Exemplos de ontologias de alto nível são: i) *Basic Formal Ontology* (BFO) (Grenon, Smith & Goldberg, 2004); ii) *Descriptive Ontology for Linguistics and Cognitive Engineering* (DOLCE) (Masolo et al, 2003); iii) *General Formal Ontology* (GFO) (Herre et al, 2006); iv) *Unified Foundational Ontology* (UFO) (Guizzardi & Wagner, 2004).

O restante da presente seção descreve as principais ontologias de referência disponíveis, além de considerações sobre as respectivas ontologias de origem (filosóficas e de alto-nível).

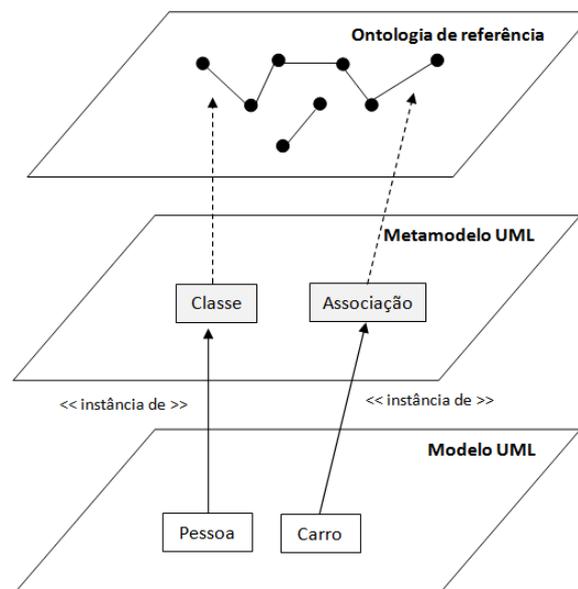


Figura 1: representação gráfica de modelo, metamodelo e ontologia de referência
Fonte: adaptado de OMG (2003, p. 29)

3. Ontologias de referência para modelagem conceitual

Nessa seção, as ontologias de referência descritas são: a ontologia de *Chisholm*, a ontologia de *Bunge*, a BWW, a DOLCE e a UFO. Existem outras ontologias de alto-nível, como a *KR Ontology* (Sowa, 2000), a SUMO (IEEE Standard⁷), as quais, apesar de importantes, não são descritas pois não tem sido associadas a modelagem conceitual de SI.

A ontologia de *Chisholm*⁸ é uma ontologia filosófica de senso comum baseada em uma forma extrema de realismo. O *realismo* corresponde a presunção de que a realidade do mundo existe independente da vontade humana (SMITH, 1995). A ontologia de Chisholm (1996) é organizada em categorias, onde são agrupadas as coisas do mundo. Milton e Kazmierczak (2004), descrevem a semântica das categorias da ontologia de Chisholm, enfatizando aquelas relevantes para a pesquisa em modelagem (FIG. 2).

⁶ Existem variações com significado similar: ontologias de base, ontologias fundamentais, dentre outros.

⁷ Institute of Electrical and Electronics Engineers. Retrieved January 20, 2005, from <http://www.ieee.org>

⁸ Roderick Chisholm, filósofo norte-americano (1916 - 1999).

Conceitos	Indivíduo	Indivíduos são objetos discerníveis e transientes.
	Atributo	Atributos se tornam aparentes através de indivíduos.
	Classes	As classes são representados através de atributos, os quais determinam as condições de associação.
	Relação	Relações são pares ordenados de atributos, os quais conectam indivíduos de forma unidirecional.
	Relações temporais	O tempo não é expresso como uma sequência de eventos, mas através de relações temporais que ordenam os estados.
	Estados	Indivíduos exemplificam atributos, e tais exemplos são estados.

Figura 2: Principais categorias da ontologia de *Chisholm*

A ontologia de *Bunge*⁹ consiste de uma ontologia filosófica na qual o mundo material existe independentemente de conhecimento que as pessoas possam ter sobre ele. Considera a possibilidade do conhecimento humano objetivo, desde que baseado no método científico através do qual se pode obter a verdade. Trata-se portanto, de uma abordagem realista, em que a ontologia é descrita em abstrações formalizadas por lógica. As principais premissas da ontologia de Bunge (1977) são apresentadas na FIG. 3:

Coisas	O mundo é feito de <i>coisas</i> que possuem <i>propriedades</i> .
	Cada <i>coisa</i> no mundo possui uma ou mais <i>propriedades</i> .
Propriedades	<i>Propriedades</i> não podem ter <i>propriedades</i> .
	<i>Propriedades</i> não podem existir por si mesmas, e devem estar ligadas a alguma <i>coisa</i> .
	As <i>propriedades</i> podem ser <i>intrínsecas</i> , as quais dependem apenas de uma coisa; e <i>mútuas</i> , que dependem de duas ou mais <i>coisas</i> .
Interações entre coisas e propriedades	Dois coisas interagem quando a história de uma delas, manifestada como uma sequência de estados, poderia ter sido diferente se a outra não existisse
	A existência de <i>propriedades mútuas</i> entre duas coisas pode indicar que elas interagem.
	Propriedades mútuas que manifestam interações entre duas coisas são chamadas <i>propriedades mútuas compulsórias</i> .
	Coisas podem se associar e formar outra coisa: uma <i>coisa composta</i> combina pelo menos duas outras coisas; caso contrário ela é uma <i>coisa simples</i> .
	Cada <i>coisa composta</i> possui propriedades emergentes, ou seja, propriedades ausentes nos componentes da coisa composta.

Fig. 3: Principais categorias da ontologia de *Bunge*

A pesquisa pioneira sobre ontologias de SI é, na verdade, a uma variação da teoria de *Bunge* denominada BWW (Wand & Weber, 1990, 1993; Wand, Storey & Weber, 1999). Trata-se de uma proposta que estrutura facetas genéricas do mundo real, com o objetivo de fornecer bases teóricas para avaliar práticas de modelagem conceitual e investigar o poder de representação das linguagens.

Wand e Weber (1993) descrevem critérios para verificar o nível de *homomorfismo*¹⁰ entre os constructos da linguagem de modelagem (L) e os constructos ontológicos (O) definidos pela BWW. Os critérios são denominados *completeza ontológica* e a *transparência ontológica*. No primeiro caso, uma linguagem de modelagem é considerada *ontologicamente completa*

⁹ *Mario Bunge*, filósofo e físico argentino (1919 - ?).

¹⁰ A despeito da definição matemática, no contexto do trabalho de Wand e Weber, o homomorfismo garante que a interpretação de um modelo por pessoas corresponde a abstração que os diagramas do modelo pretendem representar.

quando fornece pelo menos um constructo para cada constructo ontológico. O segundo caso, diz respeito a interpretação do significado de cada constructo da linguagem. Uma linguagem é *ontologicamente transparente* se não apresenta: i) *sobrecarga*, ou seja, um constructo (L) corresponde a dois (O); ii) *redundância*, em que dois constructos (L) correspondem a um (O); iii) *excesso*, em que nem todos os constructos (L) tem (O) correspondentes.

A ontologia DOLCE (Masolo et al, 2003) tem por objetivo descrever categorias ontológicas subjacentes a linguagem natural e ao senso comum humano. Integra esforço de pesquisa do LADSEB-CNR¹¹, para desenvolvimento de linguagens de representação de ontologias, denominado *WonderWeb Foundational Ontologies Library*. Faz parte do projeto a especificação de uma ontologia baseada em Lógica de 1ª Ordem e Lógica Modal, a DOLCE. Outra contribuição da equipe de pesquisa da instituição é metodologia *Ontoclean*.

A metodologia *OntoClean* (Guarino & Welty, 2002) fornece princípios ontológicos formais baseados em ontologias da metafísica. É utilizada para detectar modelagem inconsistente, propondo correções através meta-propriedades que impõe restrições a relações taxonômicas. Gangemi et al (2001) descrevem as quatro noções formais genéricas que fundamentam tais metapropriedades: *rigidez*, *identidade*, *dependência* e *unidade*. Essas noções permitem entender a natureza das relações e encontrar problemas estruturais. Os passos para a correção de erros na taxonomia são: i) selecionar propriedades; ii) avaliar a taxonomia a partir das metapropriedades; iii) complementar a taxonomia com conceitos e relações, de acordo com resultados da avaliação.

A GFO (Herre et al, 2006) é uma ontologia de alto nível utilizada em aplicações de Biomedicina no âmbito do *Onto-Med Research Group*¹². Além dessa aplicação principal, a GFO é também utilizada como fundamento ontológico para modelagem conceitual. A ontologia é estruturada em três níveis: *abstrato*, *núcleo* e *básico*. Esses níveis incluem objetos e processos, além de módulos especializados sobre medicina, economia e sociologia.

A UFO (Guizzardi & Wagner, 2004) é resultado da combinação da GFO, da DOLCE e de princípios da metodologia *OntoClean*. Como trata-se de uma síntese de outras ontologias, a UFO oferece maior número de constructos para avaliação de linguagens de modelagem conceitual. Consiste de três camadas principais: a UFO-A, que define o núcleo ontológico; a UFO-B, que é composta por termos relacionados a processos; e a UFO-C, que define termos relacionados às esferas sociais e aspectos lingüísticos.

Apesar das diversas abordagens, não há uma ontologia padrão, consensualmente aceita para a tarefa de modelagem conceitual. Além disso, existem críticas a forma com que as abordagens são conduzidas. Wyssusek e Klaus (2005) por exemplo, descrevem problemas na adaptação da ontologia de Bunge para a BWB. Segundo os autores, a ontologia filosófica foi tratada como teoria formal e não como teoria ontológica, de forma que a BWB não considera o conceito de compromisso ontológico, fundamental em Bunge. Além disso, o formalismo e a terminologia da ontologia filosófica foram transportados para o domínio da modelagem. Assim, obtém-se uma interpretação do modelo em termos de uma ontologia científica, quando se deveria obter o contrário.

4. Exemplos de ontologias para modelagem

Conforme mencionado anteriormente (seção 2.2), existem diversas iniciativas que utilizam ontologias de referência para analisar linguagens de modelagem de SI. O restante da presente seção descreve propostas que adotam essa abordagem a apresentando como exemplo os trabalhos de Evermann (2003), de Fettke e Loss (2005), de Shanks, Tansley e Weber (2003), de Opdahl e Henderson-Sellers (2004), de Holten, Dreiling e Becker (2004), de Milton (2000) e de Guizzardi e Wagner (2004). A revisão de literatura não é exaustiva e seu objetivo é exemplificar a pesquisa sobre o tema.

¹¹ *Institute of Systems Science and Biomedical Engineering, Consiglio Nazionale delle Ricerche, Itália.*

¹² *Onto-Med.* Retrieved February 20, 2009, from <http://www.onto-med.de/>

Evermann (2003) propõem o mapeamento entre constructos da linguagem UML e constructos ontológicos. A ontologia de Bunge é utilizada para a identificação de elementos básicos que devem ser confrontados com constructos da linguagem UML. Desse processo, originam-se regras que restringem o uso dos constructos da linguagem de acordo com a representação da ontologia. A análise ontológica consiste das seguintes fases: i) identificação de conceitos ontológicos e de constructos específicos da UML; ii) mapeamento de conceitos ontológicos para constructos da linguagem; iii) identificação de suposições ontológicas, relacionamentos e restrições entre os conceitos; iv) mapeamento entre relacionamentos ontológicos.

Fettke e Loss (2005) utilizam a BWW como ontologia de referência para avaliação de modelos de SI relacionados a controle e planejamento em indústrias. Apresentam um estudo de caso empírico com o modelo *Architecture of Integrated Information Systems* (Scheer, 2000). Através da avaliação, a qual os autores denominam *normalização ontológica*, obtém-se uma visão unificada dos fatos representados pelo modelo, reduzindo inconsistências. As etapas do processo de normalização ontológica são: i) mapeamento entre o modelo e a ontologia; ii) identificação das deficiências do modelo, de acordo com a BWW (incompleteza, redundância, excesso e sobrecarga); iii) avaliação dos resultados e correção do modelo.

Shanks, Tansley e Weber (2003) relatam estudos empíricos sobre duas características da realidade representadas em modelos de SI: i) a distinção entre *coisas* do mundo e coisas que são partes de outras; ii) a distinção entre coisas e *propriedades*. Segundo os autores, essas situações são representadas de forma distinta em diferentes metodologias para desenvolvimento de SI. As representações variam pelas características de uma relação e pelo tipo de representação. Utilizando o BWW como referência, buscam entender como diferentes tipos de representação facilitam ou dificultam a compreensão humana. O objetivo é reduzir o custo de aprender por tentativa e erro sobre as limitações das diferentes linguagens de modelagem.

Opdahl e Henderson-Sellers (2004) propõem o uso de um *template* baseado no BWW, formalizado em *Object Constraint Language* (OCL)¹³, para avaliação de constructos da UML. Os constructos são divididos em suas partes atômicas, detalhando-se cada classe, propriedade, estado e evento em uma taxonomia. Em seguida, o *template* é utilizado para definir um par de constructos, os quais são comparados entre si para avaliação da sobreposição resultante. Definir um par de constructos consiste em representá-los através de um conjunto fixo de *entradas de alto nível*. Essas entradas são de quatro tipos: i) *entrada de instanciação*, usada para definir se um constructo representa os processos organizacionais no nível de classes das coisas ou no nível das coisas; ii) *entrada de classe*, usada para definir quais classes de coisas podem ser representadas por um constructo; iii) *entrada de propriedade*, usada para definir quais propriedades podem ser representadas por um constructo; iv) *entrada duração*, usada para definir se um constructo representa um evento, um estado, um processo ou o total de tempo de coisas.

Holten, Dreiling e Becker (2004) apresentam um método que define uma ontologia como metamodelo de uma linguagem de modelagem para gestão de SI. A ontologia é construída através de operações sobre a linguagem denominadas *ações linguísticas*. As ações linguísticas são de três tipos: i) *instanciação*, na qual as categorias que definem um conjunto de objetos são criadas como uma relação *instância-de*; ii) *subordinação*, a qual define relações *é-um* entre categorias de objetos; iii) *composição*, a qual é usada para criar categorias. Os componentes do metamodelo resultam da aplicação de ações linguísticas aos constructos da linguagem de modelagem original.

Milton (2000) utiliza a ontologia de Chisholm como ontologia de referência para avaliar a adequação de um modelo conceitual à realidade, bem como as diferenças e as similaridades entre linguagens que especificam esse modelo. Trata-se de um método que compara a ontologia de referência com os metamodelos ontológicos subjacentes à linguagens de

¹³ A OCL é uma linguagem utilizada para criar restrições que limitam o significado de constructos da UML.

modelagem conceitual. O objetivo é avaliar o nível de concordância entre a linguagem e a ontologia de referência. O método consiste nos seguintes passos: i) determinar o conjunto de conceitos da ontologia de referência selecionada; ii) determinar o conjunto de conceitos subjacentes à linguagem; iii) executar a comparação e avaliar os resultados. O autor enfatiza que o método independe da ontologia de referência selecionada e pode ser aplicado a um conjunto de linguagens de modelagem.

Guizzardi (2005) apresenta uma proposta para avaliar linguagens de modelagem, a partir do estudo da UML. O objetivo é avaliar, corrigir inconsistências e propor extensões ao metamodelo da UML, através da comparação com uma ontologia de referência denominada *ontologia de base*. Os constructos do metamodelo UML confrontados com a ontologia de referência são: i) o elemento *classe* e a *taxonomias de classes*; ii) os *classificadores*, por exemplo, classes, associações, tipos de dados e interfaces; iii) o elemento *propriedade*, por exemplo, atributos; iv) as representações de *relações meronímicas*, por exemplo, composição e agregação. O autor apresenta exemplos de fragmentos do metamodelo UML, comparando-os a ontologia de referência. Finalmente, propõe revisões aos constructos do metamodelo e as respectivas formas de implementação.

5. Discussão

A maioria das publicações sobre ontologias é proveniente da computação, mesmo que já existam desde a década de 90 estudos sobre o assunto na CI, como por exemplo os trabalhos de Vickery (1997), de Søerguel (1997), de Almeida (2002), de Gilchrist (2003), de Campos (2004), dentre outros. O assunto merece destaque e trata-se de oportunidade de pesquisa em CI. Ao longo do restante dessa seção discutem-se interfaces, enfatizando-se a natureza das etapas típicas do desenvolvimento de SI, as habilidades para conduzir tais etapas, dentre outras.

Conforme já mencionado, desde os anos 80 as pesquisas sobre ontologias em SI têm sido conduzidas na Engenharia de *Software* e na Representação do Conhecimento, áreas que objetivam o desenvolvimento de SI. O papel das ontologias nesse contexto é similar: representar o conhecimento de um domínio para uso na codificação do sistema. A despeito do desenvolvimento de sistemas ser considerado em geral uma atividade típica de tecnologia da informação, não consiste apenas de etapas técnicas. A etapa de modelagem conceitual objetiva gerar modelos para entendimento comum, ou seja, para comunicação entre pessoas e não entre máquinas (Mylopoulos, 1992). Na modelagem conceitual, habilidades relacionadas a comunicação e a capacidade de abstração são essenciais para a obtenção de bons resultados. Tal atividade resulta em interpretações inadequadas da realidade, caso seja conduzida sem métodos bem fundamentados, seja baseada puramente em decisões para um único contexto, em opiniões e em experiência prévia.

A modelagem inconsistente resulta assim em SI com várias deficiências, as quais muitas vezes se tornam aparentes apenas a longo prazo. Tal prática se coloca no sentido oposto a todo o esforço despendido para melhor organizar e manipular a informação e o conhecimento, tornando-os ativos reais nas empresas. Os SI são importantes mecanismos para retenção e para disseminação do conhecimento nas instituições modernas. Procedimentos nesse sentido são o objetivo do conjunto de práticas que se convencionou chamar de *gestão do conhecimento*. Em última instância, obtém-se SI incapazes de interoperar, tornando as instituições limitadas em sua capacidade de compartilhar e de re-utilizar, em outros contextos, o conhecimento apreendido nos modelos. Do ponto de vista das práticas usualmente aceitas, o problema parece maior quando se observa que métodos bem fundamentados são a exceção e não a regra.

Enquanto a Ciência da Informação é um campo dedicado a abordar os problemas comunicação de conhecimento e de seus registros entre seres humanos (Saracevic, 1995), a Ciência da Computação enfatiza a busca por melhorias nos processos algorítmicos. O objetivo distinto dessas áreas de pesquisa reflete em seus profissionais, em suas habilidades e em suas possibilidades de atuação, no âmbito de projetos de SI. A formação em computação enfatiza a

codificação e os procedimentos técnicos da atividade de desenvolvimento. Os profissionais da CI, em geral recebem fundamentação sobre a natureza da informação e do conhecimento, estudam teorias sobre organização da informação, princípios de classificação, fundamentos dos instrumentos de organização (como por exemplo, vocabulários controlados), estudam necessidades de usuários, dentre outros assuntos. As habilidades do profissional de CI são essenciais na criação de modelos, enfim, no entendimento da realidade, de forma que se possa obter SI eficientes quanto aos seus objetivos no contexto social em que atuam.

Acredita-se que apenas a partir da habilidade de entender a realidade é possível pensar em melhorias de longo prazo nos SI. Só assim é possível reduzir as diferenças entre conceitualizações distintas de SI, criadas por profissionais diversos. A questão central, que permeia todas aquelas citadas até aqui, diz respeito a necessidade de conhecimento, pelos participantes das atividades de modelagem, de técnicas sistemáticas de representação. Nesse contexto, afirma-se que contribuições da pesquisa da Ciência da Informação são relevantes e que o profissional de CI tem possibilidades pouco exploradas de atuação profissional.

6. Conclusão e perspectivas

Ao longo do presente artigo, forneceram-se fundamentos básicos da pesquisa sobre ontologias aplicadas a modelagem conceitual. Apresentou-se um breve histórico dos modelos para desenvolvimento de SI, destacando-se a importância da modelagem. Descreveu-se a relação das ontologias com a modelagem, propondo-se a utilização das ontologias de alto-nível e das ontologias filosóficas para definir ontologias de referência. Apresentaram-se ainda iniciativas diversas, as quais comprovam a viabilidade do uso de ontologias na modelagem conceitual. Discutiu-se a importância de técnicas sistemáticas de representação, bem como habilidades dos profissionais de CI que facilitam sua adoção.

Apesar da discussão e da importância cada vez maior dedicada à pesquisa de ontologias em CI, o assunto ainda encontra resistências na área. A justificativa, muitas vezes, é de que os principais autores não são da CI. Acredita-se, entretanto, que tal situação não é impedimento para que se busque incentivar abordagem conjunta ao assunto, dada a sua importância. Enfatiza-se, conforme já mencionado, que a pesquisa em ontologias é altamente interdisciplinar, abrangente e multifacetada. À comunidade de CI, interdisciplinar por natureza, cabe avançar sobre suas fronteiras no sentido de promover a evolução que se espera em contextos científicos.

Discutir contribuições da CI a outras áreas, interdisciplinaridade e interfaces exige discussão e aprofundamento. De forma a fundamentar discussões e estabelecer marcos teóricos, evitando abordagens meramente multidisciplinares, é necessário a busca por resultados empíricos. O presente artigo é parte de esforços de pesquisa em andamento no âmbito de programas brasileiros de pós-graduação em Ciência da Informação, os quais avaliam modelos de SI reais de órgãos da administração pública, visando melhorias na capacidade de governança eletrônica. Espera-se que esses esforços sejam incentivos para a busca por novas perspectivas em CI, e demonstrem a relevância da área para a comunidade científica.

Referências

- Almeida, M.B. (2002). Inter-operabilidade entre fontes de dados heterogêneas: um meta-modelo baseado em ontologias. Dissertation, Universidade Federal de Minas Gerais, Belo Horizonte, BR, Escola de Ciência da Informação.
- Abrial, J. R. (1974). Data semantics. In J.W. Klimbie & K.L. Koffeman (Eds.). Proceedings of the IFIP Working Conference Data Base Management (pp. 1–60). Amsterdam: North-Holland.
- Booch, G. (1993). Object-oriented analysis and design with applications, 2nd ed. Redwood: Benjamin Cummings.
- Bosak, R., Richard, F. Clippinger, R. F., Dobbs, C., Goldfinger, R., Jasper, R. B., Keating, W., Kendrick, G., & Sammet, J.E. (1962). An information algebra: Phase 1 report - language structure group of the CODASYL development committee. Communications of the ACM, 5(4), 190–204.
- Bunge, M. (1977). Ontology I: The furniture of the world: Treatise on basic philosophy (Vol. 3–4). Boston: Reidel.

- Campos, M. L. A. (2004). Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. *Ciência da Informação*, 33(1), 22-32.
- Chen, P. (1976). The entity-relationship model: towards a unified view of data. *ACM Transactions on Database Systems*, 1(1), 9–36.
- Chisholm, R. (1996). *A realistic theory of categories: an essay on ontology*. Cambridge: Cambridge University Press.
- Codd, E. F. (1979). Extending the database relational model to capture more meaning. *ACM Transactions on Database Systems*, 4(4), 397–434.
- Evermann, J. M. (2003). *Using design languages for conceptual modelling: the uml case*. PhD Thesis, Westfälische Wilhelms University, Münster, GE, The Faculty of Graduate Studies, Management and Information Systems.
- Fettke, P., & Loss, P. (2005). Ontological analysis of reference models. In P. Green & M. Rosenmann (Eds.). *Business systems analysis with ontologies* (pp. 56–81). Hershey: Idea Group.
- Fonseca, F. (2007). The double role of ontologies in information science research. *Journal of the American Society for Information Science and Technology*, 58(6), 786–793.
- Frigg, R. (2006). *Models in science*. Retrieved July 18, 2008, from <http://plato.stanford.edu/entries/models-science/>
- Gangemi, A., Guarino, N., Masolo, C., & Oltramari, A. (2001). Understanding top-level ontological distinctions. In *IJCAI Workshop on Ontologies and Information Sharing*. Seattle, WA.
- Genesereth, M. R., & Nilsson, L. (1987). *Logical foundation of AI*. San Francisco: Morgan Kaufman.
- Gilchrist, A. (2003). Thesauri, taxonomies and ontologies: an etymological note. Retrieved March 2, 2007, from <http://dois.mimas.ac.uk/DoIS/data/Articles/julkokltny:2003:v:59:i:1:p:7-18.html>
- Grenon, P., Smith, B., & Goldberg, L. (2004). *Biodynamic ontology: applying BFO in the biomedical domain*. Retrieved January 15, 2008, from <http://ontology.buffalo.edu/medo/biodynamic.pdf>
- Gruber, T. (1993). *What is an ontology?* Retrieved September 14, 2002, from <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>
- Guarino, N. (1998). Formal ontology and information systems. In N. Guarino (Ed.). *Formal ontology in information systems* (pp. 3–15). Amsterdam: IOS Press.
- Guarino, N., & Welty, C. (2002). Evaluating ontological decisions with OntoClean. Retrieved March, 24, 2006 from <http://citeseer.ist.psu.edu/guarino02evaluating.html>
- Guizzardi, G. (2005). *Ontological foundations for structural conceptual models*. PhD Thesis, University of Twente, Twente, NL, Centre for Telematics and Information Technology.
- Guizzardi, G., & Wagner, G. (2004). *A Unified Foundational Ontology and some applications of it in business modeling*. Retrieved July, 22, 2008, from <http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-125/paper2.pdf>
- Herre, H., Heller, B., Burek, P., Hoehndorf, R., Loebe, F., & Michalek, H. (2006). *General Formal Ontology; A Foundational Ontology Integrating Objects and Processes*. Retrieved January 11, 2009, from <http://www.onto-med.de/ontologies/gfo/>
- Holten, R., Dreiling, A., & Becker, J. (2005). Ontology-driven method engineering for information systems. In P. Green & M. Rosenmann (Eds.). *Business systems analysis with ontologies* (pp. 174–217). Hershey, PA: Idea Group.
- Jacobson, I., Christerson, M., Jonsson, P., & Overgaard, G. (1992). *Object-oriented software engineering: A use case driven approach*. Boston: Addison-Wesley.
- Jardine, D. A. (1976). The ANSI/SPARC DBMS model. *Proceedings of the second SHARE Working Conference on Database Management Systems*. Amsterdam: North Holland.
- Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A., & Schneider, L. (2003). *WonderWeb library of foundational ontologies: preliminary report*. Retrieved July 11, 2008, from
- Milton, S. (2000). *An ontological comparison and evaluation of data modelling frameworks*. PhD Thesis, University of Tasmania, Hobart, AU, School of Information Systems.
- Milton, S., & Kazmierczak, E. *An ontology of data modeling languages: a study using a common-sense realistic ontology*. (2004). Retrieved May 17, 2008, from <http://www.dis.unimelb.edu.au/staff/smlton/pdf/JDM2003.pdf/>

- Mylopoulos, J. (1992). Conceptual modelling and telos. In P. Loucopoulos & R. Zicari (Eds.). *Conceptual modelling, databases and case: An integrated view of information systems development*. New York: John Wiley & Sons.
- OMG (2003). UML 2.0 Infrastructure Specification. Retrieved Dec 3, 2008, from <http://www.omg.org/docs/ptc/03-09-15.pdf>
- Opdahl, A. L., & Henderson-Sellers, B. (2004). A template for defining enterprise modeling constructs. *Journal of Database Management*, 15(2), p. 39-73.
- Peckham, J., & Maryanski, F. (1988). Semantic data models. *ACM Computing Surveys*, 20(3), 153–189.
- Reed, S. L., & LENAT, D. B. (2002). Mapping Ontologies into Cyc. Retrieved November 20, 2005, from http://www.cyc.com/doc/white_papers/mapping-ontologies-into-cyc_v31.pdf
- Rumbaugh, J., Blaha, M., Premerlani, W., Eddy, F., & Lorensen, W. (1991). *Object-oriented modeling and design*. New York: Prentice Hall.
- Saracevic, T. (1995). Interdisciplinary nature of Information Science. *Ciência da Informação*, 24(1), p. 36-41.
- Scheer, A. W. (2000). *ARIS: Business Process Frameworks*. 3rd ed. New York: Springer-Verlag.
- Shanks, G., Tansley, E., & Weber, R. (2003). Using Ontology to Validate Conceptual Model. *Communications of the ACM*, 46(10), 85-89.
- Smith, B. (1995). Formal Ontology, Common Sense and Cognitive Science. *International Journal of Human-Computer Studies*, 43(5/6), 641-667.
- Smith, B. (2003). Ontology and information systems. Retrieved March 20, 2005, from [http://www.ontology.buffalo.edu/ontology\(PIC\).pdf](http://www.ontology.buffalo.edu/ontology(PIC).pdf)
- Smith, B. (1998). The basic tools of formal ontology. In N. Guarino (Ed.). *Proceedings of formal ontology in information systems* (pp. 3–15). Amsterdam: IOS Press.
- Smith, B., & Welty, C. (2001). Ontology: Towards a new synthesis. In B. Smith & C. Welty (Eds.). *Proceedings of the International Conference on Formal Ontology in Information Systems* (pp. 3–9). New York: ACM Press.
- Søerguel, D. (1997). Functions of a thesaurus, classification and ontological knowledge bases. Retrieved December 12, 2003, from <http://www.clis.umd.edu/faculty/soergel/soergelfctclass.pdf>
- Sowa, J. F. (2000). *Knowledge Representation: logical, philosophical, and computational foundations*. Pacific Grove: Brooks-Cole.
- Vickery, B. C. (1997). Ontologies. *Journal of Information Science*, 23(4), 227–286.
- Wand, Y., & Weber, R. (1993). On the ontological expressiveness of information systems analysis and design grammars. *Journal of Information Systems*, 3(4), 217-237.
- Wand, Y., & Weber, R. (1990). Mario Bunge's ontology as a formal foundation for information systems concepts. In: P. Weingartner & J.W.G. Dorn (Eds.). *Studies on Mario Bunge's treatise*. Amsterdam: Rodopi.
- Wand, Y., Storey, V.C., & Weber, R. (1999). An ontological analysis of the relationship construct in conceptual modeling. *ACM Transactions on Database Systems*, 24(4), 494–528.
- Wyssusek, B., & Klaus, H. (2005). On the Foundation of the Ontological Foundation of Conceptual Modeling Grammars. Retrieved July 18, 2007, from <http://kybele.escet.urjc.es/PHISE05/papers/sesionIV/WyssusekKlaus.pdf>
- Young, J. W., & Kent, H. K. (1958). Abstract formulation of data processing problems. *Journal of Industrial Engineering*, 9(6), 471–479.