

VI CIFORM

ONTOLOGIAS: ALTERNATIVA PARA A REPRESENTAÇÃO DO CONHECIMENTO EXPLÍCITO ORGANIZACIONAL

JOCELMA A. RIOS¹

(jocelma.rios@cpunet.com.br)

Os estudos sobre a gerência estratégica da informação têm recebido cada vez mais importância, em virtude do crescimento exponencial do número de fontes de dados, dificultando a recuperação da informação. A questão é que o valor econômico advindo da geração, uso e venda da informação está crescendo muito mais depressa que o valor agregado pela produção tradicional de bens e de serviços. Entretanto, transformar toda esta informação em conhecimento aplicável não é uma tarefa fácil, por se tratar de um grande volume, disperso em bases descentralizadas e não integradas, o que impacta num alto custo com armazenamento, classificação e controle. Além de desviar atenção para aquilo que não é importante, aumenta a dificuldade de encontrar o que é realmente valioso. Em função da problemática instalada, surge a necessidade de codificar o conhecimento, para que o torne acessível aqueles que dele precisam. Para isso, estudos vêm sendo feitos com a utilização de ontologias, no sentido de unificar a conceituação sobre os diversos aspectos de um dado domínio de conhecimento, permitindo a troca e reutilização de informações, tanto para entre sistemas quanto entre pessoas. Criar ontologias que possam ser reutilizadas e, por conseguinte, ferramentas que possibilitem isto, têm sido os principais objetivos de pesquisas nesta área. É mostrado, neste trabalho, que devido ao emprego de uma sintaxe e semântica formal, para a construção de ontologias, definindo conceitos, propriedades e relações do domínio, possibilita a formalização de requisitos complexos, gerando uma especificação consistente para os sistemas baseados no conhecimento (SBCs), de modo a eliminar ambigüidades e consistências.

Palavras-chave: gerência do conhecimento, ontologia, engenharia de software, reuso, integração de sistemas.

¹ Mestranda em Redes de Computadores e Computação pela UNIFACS

INTRODUÇÃO

Ontologia é um conceito antigo já utilizado pela filosofia e outras áreas, desde Aristóteles, filósofo grego e discípulo de Platão, responsável pela classificação dos seres vivos, até então conhecidos. Na visão de Uschold e Gruninger (1996), **ontologia é um termo usado para se referir a uma concepção compartilhada de algum domínio de interesse, e pode ser utilizada para unificar o processo de solução de problemas relativos ao domínio em questão**. O grande interesse em ontologias, hoje, dá-se pela necessidade de haver, cada vez mais, uma maior **troca e reutilização de informações entre os sistemas e as pessoas**. Contudo, um fator impeditivo para o compartilhamento de informações é o uso de diferentes conceitos para um mesmo domínio, já que o conhecimento era adquirido para resolver tarefas específicas (FALBO, 1999). Além disso, **há uma grande variedade de ferramentas, modelos e linguagens de desenvolvimento, criando verdadeiras ilhas de informação**. Essas diferenças dificultam a troca de informações entre diferentes sistemas, bem como a interoperabilidade e o reuso. A partir daí, surge a **necessidade de codificar o conhecimento** para que o torne acessível aqueles que dele precisam. Nas organizações, uma codificação ou representação formal e explícita do conhecimento promove o gerenciamento do conhecimento, categorizando, descrevendo, modelando, estimulando e inserindo regras e padrões definidos.

As ontologias servem para criar **terminologias únicas**, de modo que o conhecimento possa ser compartilhado e reutilizado. Criar ontologias que possam ser reutilizadas e, por conseguinte, ferramentas que possibilitem isto, têm sido os principais objetivos de pesquisas nesta área (BUSSLER, FENSEL, MAEDCHE, 2002; JOSHI, HOLSAPPLE, 2002). Para Chandrasekaran e outros (1999) *apud* Oliveira (1999), o uso de ontologias vêm confirmar que os esforços em torno dos mecanismos de representação do conhecimento (*eg.* regras, *frames*, redes neurais, etc.) de nada adiantam, se não existe um bom conteúdo e organização sobre o conhecimento do domínio em que se deseja trabalhar.

CONCEITOS BÁSICOS

O que é uma ontologia? **Uma ontologia é a especificação formal explícita de uma conceitualização compartilhada** (GRUBER, 1993)². Guarino (1997), contrapõem-se a essa definição, acrescentando que o grau de especificação de uma conceitualização de uma linguagem,

² Segundo Oliveira (1999), a definição de ontologia, proposta por Tom Gruber, é um marco nos estudos sobre o tema. Guarino (1997) acrescenta que esta definição é considerada um senso comum, entre os estudiosos da área.

utilizada para uma base de conhecimento depende do propósito desejado para a ontologia; concepção compartilhada por Uschold (1996).

A ontologia pode ser definida como o ramo da metafísica que trata da natureza do ser. O termo, que foi adaptado pela comunidade de inteligência artificial, serve para se referir a um conjunto de conceitos ou termos usados para descrever algumas áreas do conhecimento ou para construir uma representação deste. As definições encontradas para ontologia são as mais variadas dependendo da área em questão, como pode ser mostrado a seguir:

Quadro 1: Conceitos de ontologia

Área	Conceito
Filosofia	É a ciência que trata de seres em geral, enquanto seres, seus nomes e propriedades, concebidos como tendo uma natureza comum que é inerente a todos e a cada um dos seres.
Linguagem e cognição	Refere-se a tudo que existe no mundo composto por objetos, mudanças e relações entre eles. Ontologia pode ser baseada no mundo, na mente, no intelecto, na cultura ou na linguagem.
Medicina	É uma doutrina que estuda o ser da doença, como se a enfermidade existisse em conformidade a um tipo bem definido, a uma essência.
Sistemas de informação	Segundo Chandrasekaran, Josephson e Benjamins, (1998), ontologias são teorias de conteúdo sobre os tipos de objetos, propriedades de objetos e relacionamentos entre objetos que são possíveis em um domínio de conhecimento específico.
Inteligência artificial	Guarino (1997) define a ontologia como uma caracterização axiomática do significado do vocabulário lógico, e, para Sowa e Dietz (1999), a ontologia define os tipos de coisas que existem no domínio de uma aplicação.

Segundo Uschold (1996), uma ontologia é, virtualmente, uma manifestação do entendimento do domínio, compartilhada e comum entre integrantes de uma comunidade. Como exemplo, um artigo sobre o Sol, a Terra, as estrelas e os planetas, poderá utilizar ontologias diferentes de acordo com o contexto. Na ontologia da astronomia moderna, a associação “Sistema Solar” é um exemplo (instância) da associação “Sistema Planetário”. Na antiguidade, esta representação não seria compreendida, já que o Sol não era considerado uma estrela, nem a Terra um planeta, portanto a notação geral “Sistema Planetário” não teria sentido.

É necessário esclarecer, que **ontologia e base de conhecimento não são sinônimos**. Sowa e Dietz (1999) consideram a base do conhecimento como uma definição informal de uma coleção de informação que inclui uma ontologia como componente, ou seja, a ontologia é o alicerce sobre o qual o conhecimento pode ser construído. A ontologia provê, ainda, um conjunto de conceitos ou termos com o objetivo de descrever algum domínio, enquanto a base de conhecimento utiliza estas estruturas para representar o que é verdade sobre algum mundo real ou hipotético (OLIVEIRA, 1999). Oliveira (1999) conclui ainda, que mesmo que os sistemas usem a mesma linguagem de representação e tenham bons meios de comunicação

entre eles, compartilhamento do conhecimento pode ser impossível devido à não utilização de uma terminologia comum de domínio.

ELEMENTOS

Tomando como base o conceito de ontologia, proposto por Gruber (1993), que a define como uma representação explícita de uma conceituação compartilhada, Falbo (1999) discorre sobre o conteúdo previsto em uma ontologia, efetuando comparações sobre as abordagens citadas nos trabalhos de Gruber (1993), Guarino (1994, 1997) e outros autores. O autor conclui que uma ontologia consiste de conceitos, relações entre estes conceitos, definições, propriedades e restrições (representadas através de axiomas). Gruninger e Fox (1995) salientam, que os axiomas são especialmente importantes para a definição da semântica dos termos contidos na ontologia, pois determinam as regras para sua interpretação.

Na visão de Almeida e Bax (2003), para representar amplamente um domínio, uma ontologia é formada por uma hierarquia de classes (organizadas através de uma taxonomia³), conceitos destas classes (*thesaurus*), axiomas e instâncias (exemplos específicos de uma classe). A hierarquia de classes é útil no sentido de prever a possibilidade de implementação dos conceitos de herança (*eg.* a sub-classe “apartamento” herda características da classe “moradia”).

Para este estudo, foi considerada uma ontologia que possua, além dos elementos anteriormente citados, a especificação de atributos (propriedades que identificam unicamente uma instância) das classes, conjunto de valores que estes atributos poderão assumir, valor padrão, cardinalidade e restrições.

OBJETIVOS

A cada sistema desenvolvido, nova base de conhecimento é extraída a partir do domínio no qual o sistema atuará. Mas, da maneira como o conhecimento, presentemente, é extraído e modelado, dificulta a interoperabilidade deste sistema com o legado, além de reduzir a possibilidade de reuso, já que o conhecimento obtido possui fortes vínculos com a sua aplicação (FALBO, 1999). As ontologias são especificações formais que dão suporte ao compartilhamento e reuso do conhecimento. As ontologias estabelecem uma junção entre membros de uma comunidade de interesse, podendo estes ser humanos ou agentes autômatos.

³ A taxonomia é definida como a técnica de classificação, que é uma distribuição em hierarquia de classes. De modo mais aderente ao trabalho, a taxonomia pode ser definida como ciência que lida com a descrição, identificação e classificação dos organismos, individualmente ou em grupo.

Elas podem ser usadas explicitamente representando informação semântica e semi-estruturada, o que permite um suporte sofisticado à aquisição, manutenção e acesso ao conhecimento. Para isto, torna-se necessário uma metodologia de acesso inteligente a grandes volumes de informação textual semi-estruturada, armazenada em documentos, proveniente de diversas fontes, como os portais corporativos, com o objetivo de potencializar o uso das ontologias na gerência do conhecimento organizacional.

Em Uschold (1996) e Uschold e Gruninger (1996), foram identificadas diferentes **categorias de uso** para as ontologias:

- a) **Comunicação** entre pessoas e organizações, com diferentes necessidades e pontos de vista de um contexto particular, através da redução da ambigüidade;
- b) **Interoperabilidade** entre sistemas, incluindo sistemas multi-agentes, unificando linguagens de representação, bases de conhecimento e ferramentas utilizadas;
- c) **Engenharia de software**, na especificação, manutenção e reuso.

De uma maneira geral, as ontologias são especialmente úteis na **gerência do conhecimento para recuperação da informação**, pois unifica termos, conceitos, categorias e relações de um mesmo domínio, permitindo o reuso. Segundo Davis e outros (1993) *apud* FALBO (1999), a noção de representação de conhecimento pode ser melhor compreendida através dos papéis que desempenha.

CLASSIFICAÇÃO

Na visão de Uschold (1996), as ontologias podem ser classificadas em **três categorias**, quanto ao tipo de conhecimento que representam:

- a) **Ontologia de domínio⁴**: conceituações de domínios particulares;
- b) **Ontologia de tarefas**: conceituações sobre a resolução de problemas independentemente do domínio onde ocorram;
- c) **Ontologia de representação**: conceituações que fundamentam os formalismos de representação do conhecimento.

Em Guarino (1994, 1997), é proposta outra abordagem, classificando as ontologias em **níveis**:

- a) **Ontologia genérica**: conceitos muito genéricos, independentes de um problema ou domínio particular;

⁴ Domínio, segundo OLIVEIRA (1999), é considerado um problema ou uma área de tarefa na qual muitas aplicações semelhantes serão desenvolvidas.

- b) **Ontologia de domínio:** descrevem o vocabulário relativo a um domínio específico, através da especialização de conceitos presentes na ontologia de alto nível;
- c) **Ontologia de tarefa:** descreve o vocabulário pertinente a uma tarefa genérica ou específica através da especialização de conceitos presentes na ontologia de alto nível;
- d) **Ontologia de aplicação:** descreve conceitos dependentes do domínio e da tarefa particulares.

Uschold (1996) acrescenta que as ontologias podem ser classificadas quanto ao grau de formalidade, sendo: altamente informal, estruturada informal, semiformal ou rigorosamente formal.

LINGUAGENS DE REPRESENTAÇÃO

Para representar uma conceitualização compartilhada, é necessária uma linguagem de representação conhecida. Segundo Almeida e Leite (2004), as origens dos estudos de uma linguagem única advém das pesquisas em volta das linguagens do tipo **markup**, realizadas pelo CERN (*Conseil Européen pour la Recherche Nucléaire*) e, posteriormente, com a criação da WWW (*World Wide Web*) em 1989, criando o HTML. Existem hoje muitas linguagens de representação definidas, baseadas na **Lógica de Primeira Ordem** e no **XML**, mas com diferentes expressividades e propriedades computacionais. Considera-se a tendência de que o XML passe a ser a linguagem padrão para troca de dados na Web, e portanto, é desejável também a utilização de sua sintaxe para modelar as ontologias. O HTML, apesar de também ser uma linguagem de marcação, além de precursor do XML, foi descartado para esta tarefa por apresentar duas fortes limitações: falta de estrutura e impossibilidade de validação da informação exibida (BREITMAN, LEITE, 2004).

Partindo deste pressuposto, muitas linguagens foram propostas, como exemplos incluem: **RDF**⁵ (*Resource Description Framework*), **RDF-Schema** (*Resource Description Framework Schema Language*), **OIL**⁶ (*Ontology Interchange Language*), **DAML+OIL**⁷ (*DARPA Agent Markup Language + OIL*), **OWL**⁸ (*Web Ontology Language*) e **OML**⁹ (*Ontology Markup Language*). Há ainda as linguagens com suporte gráfico para a representação de ontologias, como é o caso da **LINGO** (Linguagem Gráfica para descrever

⁵ Cf. www.w3c.org/rdf

⁶ Cf. www.ontoknowledge.org/oil

⁷ Cf. www.w3.org/TR/daml+oil-reference

⁸ Cf. www.w3.org/2004/OWL

⁹ Cf. www.ontologos.org/OML/OML%200.3.htm

Ontologias) e da **CML** (*Conceptual Modelling Language*), proposta na metodologia CommonKADS (SCHREIBER, 1994). A representação gráfica é benéfica para a compreensão da ontologia criada por parte de seus usuários.

Algumas destas linguagens propostas estão detalhadas a seguir:

Quadro 2: Linguagens de representação de Ontologias

Linguagem	Descrição
RDF / RDF-Schema	Implementa metadados, baseando-se na notação do XML como sintaxe de codificação. A linguagem RDF foi criada com o objetivo de facilitar o intercâmbio de informações, que podem ser interpretadas por máquinas, entre aplicativos via <i>Web</i> , além de adicionar semântica formal, representar metadados, e não dados propriamente ditos, e facilitar a representação do conhecimento. O RDF possui um modelo de representação simples, por isso mais flexível, que permitir a interpretação semântica do conhecimento modelado, com a utilização de conectivos lógicos, de negação, disjunção e conjunção (BREITMAN, LEITE, 2004). A especificação RDF é dividida em duas partes principais: RDF e RDF-Schema. A primeira define como descrever recursos através de suas propriedades e valores, enquanto a segunda define propriedades específicas, restringindo sua utilização.
DAML+OIL	Diante da limitação do RDF, o projeto On-to-Knowledge propôs a linguagem OIL, cuja contribuição decorre do uso de semântica formal e um mecanismo de inferência, fornecido através da lógica de descrição (BREITMAN, LEITE, 2004). Um sucessor do OIL é DAML+OIL. A linguagem DAML foi desenvolvida como uma extensão para XML e RDF, acrescentando construtos mais expressivos. A última versão da linguagem, a DAML+OIL, provê um conjunto de construções com o objetivo de criar ontologias e marcar informações de forma que seja compreendido e legível por máquina. DAML+OIL provê uma infra-estrutura básica que permite às máquinas fazerem a mesma classificação de inferências simples que os seres humanos fazem.
OWL	A OWL pode ser utilizada por aplicações que precisam processar o conteúdo da informação, ao invés de apenas disponibilizar conteúdo. Para a representação, utiliza a lógica descritiva para explicitação do conhecimento. A OWL facilita a leitura de conteúdo <i>Web</i> suportado por XML, RDF e RDF-Schema, provendo um vocabulário adicional com uma semântica formal. A OWL ainda provê três outras linguagens: OWL <i>Lite</i> , OWL <i>DL</i> , e OWL <i>Full</i> .
Lingo	A linguagem Lingo foi proposta por Falbo (1999), a fim de garantir independência de semântica numa linguagem gráfica. Na concepção de Falbo, a representação gráfica é fundamental para facilitar a comunicação entre os elementos do domínio de interesse. Pela característica da linguagem, as notações da Lingo são capazes de capturar certos axiomas de forma implícita, pois utiliza diferentes tipos de notação para diferentes tipos de associação.

Todas estas utilizam XML, com pequenas diferenças nas *tags*. O XML, além de incluir informação semântica, já é uma sintaxe comum que muitas ferramentas utilizam. Entretanto, com o advento da Web Semântica, onde é esperado que o dado seja legível por máquina, e não somente pela interpretação do ser humano, o XML mostra algumas limitações. A maior limitação do XML está em só descrever gramática, ou seja, tem-se a liberdade para definir e usar *tags*, atributos e outras primitivas da linguagem de um modo arbitrário, designando diferentes semânticas para descrever o modelo do domínio conceitual que se deseja, mas, visto que o XML não impõe regras para tal descrição, e que há muitos meios de denotar coisas semanticamente equivalentes, ele se torna difícil para se reconstruir o significado semântico de um documento XML.

Para que seja feita a opção por quaisquer das linguagens disponíveis, é preciso definir critérios, como facilidade de integração e compreensão, para efetuar uma comparação mais ajustada. Optar por uma linguagem mais popular pode ser uma boa estratégia, pois a dificuldade na compreensão da sintaxe da linguagem pode inviabilizar a reutilização ou a manutenção da ontologia (FERNANDEZ-LOPEZ *et al.*, 1999).

FERRAMENTAS DE EDIÇÃO

A criação de ontologias comuns para um conjunto de agentes não é uma tarefa trivial, já que significa tornar explícito algo que normalmente é implícito. As ferramentas de apoio à ontologia podem simplificar consideravelmente a criação de metadados para recursos disponíveis na Web. Um editor de metadados, bastante simplificado, é o *Reggie Metadata Editor*, baseado em Java, desenvolvido no *Resource Discovery Unit* do DSTC, que exporta para HTML 3.2, HTML 4.0 e RDF.

O *Protégé*¹⁰ é uma ferramenta, de interface gráfica, que dá suporte à aquisição de ontologia e conhecimento. O Protégé suporta a geração de interface para o usuário baseada em ontologia, a qual torna ideal para o uso de ontologia baseada em metadados. Uma outra característica, deste editor, é contemplar uma arquitetura modulada, permitindo a inserção de novos recursos, além de possuir código aberto.

O *OilEd*¹¹, um exemplo de editor simples, permite o usuário criar ontologias utilizando as linguagens OIL e DAML+OIL. A intenção por trás do OilEd é disponibilizar um editor de código aberto, simples, que demonstra e estimula o uso destas linguagens. O OilEd permite a verificação da consistência e classificação automática, mas não é um ambiente completo para desenvolvimento de ontologias, já que não suporta o desenvolvimento em larga escala, a migração e a integração de ontologias, bem como seu versionamento, argumentação e muitas outras atividades que envolvem a construção de ontologias.

Além dos já citados, existem outros editores de ontologias que estão sendo utilizados, como o *Ontolingua* (GRUBER, 1992), o *WebOnto* (DOMINGUE, 1998), *WebODE* (CORCHO *et al.*, 2005) e o *OntoEdit*¹², que possui ambiente gráfico, que permite a criação de ontologias utilizando linguagens como XML, RDF, DAML+OIL, entre outras (SURE, STAAB, STUDER, 2002).

¹⁰ Cf. <http://protege.stanford.edu/doc/users.html>

¹¹ Cf. www.img.cs.man.ac.uk/oil

¹² Cf. www.ontoprise.de/co-produ_too3.htm

ONTOLOGIAS APLICADAS

Existem muitos projetos sendo desenvolvidos, e alguns já finalizados, com objetivos em criar ontologias genéricas que possam ser utilizadas por diversos sistemas em todo o mundo. Um exemplo disso é o projeto *Harmony*¹³, desenvolvido através de uma parceria internacional entre o DSTC (*Distributed Systems Technology Centre*), JISC (*Joint Information Systems Committee*) e o NSF (*National Science Foundation*).

Um outro projeto é o *On-To-Knowledge*¹⁴, que aplica ontologias para disponibilizar informações eletronicamente, com objetivo de melhorar a qualidade da gerência do conhecimento em grandes e distribuídas organizações. O foco é criar facilidades para aquisição, manutenção e acesso *on-line* a informações vindas de diversas fontes, promovendo a sua reutilização.

Diversos outros projetos foram desenvolvidos nas mais variadas áreas:

Quadro 3: Aplicações de Ontologias

Área	Descrição do projeto
Educação	Para os conteúdos educacionais, são de fundamental importância indicações de autoria, a apresentação, a disponibilização, a estratégia pedagógica, a avaliação, assim como a evolução contínua destes conteúdos, dado o caráter dinâmico do conhecimento. A modelagem conceitual de conteúdos educacionais busca capturar as partes relevantes do assunto que se deseja ensinar, e estrutura adequadamente de acordo com determinada estratégia pedagógica. Ambientes de gerenciamento de educação " <i>on-line</i> " armazenam o conteúdo de forma estruturada, controlam o acesso ao material didático e permitem alguma forma de avaliação (INFO-EDUCACAO, 2000).
P&D (Pesquisa e Desenvolvimento)	Em Pacheco e Kern (2001), é descrita a problemática da integração de sistemas de informações sobre ciência e tecnologia, que resultou na Plataforma Lattes, tendo sido definida pelo consenso de peritos de várias instituições de ensino superior. Neste processo de desenvolvimento, foi preciso criar uma ontologia comum, denominada LPML (Linguagem de Marcação da Plataforma Lattes), para que os termos utilizados fossem compreendidos por todas as aplicações que fazem uso dos dados armazenados pela plataforma Lattes. Em 2002, Bonifácio, apresentou um modelo de metadados com semântica para o Currículo Lattes, baseado numa ontologia especificada na linguagem DAML+OIL.
B2C	As operações B2C têm crescido na <i>Web</i> , apesar da falta de interoperabilidade de sistemas e sobrecarga de informação. Foi desenvolvida por Guimarães (2002), uma ontologia para produtos e outra para lojas, onde foi utilizada a linguagem DAML+ OIL para implementação das ontologias. Estas ontologias tiveram o objetivo de auxiliar a busca de determinado produto em várias lojas, bem como a recomendação de produtos a clientes.
SIG (Sistemas de Informações Geográficas)	Foi sugerido um sistema que utiliza um repositório de objetos geográficos interoperáveis, sendo os objetos extraídos de múltiplos bancos de dados independentes e com um mapeamento baseado em orientação a objetos para criar os objetos a partir das ontologias. Esta abordagem proporciona um alto grau de interoperabilidade ou integração parcial de informações quando a integração completa não é possível e o reaproveitamento de classes já desenvolvidas. Estas aplicações são caracterizadas por um extensivo uso de ontologias

¹³ O projeto *Harmony* foi desenvolvido no período de 1999 a 2002. Cf. HARMONY (www.metadata.net/harmony), DSTC (www.dstc.edu.au/), JISC (www.jisc.ac.uk/) e NSF (www.nsf.gov/).

¹⁴ Cf. www.ontoknowledge.org.

	explícitas desde sua concepção até seu uso, inclusive (FONSECA, EGENHOFER, 1999).
Web Semântica	É uma extensão da <i>Web</i> atual, na qual é dada à informação um significado, permitindo a colaboração, na tentativa inversa de solução para a diversidade de informações disponibilizadas na <i>Internet</i> , de forma semi-estruturada e não-estruturada. (BREITMAN, LEITE, 2004). Ao invés de pensar na informação para os humanos, a idéia é pensar na máquina (BUSSLER, FENSEL, MAEDCHE, 2002). O projeto da <i>Web Semântica</i> foi gerenciado pelo consórcio W3C e projetado para ser um sucessor do projeto Metadados, cujos princípios são os mesmos (incluir "informação sobre a informação" na <i>Web</i>), mas que trabalhava com a linguagem HTML, o qual não permite criar categorias semânticas.

PROCESSOS DE AQUISIÇÃO E CONSTRUÇÃO

A engenharia de ontologias requer sistematização e empenho de todos os envolvidos (engenheiros e usuários), exigindo critérios de controle de qualidade, verificação e validação. Em função disso, alguns pesquisadores identificaram critérios mais propícios para avaliar as ontologias construídas. Dessa forma, antes de iniciar o processo de construção de uma ontologia, deve-se, primeiro, definir os usuários potenciais, o propósito da ontologia, e como a mesma será utilizada. Passar por cima destas etapas pode provocar desentendimentos entre os envolvidos, gerando re-trabalho, e em alguns casos, desistência.

Construir uma ontologia não é uma tarefa trivial (HOLSAPPLE, JOSHI, 2002). O projeto e o desenvolvimento de ontologias vêm deixando de ser uma arte para se transformar em uma ciência (FALBO, 1999). Para construir uma ontologia, é necessário capturar o conhecimento de um domínio de forma que não o restrinja ao extremo, nem o generalize demasiadamente, a fim de promover um entendimento compartilhado. O comprometimento de grupos de pessoas e o uso de metodologias, processos ou métodos bem definidos se faz importante.

O uso de uma metodologia¹⁵ para construir ontologias busca criar um processo unificado. A unificação de diversos métodos, oriundos de metodologias distintas em um processo abrangente, permite o projeto e o desenvolvimento de ontologias e bases de conhecimento. Além disso, é necessário estabelecer o grau de formalidade requerido, que em geral, é incrementado à medida que aumenta o grau de automação da tarefa em questão (USCHOLD, 1996). É preciso salientar que dependendo do domínio, da formalidade requerida e do propósito da ontologia, será necessário fazer inúmeras iterações, de modo que a ontologia passe por revisões, evoluindo a cada iteração.

Apesar da ontologia ser um objeto de estudos recentes, diversas propostas nesse sentido já foram apresentadas em Fernandez-Lopez, Gómez-Perez e Juristo (1997), Guha e

¹⁵ Metodologia é um conjunto de atividades e resultados associados a essas atividades com o objetivo de garantir a geração de um produto final, seja este produto um software, uma ontologia ou uma ontologia populada com sua base de conhecimento.

Lenat (1990), Gruninger e Fox (1995), Guarino e Welty (2000), Holsapple e Joshi (2002), Staab e outros (2001), Uschold (1996), Uschold e Gruninger (1996) e em diversos outros trabalhos, bem como sua aplicação em projetos acadêmicos e em parceria de empresas.

Neste trabalho, é proposta uma metodologia adaptada com base nas metodologias propostas em Uschold e Gruninger (1996) e Uschold (1996), e contribuições da proposta de Sure, Staab e Studer (2002). O processo de aquisição da ontologia ocorre durante a fase de levantamento de requisitos do desenvolvimento de um SBC, e contempla atividades de definição de propósito e escopo, construção, refinamento, avaliação e documentação, conforme é mostrado a seguir:

Definição do propósito e do escopo: antes de qualquer coisa, é importante esclarecer a razão pela qual a ontologia está sendo discutida. Para que um dado grupo de pessoas se dedique a esta atividade, é imperativo que saibam o benefício que obterão, após aquele esforço. Por isso, é importante a inclusão de atividades ligadas ao *endomarketing* do projeto. A fim de envolver os *stakeholders*, pode-se utilizar técnicas *brainstorm* e os grupos de discussão, que são técnicas de grande valia para extrair os reais requisitos referentes à ontologia a ser gerada. Um outro ponto importante é a correta definição dos envolvidos no processo, especialmente os especialistas do domínio.

Construção: com a definição clara dos objetivos a se alcançar, com a construção de ontologias, pode-se iniciar a atividade de captura da ontologia, propriamente dita. Esta atividade consta da identificação dos conceitos-chave e relacionamentos do domínio, uma descrição detalhada e precisa destes, bem como a identificação de termos que se referem a tais conceitos e relacionamentos. Para promover uma participação mais ativa por parte dos usuários, a construção da ontologia deve seguir uma abordagem *top-down*. Esta abordagem parte de uma visão geral sobre o domínio para os casos mais específicos, possibilitando uma melhor compreensão de todo o processo e do domínio como um todo. Após a captura da conceitualização do domínio, deve-se pensar na codificação. Contudo, antes disso, é necessário estabelecer uma meta-ontologia, especificando termos, conceitos, símbolos e relações utilizadas para representar a ontologia. A linguagem de representação a ser escolhida deve ser compatível com a meta-ontologia especificada. Daí, então, pode-se iniciar a codificação. A última atividade desta etapa é a integração com outras ontologias já existentes. Este não é um problema de fácil resolução, pois se já é difícil chegar a um senso comum num dado domínio, mais complexo e trabalhoso será quando forem agregados outros domínios.

Refinamento: para chegar à ontologia desejada (mais aderente ao domínio de conhecimento em questão), serão necessárias *n* iterações. Nesta atividade, é de suma

importância a participação dos usuários, o que intensifica a necessidade de uma linguagem de representação conhecida por todos os envolvidos. Na primeira iteração da atividade de construção, obtém-se uma ontologia de alto nível e, a cada nova iteração, a ontologia obtida se aproxima dos casos particulares do domínio tratado.

Avaliação: esta etapa consiste em verificar se a ontologia obtida a partir da etapa anterior está de acordo com o projetado durante a etapa de definição de escopo, ou seja, deve-se confirmar se a ontologia pode responder a todas as questões de competências estipuladas, se representa fielmente o domínio, se é facilmente compreendida por todos os *stakeholders*, se é independente da linguagem de codificação utilizada, se é passível de *updates* e, finalmente, se possui pouco compromisso ontológico dissociando-a da base de conhecimento. O ideal é estipular antecipadamente os critérios sobre os quais se quer avaliar a ontologia. A abrangência de uma ontologia é percebida através de suas características referentes à clareza, coerência, extensibilidade, independência de codificação e compromissos ontológicos. Pode-se concluir, então, que o melhor método para construir uma ontologia irá depender do grau de formalidade requerida, que por sua vez depende do propósito de tal ontologia.

Documentação: dependendo do tipo e propósito da ontologia, devem ser estabelecidas regras diferenciadas para documentá-la. A falta de uma documentação formal e compreensível pode dificultar o compartilhamento do conhecimento, objetivo básico da ontologia.

A experiência com o uso de ontologias no desenvolvimento de SBCs mostra que o aspecto motivacional é o maior obstáculo para o sucesso do sistema, seguido pela comunicação, aspecto negligenciado em boa parte dos casos.

CONCLUSÃO

Foi mostrado, neste trabalho, que devido ao emprego de uma sintaxe e semântica formal, para a construção de ontologias, definindo conceitos, propriedades e relações deste domínio, possibilita a formalização de requisitos complexos, gerando uma especificação de *software* consistente, de modo a eliminar ambigüidades e consistências.

Sem uma conceitualização compartilhada de um domínio de conhecimento, não é possível representá-lo, e a falta de representação impossibilita a padronização e o reuso do conhecimento. Entretanto, deve-se salientar que é impossível representar o mundo real, ou mesmo uma parte dele, de forma minuciosa e detalhada. Por isso, para representar um domínio, é desnecessário ater-se a detalhes muito específicos, pois tentar armazenar num repositório todo o conhecimento de especialistas, pode transformar este repositório em um

caos informatizado. Neste contexto, pode-se concluir que a simplicidade é benéfica e almejada na construção de ontologias, mesmo porque, qualquer representação de um domínio é apenas uma aproximação deste. Por isso, para representar a conceituação de um domínio, é necessário concentrar a atenção em um número limitado de conceitos, suficientes e relevantes, para se criar uma abstração do fenômeno em estudo.

Hoje, a base de conhecimento construída possui pouco compartilhamento e reuso, uma vez que a máquina de inferência tinha foco genérico e a estratégia para resolver um problema era embutida como parte da base de conhecimento. No futuro, os sistemas inteligentes poderão dispor de bases unificadas de ontologias, e serão construídos utilizando-se componentes extraídos destas bases.

BIBLIOGRAFIA E REFERÊNCIAS¹⁶

- ALMEIDA, M. BAX, M. **Uma visão geral sobre ontologias**: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. Ciência da Informação, Brasília, v. 32, n. 3, p. 7-20, 2003.
- BONIFACIO, A. **Ontologias e Consultas Semânticas**: uma aplicação ao caso Lattes. 2002. 85 f. Dissertação (Programa de Pós-graduação em Computação) – Universidade Federal do Rio Grande do Sul, Instituto de Informática, Porto Alegre.
- BREITMAN, K., LEITE, J. **Ontologias**: como e porquê criá-las. XXIV Congresso da Sociedade Brasileira de Computação – SBC' 2004. Salvador, 2004.
- BUSSLER, C., FENSEL, D., MAEDCHE, A. **Conceptual Architecture for Semantic Web Enabled Web Services**. SIGMOD Record, v. 31, n. 4, dez. 2002.
- CHANDRASEKARAN, B., JOSEPHSON, J., BENJAMINS, V. **Ontology of Tasks and Methods**. In: *Proceedings of KAW'98, Eleventh Workshop on Knowledge Acquisition, Modeling and Management*, Inn, Banff, Alberta, Canada, abr. 1998.
- CORCHO *et al.* **Building legal ontologies with METHONDOLOGY and WebODE**. *Law and the Semantic Web. Legal Ontologies, Methodologies, Legal Information Retrieval, and Applications*. Mar. 2005. Springer-Verlag, LNAI 3369.
- CORCHO, O., GOMEZ-PEREZ, A. **Guidelines to Study Differences in Expressiveness between Ontology Specification Languages**. *12th Workshop on Knowledge Acquisition, Modeling and Management (KAW 1999)*. Banff, Canada. Nov. 1999.
- DOMINGUE, J. **Tadzebao and WebOnto**: discussing, browsing and editing ontologies on the web. KAW'98. Banff, Canada, 1998.
- FALBO, R. **Integração do conhecimento em um ambiente de desenvolvimento de software**. 1998. 205 f. Tese (Doutorado em Engenharia de Sistemas e Computação). COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- FERNANDEZ-LOPEZ, M., *et al.* **Building a Chemical Ontology using Methondology and the Ontology Design Environment**. *IEEE Intelligent Systems*, v. 14, n. 1, *Special Issue on Ontologies*, p. 37-46, jan./fev. 1999.

¹⁶ Baseada na NBR-6023/2002, da Associação Brasileira de Normas Técnicas (ABNT).

- FERNANDEZ-LOPEZ, M., GOMEZ-PEREZ, A., JURISTO, N. *METHONDOLOGY: from ontological art towards ontological engineering*. Workshop on Ontological Engineering, Spring Symposium Series, AAAI97, Stanford, 1997.
- FONSECA, F., EGENHOFER, M. **Sistemas de Informação Geográficos Baseados em Ontologias**; Informática Pública 1 (2): 47-65, 1999.
- GRUBER, T. *Ontolingua: a mechanism to support portable ontologies*. 1992.
- _____. *Toward Principles for the Design of Ontologies Used for Knowledge Sharing*. In . 1993.
- GRUNINGER, M. FOX, M. *Methodology for Design and Evaluation of Ontologies*. Workshop on Basic Ontological Issues in Knowledge Sharing. International Joint Conference on Artificial Intelligence, 1995.
- GRUNINGER, M., LEE, J. *Ontology: applications and design*. Communications of the ACM. fev. 2002, v. 35, n. 2. p. 39-41.
- GUARINO, N. *The Ontological Level*. In: R. Casi, B. Smith and G. White (Eds.), *Philosophy and the Cognitive Science*. Viena: Holder-Pichler-Tempsky, 1994.
- _____. *Understanding, Building, and Using Ontologies*. International Journal of Human-Computer Studies. v. 46, issue 2-3, fev./mar. 1997, p. 293-310. Duluth: Academic Press, Inc., 1997.
- GUARINO, N., WELTY, C. *A Formal Ontology of Properties*. 2000.
- GUHA, R., LENAT, D. *Cyc: A midterm report*. AI Magazine, 1990.
- GUIMARAES, F. *Utilização de Ontologias no domínio B2C*. PUC-RJ, 2002.
- HOLSAPPLE, C., JOSHI, K. *A collaborative approach to Ontology Design*. Communications of the ACM. fev. 2000, v. 45, n. 2. 2002.
- INFO-EDUCACAO. *Informática na Educação: Teoria & Prática*. UFRGS, v. 3, n. 1, Porto Alegre, set. 2000.
- OLIVEIRA, K. **Modelo para construção de ambientes de desenvolvimento de software orientados a domínio**. 1999. 222 f. Tese (Doutorado em Engenharia de Sistemas e Computação). COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- PACHECO, R., KERN, V. **Uma Ontologia comum para a integração de bases de informações e conhecimento sobre ciência e tecnologia**. Ciência da Informação. Brasília, v. 30, n. 3, p. 56-63, set./dez. 2001.
- SCHREIBER, A. et al. *CML: the CommonKADS conceptual modeling language*. In: Steels et al. (eds): *A Future of Knowledge Acquisition, Proc. 8th European Knowledge Acquisition Workshop (EKAW94)*, Hoegaarden, Lecture Notes in Artificial Intelligence 867, Springer-Verlag, 1994.
- SOWA, J., DIETZ, D. *Knowledge Representation: logical, philosophical, and computational foundations*. Brooks Cole, 1. ed. ago. 1999.
- STAAB, Y. et al. *Knowledge Processes and Ontologies*. Intelligent Systems. IEEE, v. 16, n. 1, p. 26-34, jan./fev. 2001.
- SURE, Y., STAAB, S., STUDER, R. *Methodology for Development and Employment of Ontology based Knowledge Management Applications*. SIGMOD Record, v. 31, n. 4, p. 18-23, dez. 2002.
- USCHOLD, M. *Building Ontologies: towards a unified methodology*. Expert Systems '96, The 16th Annual Conference of the British Computer Society Specialist Group on Expert Systems. Cambridge, 1996.
- USCHOLD, M., GRUNINGER, M. *Ontologies: principles, methods and applications*. Knowledge Engineering Review, 1996.