

ONTOLOGIA: o termo e a idéia

ONTOLOGY: the term and the idea

Marcelo Schiessl - schiessl@ig.com.br
Universidade de Brasília

Resumo

Aborda o conceito de Ontologia e relaciona alguns problemas que motivam os pesquisadores desta área. Contextualiza a sobrecarga informacional em virtude do crescente volume de texto eletrônico armazenado em bases de dados, em especial na Web, e sua implicação na manutenção do fluxo da informação. Além disso, introduz a idéia da Web Semântica como uma solução para a organização dos conteúdos das páginas nos sítios da Web e o papel preponderante das ontologias, dentro deste contexto, que auxiliam a automatização da interpretação destes conteúdos e otimiza a recuperação da informação.

Palavras-chave: Ontologia. Web semântica. Recuperação de informação. Representação do conhecimento.

1 INTRODUÇÃO

Desde a invenção de Gutenberg^[1], o homem começou a acumular informações em escala que poucos poderiam imaginar. Da impressão da Bíblia às pesquisas científicas, fontes preciosas de conhecimento potencial foram sendo armazenadas por séculos.

Com a passagem da era industrial à era da informação, o foco de profissionais vem se deslocando para o tratamento de mensagens e signos. Como consequência, a informação adquire valor de um produto, um bem, que pode ser comercializado conforme a necessidade do seu uso.

Na última década, em decorrência dos avanços tecnológicos e da popularização da Internet, vimos a explosão informacional e a passagem da cultura impressa à cultura eletrônica. Assim, a criação e distribuição da informação ocorrem em qualquer lugar do planeta e a qualquer hora.

Considerando a Internet como uma biblioteca digital globalizada, uma pessoa divulga livremente qualquer informação, verdadeira ou não, sem que haja controle sobre seu conteúdo, nem a estruturação e a organização necessárias para sua recuperação.

Neste contexto, foi proposta a Web^[2] Semântica, concebida por Berners-Lee, considerado o Pai da Web, que se dispõe a melhorar e padronizar sistemas e linguagens utilizados na Web com o objetivo de torná-la um Sistema de Informação mais eficiente.

No entanto, para que se utilize desta informação de maneira eficiente é necessária a adequação de computadores e de artefatos computadorizados para que eles possam “entender” o conteúdo dos documentos eletrônicos na rede.

As ontologias, apesar da controvérsia em torno de seu conceito, automatizam a interpretação de significado da informação contida nos documentos. Além disso, espera-se que as ontologias forneçam definições ao vocabulário utilizado para representar um dado conhecimento.

No decorrer deste texto, abordam-se assuntos relacionados à Ontologia, quais sejam sua história, conceitos, características e aplicações.

2 CONTEXTO

Em 1989, coordenado por Tim Berners-Lee, no laboratório do CERN^[3], originou-se a *World Wide Web*^[4], *www* ou simplesmente Web, que se propunha ao compartilhamento de informações contidas em documentos, baseados em hipertextos, por uma rede de computadores.

A Web tornou-se uma das invenções mais extraordinárias dos últimos tempos que revolucionou a maneira como lidamos com a informação.

Vale a pena ressaltar que a Internet é a rede mundial de computadores e a Web (substantivo próprio) é o sistema de informação hipermídia que utiliza a Internet como meio de transmissão.

Dessa forma, a informação textual, sonora e imagética pode estar situada, em formato eletrônico, no ciberespaço, que pode ser em qualquer lugar do mundo, e se acumula em crescentes volumes de informação através de portais e sites e que atendem, cada vez mais, a um maior número de pessoas. Entretanto, por mais paradoxal que possa parecer, o aumento destes estoques informacionais não garante o acesso à informação na velocidade, qualidade e precisão que o usuário deseja.

Muitos conteúdos das páginas da Web são textos. Eles são construídos com a utilização de palavras que podem ter significados distintos em contextos diferentes. Textos são elaborados para pessoas, não máquinas. Entretanto, para a recuperação de informação na Web, de forma eficiente, necessita-se que esta tarefa seja feita de maneira automática, ou seja, contando com o auxílio de computadores. Assim, é necessário que aparatos computacionais sejam capazes de tratar as ambigüidades inerentes às palavras que são utilizadas em consultas e buscas de informação na Web.

Nesse contexto, a Web Semântica procura estruturar o significado dos conteúdos das páginas da Web, criando um ambiente onde programas agentes, que funcionam sem o controle direto ou supervisão permanente, vagam de página em página com o objetivo de realizar as demandas mais sofisticadas de usuários (BERNERS-LEE; HENDLER; LASSILA, 2001). Por conseguinte, a Web Semântica não é uma separação da Web, mas uma extensão dela, na qual a informação tem um significado semântico bem definido que possibilita pessoas e máquinas trabalharem cooperativamente.

Dessa forma, computadores devem ter acesso a coleções de informações e conjuntos de regras de inferência para que possam ser usados na condução de “raciocínio” automático.

Para automatizar a interpretação de significado das informações, os computadores utilizarão conjuntos de afirmativas que definem as relações entre conceitos e estabelecem regras lógicas de raciocínio sobre eles.

3 ONTOLOGIA: apropriação do termo filosófico

Na Filosofia, o termo Ontologia possui sua origem na Metafísica, segundo Aristóteles é a Filosofia Primeira que trata do estudo do ser enquanto ser. Apropriando-se da obra de Chauí (2003), a palavra ontologia é formada por outras duas: onto que significa “o Ser” e logia, “estudo ou conhecimento”. Assim, Ontologia significa “estudo ou conhecimento do Ser, dos entes ou das coisas tais como são em si mesmas, real e verdadeiramente”.

Porém, outras áreas do conhecimento tomaram emprestado o termo Ontologia e o sentido original na Filosofia, certamente, foi mudado. Para Gruber (2002), ontologia é uma especificação de conceitos no contexto do compartilhamento de conhecimento.

Segundo Castels (2002), o campo da inteligência artificial foi construído com a noção de processo cognitivo artificial, ou seja, uma representação da realidade de acordo com a percepção humana. Portanto, ontologia é a maneira que se divide a realidade em partes menores com o objetivo de entendê-la e processá-la.

Além disso, espera-se que uma ontologia forneça definições de vocabulário específico para representar uma dada realidade, bem como suas relações (CORAZZON, 2002). Nesse sentido, entende-se como a normalização de conceitos e o estabelecimento de relações entre eles.

A idéia da ontologia como um mecanismo de especificação é um elemento chave na construção da Web semântica, pois a delimitação de um dado conceito e a definição das relações entre eles auxiliam a máquina deduzir o significado da informação contida nos documentos.

4 REPRESENTAÇÃO DO CONHECIMENTO?

Segundo Alvarenga (2003), representar refere-se ao “ato de colocar algo no lugar de” e, ainda, que representação do conhecimento é a expressão simbólica de observações metódicas da natureza e dos fatos sociais, fazendo uso de linguagens disponíveis no contexto da comunicação e produção do conhecimento, ou melhor, são registros do pensamento em suporte documental, utilizando-se dos sentidos, da emoção, da razão e da linguagem.

As ontologias são planejadas e construídas pelo homem para atender suas necessidades de informação em relação à organização, à estruturação e à recuperação. Dessa forma, elas são parte integrante de um sistema formal que de alguma maneira necessita de uma forma de representação.

De acordo com Lima-Marques e Modesto (2004), espera-se que as ontologias forneçam definições ao vocabulário utilizado para representar um dado conhecimento.

Conforme mencionado, uma ontologia é uma especificação explícita de conceituação que representa uma visão simplificada da realidade e para a Filosofia, uma explicação sistemática da existência. Logo, para sistemas que utilizam processos cognitivos artificiais, o que existe é aquilo que pode ser representado.

Assim, entende-se que a Ontologia seja uma forma de Representação do Conhecimento, pois traduz a realidade para um modelo bem delimitado e específico daquilo que se pretende representar.

Importante destacar que a restrição dessa modelagem é que a única representação de um objeto perfeitamente precisa é o próprio objeto (DAVIS; SHROBE; SZOLOVITS, 1993) e, portanto, a ontologia é uma correspondência, limitada à visão de mundo do seu elaborador, entre o modelo e seu referente no mundo e, inevitavelmente, possui simplificações e conseqüentes imperfeições na representação da realidade.

4.1 Para que?

Os recentes avanços da tecnologia proporcionaram desenvolvimento de ontologias que tem se transferido dos laboratórios de Inteligência Artificial para outras áreas de domínio especializado. As ontologias na Web vão de grandes taxonomias categorizando sítios na Web, como é o caso do *Yahoo!*, até a categorização de produtos para vendas, Livraria Virtual Saraiva por exemplo.

Uma ontologia define um vocabulário comum para pesquisadores que necessitam compartilhar informações em um domínio. Isto inclui definições de máquinas capazes de interpretar conceitos dentro de um domínio e a relação entre eles (NOY; MCGUINNESS, 2001). Portanto, o compartilhamento do significado comum das estruturas de informação entre pessoas e programas agentes permite que estes programas consigam responder às perguntas feitas pelos usuários em um sistema de informação.

O benefício imediato verificado com esse compartilhamento pode ser observado na melhoria da comunicação entre pessoas e organizações, além da integração de diferentes perspectivas de usuários.

Além disso, a reutilização de ontologias, de domínio mais específico, facilita o desenvolvimento de outras, pois poderão ser simplesmente agregadas para descrever o domínio de interesse mais amplo.

A interoperabilidade de sistemas pode ser auxiliada com a utilização de ontologias com o objetivo de unificar o vocabulário e viabilizar a transferência mútua de informações entre diferentes sistemas.

A vantagem da utilização de ontologia na Web é evidente, principalmente se considerarmos a dificuldade existente na tarefa de recuperação da informação desejada.

4.2 Propriedades desejáveis

Uma ontologia é projetada levando-se em conta os propósitos da área de atuação e aplicação e, enfatizado por Gruber (1993), algumas características relevantes devem ser consideradas no seu planejamento, como se segue:

. Clareza: as definições devem ser objetivas e, quando viável, completas. É preferível que uma definição seja declarada através de axiomas^[5] lógicos. Todas as definições devem ser documentadas em linguagem natural.

. Coerência: uma ontologia deve permitir inferências que sejam consistentes com as definições.

. Extensível: uma ontologia deve permitir que novos termos sejam definidos, para usos especiais, baseados no vocabulário existente, de maneira que não seja necessária a revisão das definições já existentes.

. Mínimo viés de codificação: a conceituação deve ser especificada no âmbito do conhecimento, ou seja, a escolha da representação não deve se basear somente na conveniência da notação ou implementação.

. Mínimo compromisso ontológico: uma ontologia deve necessitar de um compromisso ontológico apenas o suficiente para permitir que as atividades de compartilhamento do conhecimento esperadas aconteçam.

5 METODOLOGIA: ORIENTAÇÃO NA ELABORAÇÃO DE ONTOLOGIAS

Como foi visto, uma ontologia possui características que devem ser desejadas na sua elaboração e, também, lembrar que ela é um modelo de realidade do mundo e seus conceitos devem refleti-la.

Assim, a primeira constatação é de que não há uma maneira melhor que outra na construção de uma ontologia. Por conseguinte, na tentativa de representar um modelo do mundo, que não é estático, o desenvolvimento de uma ontologia é um processo iterativo para que seus conceitos se aproximem o máximo possível dos objetos que eles representam, bem como as relações entre eles.

Desse modo, baseado em Noy e Mcguinness, (2001), sugere-se uma orientação para construção de uma ontologia.

Em primeiro lugar, identifica-se a competência da ontologia, ou seja, sua intenção, sua importância, sua abrangência, seus usuários, os panoramas de utilização e seus fundamentos.

Em seguida, com base na competência identificada para a ontologia, procede-se à captura da conceituação (conceitos e relações) do universo de discurso.

Além disso, busca-se alguma linguagem de representação com o objetivo de trazer um nível de formalização da ontologia.

Com o objetivo de aproveitar as ontologias já elaboradas procede-se a integração com outras anteriormente elaboradas, pois se busca integrar a ontologia em construção com outras já existentes para re-aproveitar conceituações.

Por último, executa-se a avaliação que verifica a ontologia construída ou em construção

para validação e, caso se identifique alguma inconsistência, realiza-se o ajuste necessário iterativamente até que ela seja validada.

Finalmente, desenvolve-se a documentação que é a atividade que registra as informações relevantes à sua identificação, construção e requisitos em documentos formais.

6 ABERTURAS E BARREIRAS

Entendendo a informação como elemento essencial à sociedade e, ainda, um fenômeno social, são notáveis as vantagens da utilização de ontologias no sentido de tornar disponível, de maneira eficiente, o conhecimento desejado. Entretanto, a cultura, a língua, a política dentre outras causas são muros erguidos contra a proliferação das ontologias e conseqüente restrição a sua utilização como veremos.

Com o propósito de facilitar o compartilhamento do conhecimento de maneira estruturada, as ontologias facilitam a comunicação, pois são instrumentos de grande valia para ajudar as pessoas a comunicar, de variadas formas, acerca de um domínio específico do conhecimento.

Ressalta-se, também, que em virtude do caráter formal da notação usada para a representação, a normalização do domínio pode eliminar incoerências envolvendo as ambigüidades inerentes à linguagem, a palavra.

Decorrente do processo de desambiguação, as ontologias estabelecem um vocabulário comum, em conformidade com o domínio, e representa o conhecimento específico de maneira explícita e em elevado nível de abstração, possuindo um potencial admirável de reutilização.

Há que se ressaltar, porém, que a escolha de uma ontologia é um processo político, uma vez que uma ontologia não abrange todos os domínios do conhecimento. Dessa maneira, o processo de escolha de uma determinada ontologia é orientada para a adequação da necessidade dos indivíduos ou grupos.

Por tratar de linguagem, que é um ente vivo em constante transformação e evolução, as ontologias devem estar em contínuo processo de manutenção para que reflitam com maior fidelidade o domínio que representam.

Além disso, as ontologias são desenvolvidas com o propósito bem definido de representar um vocabulário ou uma visão de mundo de indivíduos ou grupos e esta representação pode ser contraditória em relação a outras. Portanto, nesse contexto, o caráter extensível e a reutilização das ontologias podem sofrer sérias restrições.

7 CONCLUSÃO

A tecnologia possibilita a criação de enormes bases de dados, em meio eletrônico, distribuídas mundo afora.

A Internet proporciona que essa informação espalhada pelo mundo esteja a apenas alguns cliques no mouse de um computador pessoal conectado à rede mundial.

Porém, o caráter individual, a ambigüidade da linguagem e a ausência de normas regulamentares para a publicação de informações na Web provocam uma dificuldade na sua recuperação.

A Web Semântica foi proposta com o objetivo de estruturar os conteúdos das páginas na Web possibilitando a recuperação da informação levando em consideração o seu aspecto semântico.

Nesse contexto, a Ontologia tem um papel preponderante na normalização de conceitos e nas relações semânticas entre eles.

O histórico do termo Ontologia, mencionado primeiramente na Filosofia, tem origem na Metafísica, que é conceituado como o estudo do ser enquanto ser. Contudo, junto com a evolução do homem e do seu conhecimento, a palavra Ontologia teve o seu conceito original alterado à medida que foi emprestada a outras áreas do conhecimento.

Atualmente, no contexto da Inteligência Artificial, não se trata somente da explicação sistemática da existência, mas da especificação explícita de conceitos de um determinado domínio do conhecimento podendo ser portátil para atender às necessidades de indivíduos ou grupos.

Dessa maneira, a aplicação das ontologias, no universo da Web, contribui para que a “interpretação” automática do conteúdo semântico de páginas dos sítios na Internet seja possível e, conseqüentemente, a colaboração entre máquinas e homens seja reforçada.

Essas considerações apenas tangem um assunto que, tenho certeza, tem tanto a evoluir, quanto a contribuir para o desenvolvimento da Ciência da Informação. Não se pretende esgotar o assunto, apenas introduzir o assunto no contexto da Web e, especialmente, afirmar sua importância na gestão da informação.

Muito se avançou para que a Ontologia pudesse ser efetiva na função de Representação de Conhecimento, porém há muitas barreiras a serem transpostas já que o objetivo principal é atender um ser social, o homem, que inerente à sua natureza traz a ambigüidade de culturas, a língua e processos políticos que devem ser levados em consideração.

REFERÊNCIAS

ALVARENGA, L. Representação do conhecimento na perspectiva da ciência da informação em tempo e espaços digitais. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, n. 15, jan./jun 2003.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic WEB. **Scientific American**, p. 5-7, may 2001.

CASTEL, F. Ontological computing. **Communications of the ACM**, v. 45, n. 2, p. 29-30, 2002.

CHAUÍ, M. **Convite à Filosofia**. 13. ed. São Paulo: Ática, 2003.

CORAZZON, R. **Ontology: a resource guide for philosophers**. Itália, 2002. Disponível em: < <http://www.formalontology.it/> >. Acesso em: fev 2005.

DAVIS, R.; SHROBE, H.; SZOLOVITS, P. What is a knowledge representation? **AI Magazine**, v. 14, n. 1, p. 17-33, 1993. Disponível em: < <http://groups.csail.mit.edu/medg/ftp/psz/k-rep.html> > . Acesso em: fev 2005.

GRUBER, T. R. **Toward principles for the design of ontologies used for knowledge sharing**. USA, ago 1993. Technical report KSL 93-04.

GRUBER, T. R. **What is an Ontology?** USA, ago 2002. Disponível em: < <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html> > . Acesso em: fev 2005.

LIMA-MARQUES, M; MODESTO, A.; **Ontologias**: contribuição à arquitetura da informação. Artigo distribuído em sala de aula no curso de pós-graduação, Departamento de Ciência da Informação e Documentação, UnB, 2004.

NOY, Natalya F. ; McGUINNESS, Deborah L. **Ontology Development 101: A Guide to Creating Your First Ontology**. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001. Disponível em: < <http://www.ksl.stanford.edu/people/dlm/papers/ontology101/ontology101-noy-mcguinness.html> > . Acesso em: fev 2005.

[1] Inventor da prensa

[2] nome pelo qual a rede mundial de computadores Internet se tornou conhecida a partir de 1991, quando se popularizou devido à criação de uma interface gráfica que facilitou o acesso e estendeu seu alcance ao público em geral (HOUAISS, Antônio. **Dicionário – Houaiss de língua Portuguesa**. Rio de Janeiro: Objetiva, 2001)

[3] Organização Européia para Pesquisa Nuclear que fica em Genebra, na Suíça..

[4] Literalmente traduzido como a “Teia do tamanho do mundo”.

[5] premissa considerada necessariamente evidente e verdadeira, fundamento de uma demonstração, porém ela mesma indemonstrável, originada, segundo a tradição racionalista, de princípios inatos da consciência ou, segundo os empiristas, de generalizações da observação empírica. (HOUAISS, Antônio. **Dicionário – Houaiss de língua Portuguesa**. Rio de Janeiro: Objetiva, 2001.)

ABSTRACT

The paper presents the concept of Ontology, and lists some problems that challenge researchers in this area. It contextualizes information overload as a result of the increasingly amount of electronic texts in storage, mainly in the Web, and the implications to keep up with the information cycle. It presents an introduction to Semantic Web Idea in order to organize the content of web sites, and the preponderant role of Ontology that helps the automated interpretation of these contents and optimizes the information retrieval.

KEYWORDS: Ontology. Semantic web. Information retrieval. Knowledge representation.

Originals recebidos em: 10/06/2007

Texto aprovado em: 14/09/2007