

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/310018091>

O cenário da avaliação de ontologias: revisão de literatura

Article · August 2016

CITATIONS

0

READS

272

2 authors:



[Webert Araújo](#)

Federal University of Minas Gerais

12 PUBLICATIONS 4 CITATIONS

[SEE PROFILE](#)



[Gercina Angela de Lima](#)

Federal University of Minas Gerais

142 PUBLICATIONS 558 CITATIONS

[SEE PROFILE](#)

O CENÁRIO DA AVALIAÇÃO DE ONTOLOGIAS: REVISÃO DE LITERATURA

THE ONTOLOGY EVALUATION SCENARIO: A LITERATURE REVIEW

Webert Junio Araújo

Gercina Ângela B. de Oliveira Lima

Resumo: As ontologias são sistemas de organização do conhecimento que têm ganhado destaque nos estudos da Ciência da Informação devido ao potencial que possuem na representação e compartilhamento de estruturas do conhecimento em diferentes domínios. Porém, é preciso garantir a qualidade das ontologias por meio da avaliação de ontologias. O estudo da avaliação de ontologias se deve à necessidade de verificar a eficiência delas na representação do conhecimento e na realização de outras tarefas para as quais foram desenvolvidas. Diante desse cenário, este artigo apresenta uma revisão de literatura sobre avaliação de ontologias com o propósito de identificar o que já foi discutido sobre a temática, o andamento das pesquisas e o que é necessário para o desenvolvimento da área. Constatou-se por meio da revisão de literatura que existem muitas propostas, hipóteses e alguns experimentos na área de avaliação de ontologias, porém, pouca coisa está consolidada. Assim, percebe-se que a área ainda está em aberto e são necessários estudos para seu desenvolvimento.

Palavras-chave: Sistema de Organização do Conhecimento. Ontologia. Avaliação de ontologia.

Abstract: Ontologies are Knowledge Organization Systems that have gained prominence in studies of Information Science because of their potential in the representation and sharing of knowledge structures in different domains. However, we must ensure the quality of ontologies through their evaluation. The study of ontology evaluation is due to the need to check their efficiency in knowledge representation and performing other tasks for which it was developed. In this scenario, this article presents a literature review on ontology evaluation in order to identify what has been discussed on the topic, the progress of research and what is needed for the development of the area. It was found through the literature review that there are many proposals, some hypotheses and experiments in the field of ontology evaluation, however, almost nothing is consolidated. Thus, it is clear that the area is still open and further studies are required for their development.

Keywords: Knowledge Organization System. Ontology. Ontology evaluation.

1 INTRODUÇÃO

Desde 1995, pesquisadores identificaram a necessidade de estudos sobre avaliação de ontologias. Devido aos diversos usos que se pode fazer das ontologias e aplicações onde elas podem ser usadas, é necessário que elas sejam bem desenvolvidas para que se possa garantir a qualidade das mesmas.

O estudo da avaliação de ontologias se deve à necessidade de verificar a eficiência delas na representação do conhecimento e na realização de outras tarefas para as quais foram desenvolvidas, uma vez que, com o desenvolvimento de ontologias em várias áreas do conhecimento e sem uma metodologia padrão que orientasse a construção delas, logo surgiram ontologias ineficientes, que não representavam de maneira adequada o conhecimento a que se propunham representar.

Mesmo as metodologias e métodos para construção de ontologias mais citadas na literatura técnica, como a TOVE¹, Methontology², NeON³, Método 101⁴, Método Sensus⁵ não apresentam um consenso com relação às etapas para construção de ontologias, sendo

¹TOVE (Toronto Virtual Enterprise Ontology Project): Desenvolvida por Gruninger e Fox (1995).

²Methontology: Desenvolvida no laboratório de Inteligência Artificial da Universidade Politécnica de Madri pelos pesquisadores Gómez-Pérez, Fernandez e Vicente (1996); Fernandez, Gómez-Pérez e Juristo (1997).

³NeON: Metodologia desenvolvida por Suárez-Figueroa et al (2008)

⁴Método 101: Desenvolvido por Noy e McGuinness (2001).

⁵Método Sensus: Desenvolvido pelo grupo de linguagem natural – Information Sciences Institute da Universidade da Carolina do Sul

direcionadas a diferentes propósitos e aplicações. A literatura mostra que ainda não existem modelos unificados para o desenvolvimento de ontologias. Então, com tal deficiência, a temática avaliação de ontologias vem como uma solução, pois tem o propósito de analisar a ontologia e identificar se ela está representando de maneira adequada o modelo a que se propõe representar.

Diante desse cenário, este artigo apresenta uma revisão de literatura sobre avaliação de ontologias com o propósito de identificar o que já foi discutido sobre a temática, o andamento das pesquisas e o que é necessário para o desenvolvimento da área.

2 ONTOLOGIAS

O termo 'ontologia' vem do grego: *onto* (ser) e *logos* (palavra) e, de acordo com Moreira et al (2004), foi cunhado por Rudolf Goclenius em seu *Lexicon philosophicum* e Jacob Lorhard em seu *Theatrum philosophicum*, em 1613. Na literatura, há várias definições para o termo; algumas são consensuais e apenas se completam e outras seguem linhas diferentes. Sob o ponto de vista da Filosofia, a ontologia é uma teoria sobre a natureza da existência.

No contexto da Ciência da Informação (CI), Vickery (1997) foi um dos pioneiros neste campo a dar atenção ao assunto ontologia na organização e representação do conhecimento. Logo depois, outros pesquisadores da área começaram a estudar o tema, como Soergel (1999), Gilchrist (2003). Os estudos da temática ainda são incipientes na área, mas têm evoluído nos últimos anos, dada a potencialidade das ontologias para contribuir com a CI.

Soergel (1999) diz que "ontologia" é uma classificação de categorias básicas. O autor faz ainda uma crítica à falta de comunicação entre as áreas de conhecimento, pois, de acordo com ele, o termo "ontologia" é apenas um novo nome para "classificação", termo que já é usado pela Biblioteconomia e Ciência da Informação (BCI) há muito tempo. Assim, o autor sugere que os estudos sobre ontologias possam se basear em técnicas já consolidadas no contexto da BCI.

Segundo Almeida e Bax (2003), uma ontologia pode ser concebida como uma linguagem, pois estabelece um conjunto de termos que, posteriormente, poderão ser usados para formular consultas sobre uma base de conhecimento. Nessa mesma linha, Furgeri (2006) define ontologia como uma técnica que permite organizar a informação e

auxiliar sua recuperação, através de um vocabulário de termos categorizados que descrevem determinada realidade.

O objetivo da ontologia, de acordo com Souza e Alvarenga (2004) é “criar um vocabulário controlado para se trocar informações entre os membros de uma comunidade, sejam eles humanos ou agentes inteligentes” (SOUZA; ALVARENGA, 2004, p.137). A ideia da ontologia funcionar como uma linguagem comum para um determinado domínio e assim possibilitar o compartilhamento do conhecimento é bastante citada na literatura. Gómez-Pérez e Pazos (1995), por exemplo, dizem que as ontologias são plataformas que possibilitam o reuso do conhecimento através do estabelecimento de vocabulários comuns e interpretações semânticas dos termos. Nessa mesma linha, Uschold e Gruninger (1996) dizem que uma ontologia padroniza a linguagem de um domínio e facilita a comunicação entre os usuários, através de um vocabulário compartilhado do domínio.

Sob o ponto de vista da Ciência da Computação (CC), o termo ‘ontologia’ surgiu em 1967 em um trabalho sobre modelagem de dados de G. H. Mealy (SMITH, 2004). A ontologia é definida no contexto da CC como um artefato de engenharia, composto de um vocabulário de termos organizados em uma taxonomia, as definições dos termos e um conjunto de axiomas formais usados para criar relações e limitar as interpretações (NOY; HAFNER, 1997). Sob esse enfoque, Guarino (1998) define ontologia como

[...] um artefato constituído por um vocabulário usado para descrever uma certa realidade, mais um conjunto de fatos explícitos e aceitos que dizem respeito ao sentido pretendido para as palavras do vocabulário. Este conjunto de fatos tem a forma da teoria da lógica de primeira ordem, onde as palavras do vocabulário aparecem como predicados unários ou binários (GUARINO, 1998, p.4, tradução própria⁶).

Gruber (1993, p.1) traz uma das principais definições do termo; ele diz que “ontologia é uma especificação formal, explícita e compartilhada de uma conceitualização.” A palavra *formal*, apontada por Gruber, quer dizer que a ontologia deve possuir um nível de formalidade que possa ser compreendida pelas máquinas; já a palavra *explícita* significa que todos os recursos empregados na ontologia devem ser claros; *compartilhada* quer dizer que aquele conhecimento trazido pela ontologia deve ser consensual, deve ser aceito por um

⁶[...] an ontology refers to an engineering artifact, constituted by a specific vocabulary used to describe a certain reality, plus a set of explicit assumptions regarding the intended meaning of the vocabulary words. This set of assumptions has usually the form of a first-order logical theory, where vocabulary words appear as unary or binary predicate names (GUARINO, 1998, p. 4).

grupo de pessoas e não apenas por alguns indivíduos; quanto à palavra *conceitualização*, trata-se de um modelo abstrato de algum fenômeno do mundo através da identificação dos conceitos relevantes daquele fenômeno (STUDER et al., 1998).

Para Chandrasekaram et al (1999), uma ontologia é a representação de um vocabulário, geralmente especializado de algum domínio. De modo mais preciso, não é o próprio vocabulário que se qualifica como uma ontologia e sim a conceitualização que os termos presentes no vocabulário pretendem capturar. Percebe-se nessa definição que a relevância de uma ontologia é esclarecer a estrutura de conhecimento e não apenas listar termos de certo domínio.

3 AVALIAÇÃO DE ONTOLOGIAS

Os estudos sobre avaliação de ontologias iniciaram por volta de meados da década de 1990. Grande parte dos estudos está relacionada à área da Ciência da Computação, porém, identificam-se alguns estudos no âmbito da Ciência da Informação.

Gómez-Pérez (2001), um dos pioneiros no estudo de avaliação de ontologia, apresenta cinco critérios pelos quais uma ontologia pode ser avaliada. São eles: (1) coerência, que refere-se a saber se é possível obter conclusões contraditórias a partir de definições explícitas na ontologia; (2) completude, que visa identificar se a ontologia não possui definições incompletas; (3) concisão, quando não há definições desnecessárias e inúteis; (4) capacidade de expansão, refere-se ao esforço exigido em adicionar novas definições ou mais conhecimento para as suas definições, sem alterar o conjunto de propriedades que já estão definidos; e (5) sensibilidade, que se refere a como pequenas mudanças na definição podem alterar o conjunto de propriedades que já estão bem definidas. Ele deixa claro o objetivo da avaliação de ontologias. Afirma que:

a avaliação de ontologias refere-se à construção correta do conteúdo da ontologia, isto é, garantir que a sua definição corresponda corretamente aos requisitos da ontologia e às questões de competência ou reflète corretamente ao mundo real. O objetivo é comprovar a conformidade do modelo de mundo com o mundo modelado formalmente. Avaliação de ontologia inclui: cada definição e axioma individual; coleção de definições e axiomas que são declarados explicitamente na ontologia; definições que são importadas de outras ontologias; definições que podem ser inferidas utilizando outras definições e axiomas (GOMEZ-PEREZ, 2001, p. 393,

tradução própria⁷).

Antes de abordar as propostas de avaliação de ontologias, cabe classificá-las para clarear o entendimento. A literatura mostra que não há um consenso com relação à classificação. Isso se deve à diversidade de propostas existentes. Pesquisadores como Porzel e Malaka (2004) classificam as propostas como aquelas de cunho quantitativo e qualitativo; outros, como Sabou et al (2006) preferem classificá-las de acordo com o método de avaliação, automático ou manual. Brank, Grobelnik e Mladenić (2006) apontaram outro tipo de classificação, que vai além das classificações dicotômicas (ou seja, que possuem mais de duas categorias para classificação), separando as propostas de avaliação de ontologias em quatro classes. São elas:

- baseadas na comparação da ontologia a um 'modelo padrão' ou '*gold standard*';
- baseadas no uso da ontologia em uma aplicação/tarefa e avaliação dos resultados;
- envolvendo comparações com fontes de dados não estruturados sobre um domínio que será coberto pela ontologia;
- aquelas em que a avaliação é feita por humanos, que tentam estimar o quão bem a ontologia responde a um conjunto predefinido de critérios, padrões, requerimentos.

Além da classificação das diferentes propostas de avaliação de ontologias, há autores que dividem cada avaliação em níveis. Os principais são: o léxico, que mede a qualidade das palavras; o taxonômico, que considera as conexões hierárquicas entre os conceitos que utilizam a relação 'é-um'; o semântico, que aborda as outras relações entre os conceitos; o sintático, que considera os requisitos sintáticos da linguagem formal usada para descrever a ontologia; o contextual ou de aplicação, que considera o contexto da ontologia; o estrutural, que leva em consideração os princípios e critérios envolvidos na construção da ontologia (FERNÁNDEZ, CANTADOR, CASTELLS, 2006).

Um dos argumentos para essa divisão em níveis é a diminuição da complexidade na avaliação, visto que avaliar cada nível da ontologia separadamente pode ser mais simples do

⁷The evaluation of ontologies refers to the correct building of the content of the ontology, that is, ensuring that its definitions (a definition is written in natural language and in a formal language) correctly implement ontology requirements and competency questions or perform correctly in the real world. The goal is to prove compliance of the world model (if it exists and is known) with the world modeled formally. Ontology evaluation includes: each individual definition and axiom; collection of definitions and axioms that are stated explicitly in the ontology; definitions that are imported from other ontologies; definitions that can be inferred using other definitions and axioms (GOMEZ-PEREZ, 2001, p. 393)

que avaliar a ontologia por completo.

No decorrer dos anos, várias propostas de avaliação de ontologias foram desenvolvidas. Como poderá ser visto nas próximas subseções, a maioria dessas propostas foram construídas com base em ontologias específicas e seguindo os objetivos dessa ontologia.

3.1 Avaliações Baseadas na Comparação a um Modelo Padrão

Nesse tipo de avaliação, a ontologia é comparada a outro modelo construído de forma ideal, conhecido como *Gold Standard*, “padrão ouro” ou ainda “modelo padrão.” Esse modelo pode ser outra ontologia, que servirá de base para a avaliação. Na literatura, encontram-se alguns trabalhos de pesquisadores que utilizam essa abordagem; entre eles, Maedche e Staab (2002), Gangemi et al (2006); Brank, Grobelnik e Mladení c (2006), Dellschaft e Staab (2006), Zavitanos, Paliouras e Vouros (2011), que são apresentados a seguir.

Maedche e Staab (2002) apresentam uma abordagem baseada na medida de similaridade para comparar ontologias. A ontologia avaliada é comparada com um padrão ouro (*Gold Standard*), que pode ser outra ontologia (como é usada no trabalho dos autores) ou pode ser uma coleção de documentos de onde são retirados os termos representativos do domínio. No trabalho dos autores, a comparação de similaridades entre as ontologias é feita no nível lexical e conceitual. No primeiro, objetiva-se averiguar a semelhança entre os vocabulários das ontologias e, no segundo, objetiva-se comparar os conceitos e as relações delas.

Na proposta de Gangemi et al (2006) para avaliação de ontologias, são propostos dois modelos: uma meta ontologia (ou seja, uma ontologia que possui ontologias e elementos de ontologias em seu domínio), denominada O^2 ; e uma ontologia para a seleção de outras ontologias, denominada oQual. Com base na meta ontologia, é possível identificar três tipos principais de métricas para avaliação de ontologias: i) medidas estruturais, que enfatizam sintaxe e semântica formal nas ontologias representadas como grafos; ii) medidas funcionais, que se tratam do uso pretendido da ontologia; iii) medidas de usabilidade, relacionadas ao nível de anotação da ontologia. A meta ontologia O^2 é complementada pela ontologia oQual, que possibilita a realização de análises sobre elementos, processos e atributos da ontologia sob avaliação.

Na proposta de Brank, Grobelnik e Mladení c (2006), a abordagem para avaliação de ontologia permite a avaliação automática com base nas instâncias. É baseada no paradigma do *Gold Standard* e o foco principal é comparar quão bem determinada ontologia se assemelha ao *Gold Standard* no arranjo de instâncias dentro dos conceitos e o arranjo hierárquico entre os conceitos. Os autores dizem que a avaliação é similar às outras que são baseadas na comparação a um *Gold Standard*; porém, na abordagem deles, a avaliação destaca os atributos dos conceitos da ontologia e não a descrição em linguagem natural dos conceitos e instâncias. Ressalta-se que o pré-requisito para a execução desse tipo de avaliação é que a ontologia avaliada precisa ser baseada no mesmo conjunto de instâncias que o *Gold Standard*.

Na proposta de Dellschaft e Staab (2006) apresenta-se um esquema para avaliação de ontologias com base em um “Gold Standard”. Os autores focam no problema das métricas para avaliação de ontologias, ou seja, qual das métricas citadas na literatura permitem realizar uma avaliação multimencional da ontologia, que abrange vários níveis e não apenas um, como acontece na maioria das propostas de avaliação. Assim, faz-se uma análise de várias métricas de avaliação de ontologia e em seguida são realizadas avaliações empíricas dessas métricas para no fim propor as métricas ideais para uma avaliação multinível de uma ontologia.

Zavitanos, Paliouras e Vouros (2011) apresentam uma proposta que transforma os conceitos e propriedades da ontologia em uma representação de espaço vetorial com o propósito de avaliá-la em comparação a um “Gold Standard”. As medidas de avaliação da proposta exploram a representação de espaço vetorial e calculam a semelhança entre as duas ontologias nos níveis léxico e relacional. Isso é feito com base em três etapas: 1) Transformação, os elementos da ontologia são transformados em distribuições probabilísticas; 2) Correspondência, compara-se os elementos da ontologia com os elementos do “Gold Standard”; 3) Avaliação, a ontologia avaliada é penalizada quando seus elementos se diferem dos elementos do “Gold Standard”.

3.2 Avaliações Baseadas no Uso da Ontologia em uma Aplicação /Tarefa

Nesse tipo de avaliação, a ontologia é testada com base no seu desempenho na realização de uma tarefa específica. Em uma proposta de avaliação desse tipo, Porzel e Malaka (2004) propõem uma abordagem quantitativa para avaliação de ontologias. Os

autores se baseiam na questão fundamental: quão efetiva é uma ontologia na execução de uma tarefa bem definida? Eles defendem que, se uma ontologia foi construída para realizar determinada tarefa, a melhor forma de avaliar é testando a ontologia na execução dessa tarefa específica. Na proposta dos autores, primeiro escolhe-se uma ou mais ontologias que serão utilizadas; segundo, escolhe-se a tarefa que deseja que a ontologia escolhida execute (os autores frisam que a tarefa escolhida precisa ser suficientemente complexa para que sirva como uma referência adequada na avaliação); depois, escolhe-se o padrão ouro (esse padrão define a medida de desempenho perfeita), ou seja, um modelo com as respostas adequadas que servirá de comparação para a ontologia avaliada. Assim, um algoritmo específico (construído pela equipe de avaliação) usa a ontologia para desempenhar a tarefa proposta. O resultado da avaliação mostra três tipos de erros: erros de inserção, que indicam conceitos supérfluos encontrados na ontologia; erros de eliminação, que indicam conceitos faltantes; e erros de substituição, que indicam conceitos ambíguos. Ainda de acordo com Porzel e Malaka (2004), as ontologias podem ser averiguadas em três níveis básicos: o escopo do vocabulário, a qualidade da taxonomia e a adequabilidade das relações semânticas.

Fernández, Cantador e Castells (2006) apresentam uma ferramenta, nomeada CORE (Collaborative Ontology Reuse and Evaluation) para identificar qual ontologia de um repositório é a mais adequada para representar um domínio. Para a realização da tarefa proposta, o ambiente da ferramenta é dividido em três componentes: o primeiro recebe a descrição do problema baseada em um conjunto de termos; o segundo aplica critérios automáticos de avaliação às ontologias, com o objetivo de indicar qual se adéqua melhor ao problema; o terceiro faz uso de avaliações manuais em ontologias, as quais incorporam a avaliação colaborativa de usuários. Com base nesses três componentes, o sistema elenca qual ontologia representa melhor o domínio proposto.

A proposta de Strasunskas e Tomassen (2008) visa identificar quais recursos da ontologia influenciam na tarefa de busca na Web. Elaborou-se um modelo teórico para medir o desempenho da ontologia na realização da tarefa de busca, com o objetivo de identificar a importância da qualidade da ontologia para a realização da tarefa. Em seguida, o modelo teórico foi testado experimentalmente em quatro ontologias de diferentes domínios, o que permitiu identificar como a variação de qualidade de uma ontologia pode influenciar no desempenho de uma busca na Web.

Clarke et al (2013) fazem um experimento de avaliação baseada em tarefa da Gene Ontology (GO). Na proposta dos autores avalia-se a qualidade e utilidade da GO e suas anotações através do seu desempenho em uma tarefa comum, ou seja, utiliza-se análises enriquecidas em conjunto de genes com o objetivo de identificar quais termos da ontologia estão bem representados e quais não estão. Isso determina a efetividade das anotações da Gene Ontology fornecendo resultados biologicamente precisos. Além disso, permite entender os pontos fortes e fracos da ontologia avaliada.

3.3 Avaliações Baseadas na Comparação a uma Fonte de Dados

As avaliações desse tipo são feitas através da comparação da ontologia a uma fonte de dados, que geralmente se trata de textos do domínio coberto pela ontologia avaliada. Entre as propostas de avaliação encontradas na literatura, estão a de Patel et al (2003), Brewster et al (2004), Elhadad, Gabay e Netzer (2010), Ouyang e Qu (2011), Hlomani e Stacey (2013) e Hlomani e Stacey (2014), que são abordadas a seguir.

Patel et al (2003) desenvolveram um protótipo denominado Ontokhoj para buscar, classificar e ranquear ontologias com base em dados extraídos da web semântica. As ontologias são carregadas em um software de classificação que determina se a ontologia se refere a um tópico particular.

Brewster et al (2004) propõem uma abordagem para avaliação de ontologias através da comparação da ontologia com uma coleção de textos de determinado domínio. O objetivo é medir a proximidade entre a ontologia e o domínio que ela se propõe a representar. Os autores justificam que a coleção de textos têm-se mostrado a fonte de dados mais efetiva para a construção de várias ontologias. Eles dizem ainda que esse tipo de avaliação pode ser feito através da extração de termos em dada coleção de textos e simplesmente contagem dos termos que aparecem na ontologia e na coleção. Porém, em seus estudos, eles propõem o uso de uma arquitetura mais sofisticada que se baseia em três passos: 1) identificar palavras-chave/termos na coleção de textos; 2) expandir a consulta, ou seja, procurar termos semanticamente próximos, como os sinônimos, por exemplo (os autores utilizam o Word Net⁸ para fazer isso); 3) fazer o mapeamento dos termos

⁸Word Net: é um banco de dados lexical para o idioma Inglês, que agrupa palavras em conjuntos de sinônimos e fornece definições curtas. <http://wordnetweb.princeton.edu/perl/webwn>

identificados na ontologia que está sendo avaliada. Por fim, ao comparar a coleção com a ontologia, contam-se quantos termos na ontologia estão entre os termos da coleção.

Elhadad, Gabay e Netzer (2010) propõem um método de avaliação de ontologias que conta com o mapeamento de instâncias de uma ontologia a documentos textuais. O processo que orienta a avaliação é o seguinte: a partir de uma determinada instância da ontologia, coleta-se grande quantidade de documentos textuais que têm relação com a instância. Em seguida, faz-se uma análise para determinar se a ontologia reflete os documentos encontrados. A partir desse mapeamento, pode-se avaliar a adequação das relações das ontologias medindo o seu potencial de classificação através dos documentos textuais. De acordo com os autores, o método fornece um *feedback* concreto aos ontologistas e uma estimativa quantitativa da adequação funcional das relações de ontologias para uma melhoria na experiência da pesquisa.

Ouyang e Qu (2011) apresentam uma proposta para avaliar a qualidade da ontologia em comparação a um corpus de dados utilizando três critérios: cobertura, coesão e acoplamento. Para a realização dos experimentos empíricos, utilizaram-se cinco ontologias, três delas relacionadas ao domínio de projetos, do Departamento de Ciência da Computação da Universidade de Maryland e duas relacionadas ao meio ambiente, da NASA. Os autores apresentaram expressões para o cálculo dos critérios mencionados e fizeram os cálculos de cada ontologia em comparação com dois corpus de dados da American National Corpus (OANC). Em seguida os resultados foram tabelados e comparados em gráficos. A proposta dos autores vai além e envolve a avaliação de usuários, que com base nos resultados da comparação decide qual dos critérios (cobertura, coesão e acoplamento) possui mais peso e assim qual ontologia atende melhor aos seus requisitos.

Hlomani e Stacey (2013) se baseiam na proposta de Brewster et al (2004) para avaliar ontologias no domínio de gerenciamento de fluxos de trabalho. Os principais passos seguidos para a avaliação foram: definição do corpus, ou seja, selecionaram-se os documentos (artigos acadêmicos revisados); cálculo de similaridade entre as palavras de cada artigo e os conceitos das ontologias avaliadas; por fim, avaliação estatística, ou seja, análise de qual ontologia possuía mais similaridade com os documentos do domínio.

Em outro trabalho, Hlomani e Stacey (2014) propõem uma visão multidimensional

para avaliação de ontologias baseados na comparação a uma fonte de dados. De acordo com os autores, as propostas até então de avaliação de ontologias com base em uma fonte de dados consideram o conhecimento do domínio como sendo estático. Criticando essa forma de ver o conhecimento do domínio, eles propõem um esquema teórico para avaliação de ontologias que considera o conhecimento como sendo dinâmico, que pode variar com base em várias dimensões. Eles discutem que se o conhecimento do domínio é dinâmico, então as ontologias deveriam ser avaliadas com base nas mesmas dimensões que definem as características do domínio.

3.4 Avaliações Feitas por Humanos

As propostas de avaliação que utilizam essa abordagem geralmente são realizadas manualmente, de maneira que uma pessoa avalia a ontologia com base em alguns critérios pré-estabelecidos. Exemplos desse tipo de avaliação podem ser vistos em Lozano-Tello e Gómez-Pérez (2004), Almeida (2009) e Orme, Yao e Etzkorn (2007).

Lozano-Tello e Gómez-Pérez (2004) adotam uma abordagem para avaliação de ontologias baseada em uma taxonomia de 160 características, denominada “Estrutura Multinível de Características”, que são segmentadas em 5 dimensões. São elas: conteúdo, linguagem, metodologia, ferramentas e custos. Essas dimensões são definidas por meio de um conjunto de fatores, que são elementos requeridos para análise para obter o valor da dimensão adequada. Os fatores são os elementos fundamentais que devem ser analisados para obter o valor de cada dimensão. Dimensões, fatores e características são nomeados ‘critério’ e reunidos em uma ontologia referência. O método chamado Ontometric tenta obter o valor de cada critério por meio de fórmulas que avaliam os pesos atribuídos aos termos. O objetivo é selecionar a melhor opção entre alternativas e assim o usuário pode decidir qual ontologia é mais adequada para determinado projeto.

Orme, Yao e Etzkorn (2007) apresentam uma forma de avaliar ontologias com base no *feedback* de avaliadores humanos com experiência na área de avaliação de softwares. Eles propõem um conjunto de métricas, com base em conceitos da área de avaliação de software, para medir a complexidade e coesão de ontologias. Entre as métricas estão: validade do atributo, validade da unidade, validade instrumental, validade do protocolo, escada da unidade, etc. Após uma avaliação teórica de cada métrica proposta é realizado um teste empírico com avaliadores humanos que determinam a qualidade das ontologias com

base na atribuição de valores para cada métrica. Cada métrica reflete diferentes tipos de mudanças necessárias na ontologia.

Almeida (2009) propõe uma abordagem de avaliação baseada no conteúdo da ontologia. Ele apresenta uma pesquisa qualitativa para avaliar o conteúdo de uma ontologia de uma grande companhia energética brasileira. O processo de avaliação consiste de duas fases: a primeira trata-se do uso de um protótipo pelos especialistas do domínio; a segunda fase trata-se de três questionários baseados nas seguintes dimensões: questões de competência, qualidade da informação e objetivos educacionais da taxonomia. O protótipo trata-se de um simples motor de busca, que tem o objetivo de buscar termos correspondentes a conceitos na ontologia. Ele foi implementado com quatro funcionalidades: uma interface de busca e navegação que permite aos usuários busca por conceitos e relações; uma hierarquia de conceitos que possibilita acessar conceitos e relações através de *hyperlinks* organizados na taxonomia; uma visão parabólica dos conceitos que oferece aos usuários uma melhor visualização da estrutura como um todo e assim permite entender o contexto dos conceitos durante uma busca; por fim, uma interface de gerenciamento, em que é possível carregar diferentes ontologias no protótipo. O processo funcionou da seguinte forma: primeiro os especialistas do domínio utilizaram o protótipo por uma semana e depois responderam aos três questionários, constituídos de 81 perguntas (20 perguntas no questionário da dimensão questões de competência; 36 perguntas no questionário da dimensão qualidade da informação e 25 perguntas no questionário da dimensão objetivos educacionais). Cada pergunta do questionário possuía uma escala de 1 a 5 (sendo 1 'não corresponde às necessidades' e 5 'corresponde totalmente às necessidades'). Assim, com base no cálculo das escalas atribuídas pelos especialistas do domínio, é feita a avaliação da ontologia.

No QUADRO 3 é possível observar a relação entre as propostas e os níveis de avaliação. Apresenta-se quais níveis da ontologia cada tipo de avaliação consegue analisar.

Quadro 3 - Visão Global das Propostas de Avaliação

Proposta de avaliação				
Nível	Padrão ouro	Baseada em aplicação	Baseada em uma fonte de dados	Feita por humanos
Léxico, vocabulário, conceito, dados	X	X	X	X
Hierárquico, taxonômico	X	X	X	X
Outras relações semânticas	X	X	X	X
Contexto, aplicação		X		X
Sintático	X			X
Estrutura, arquitetura, desenho				X

Fonte: Adaptado de Brank, Grobelnik, Mladeníć, 2006, p. 2.

Constata-se, com base no quadro apresentado, que as propostas de avaliação feitas por humanos são as únicas que conseguem avaliar todos os seis níveis de uma ontologia. Já as propostas baseadas na comparação a um *Gold Standard* e as propostas baseadas no uso da ontologia em uma aplicação/tarefa permitem a avaliação de quatro níveis de uma ontologia, com a diferença em um nível de avaliação. Enquanto as propostas baseadas na comparação a um *Gold Standard* conseguem avaliar o nível 'sintático' e não permitem avaliar o nível do 'contexto e aplicação', as propostas baseadas no uso da ontologia em uma aplicação/tarefa possibilitam a avaliação do nível 'contexto e aplicação', porém, não possibilitam a avaliação do nível 'sintático'. Por fim, as propostas baseadas na comparação

da ontologia a uma fonte de dados permitem a avaliação de três níveis de uma ontologia.

É válido frisar que o fato de um tipo de proposta cobrir todos os níveis não garante a sua qualidade ou supremacia em relação aos outros tipos de propostas, visto que é necessário levar em conta outros critérios na avaliação de ontologias, por exemplo, quais são os objetivos da avaliação e o que exatamente se pretende avaliar na ontologia. Portanto, se o objetivo é apenas avaliar o desempenho da ontologia na representação de conceitos, que se encaixa no nível 'léxico, vocabulário e conceito', qualquer um dos tipos de propostas podem ser utilizadas. Assim, para a escolha da proposta mais adequada levam-se em consideração os outros objetivos e possibilidade de aplicação de determinada proposta.

Encontram-se na literatura algumas críticas com relação a cada um dos tipos de propostas para avaliação de ontologias, que podem ajudar na escolha da proposta para avaliação de determinada ontologia. Com relação às propostas baseadas na comparação da ontologia a um *Gold Standard*, alguns pesquisadores (HLOMANI; STACEY, 2013) dizem que é complexo estimar até que ponto o modelo é ideal, além de ser difícil encontrar um modelo que englobe o mesmo domínio da ontologia que está sendo avaliada. Ademais, quando os resultados da ontologia avaliada diferem dos resultados do *Gold Standard*, é difícil descobrir a origem do problema. No que se refere às propostas de avaliações baseadas em tarefas e aplicações, os críticos dizem que elas são muito específicas, pois testam a ontologia na realização de uma tarefa. Portanto, a ontologia pode obter ótimos resultados em determinada tarefa, mas quando testada em outra, os resultados podem ser diferentes. Dessa forma, dependendo dos objetivos da avaliação, esse tipo de proposta não se aplica. Com relação às propostas de avaliação realizadas por humanos, as dificuldades são estabelecer quem são os usuários certos e quais os melhores critérios para a avaliação da ontologia. Além disso, o fato da avaliação depender da intuição do especialista faz com que se torne incerto a verificação do rigor da avaliação. No que tange às propostas de avaliação de ontologias baseadas na comparação a uma fonte de dados, Hlomani e Stacey (2014) apontam para o fato da maioria das propostas considerarem o conhecimento de determinado domínio como algo constante, ou seja, que não se altera, porém não é isso que acontece, o conhecimento de um domínio é algo dinâmico. Assim, se uma ontologia foi avaliada em comparação ao conhecimento de um domínio em determinada época, se essa mesma ontologia for avaliada em comparação ao mesmo domínio de conhecimento em uma época diferente, os resultados da avaliação podem ser diferentes, posto que com o passar

do tempo o domínio de conhecimento tenha sofrido alterações. Essa dinamicidade do conhecimento pode variar de domínio para domínio.

4 CONSIDERAÇÕES FINAIS

Este trabalho não pretendeu esgotar todos os assuntos relacionados a avaliação de ontologias. Existem muitos trabalhos que abordam a questão de vários pontos de vista. O que se percebe na literatura são muitas propostas, hipóteses e alguns experimentos, porém, pouca coisa está consolidada. E como afirmam vários autores, entre eles Gómez-Pérez (2001) e Vrandečić (2009), a avaliação de ontologias ainda é um campo emergente, que precisa de um aprofundamento de ideias e de diretrizes.

Com base no levantamento feito na literatura, percebe-se que as diversas abordagens para avaliação de ontologias variam de acordo com o tipo e o propósito da ontologia que está sendo avaliada. Assim, constata-se que a temática ainda está em aberto e são necessários estudos para o desenvolvimento do campo avaliação de ontologias.

REFERÊNCIAS

ALMEIDA, Mauricio Barcellos. A proposal to evaluate ontology content. **Applied Ontology**, Birmingham, v. 4, n. 3-4, p. 245-265, 2009. Disponível em: <<http://content.iospress.com/articles/applied-ontology/ao070>>. Acesso em: 05 abr. 2014.

ALMEIDA, Maurício Barcellos; BAX, Marcello Peixoto. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da Informação**, Brasília, DF, v. 32, n. 3, p. 7-20, set./dez. 2003.

BRANK, J.; GROBELNIK, M.; MLADENIĆ, D. A survey of ontology evaluation techniques. In: PROCEEDINGS OF THE CONFERENCE ON DATA MINING AND DATA WAREHOUSES, Houston, TX, 2005, [Proceedings...]. Houston, TX: IEEE, 2005. Disponível em: <<http://http://ailab.ijs.si/dunja/sikdd2005/papers/BrankEvaluationSiKDD2005.pdf>>. Acesso em: 03 set. 2014.

BREWSTER, C. *et al.* Data driven ontology evaluation. In: INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION, 2004, Lisboa. **Anais...** Lisboa: [s.n], 2004.

Disponível em: <<http://www.cbrewster.com/papers/BrewsterLREC.pdf>>. Acesso em: 22 fev. 2014.

CHANDRASEKARAN, Balakrishnan; JOSEPHSON, John R.; BENJAMINS, V. Richard. What are ontologies, and why do we need them?. **IEEE Intelligent systems**, Piscataway, New Jersey, n. 1, p. 20-26, 1999.

CLARKE, Erik L. *et al.* A task-based approach for Gene Ontology evaluation. **Journal of Biomedical Semantics**, Londres, v. 4, n. 1, p. 1, 2013. Disponível em: <<https://jbiomedsem.biomedcentral.com/articles/10.1186/2041-1480-4-S1-S4>>. Acesso em: 10 jan. 2015.

DELLSCHAFT, K.; STAAB, S. On How to Perform a Gold Standard Based Evaluation of Ontology Learning. In: PROC. FIFTH INT'L CONF. SEMANTIC WEB, 2006, Heidelberg. **Anais... Heidelberg**: [s.n], 2006. p. 228-241. Disponível em: <<http://www.socialdynamics.it/tagora/wp-content/2007/05/Dellschaft2006fy.pdf>>. Acesso em: 03 abr. 2014.

ELHADAD, D. GABAY, M.; NETZER, Y. Automatic evaluation of search ontologies in the entertainment domain using text classification. In: _____. **Applied Semantic Technologies: Using Semantics in Intelligent Information Processing**. [S.l.]: Taylor and Francis, 2010. Disponível em: <<https://www.cs.bgu.ac.il/~elhadad/papers/OntologyEvaluation.pdf>>. Acesso em: 10 jun. 2014.

FERNÁNDEZ, Miriam; CANTADOR, Iván; CASTELLS, Pablo. CORE: A tool for collaborative ontology reuse and evaluation. In: 4TH INTERNATIONAL WORKSHOP ON EVALUATION OF ONTOLOGIES FOR THE WEB, 4., 2006, Edinburgh, UK. **Proceedings...** Edinburgh, UK, 2006. Disponível em: <<http://oro.open.ac.uk/28592/1/eon06.pdf>>. Acesso em: 03 abr. 2014.

FURGERI, Sérgio. **Representação de informação e conhecimento**: estudo das diferentes abordagens entre Ciência da Informação e a Ciência da Computação. 2006. 161 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Ciência da Informação, Pontifícia Universidade Católica de Campinas, Campinas, 2006.

GANGEMI, A. et al. A metaontology-based framework for ontology evaluation and *selection*. In: 4TH INTERNATIONAL WORKSHOP ON EVALUATION OF ONTOLOGIES FOR THE WEB, 4., Edinburgh, UK, 2006, Edinburgh, UK. **Proceedings...** Edinburgh, UK, 2006. Disponível em: <<http://ceur-ws.org/Vol-179/eon2006gangemietal.pdf>>. Acesso em: 03 abr. 2014.

GILCHRIST, Alan. Thesauri, taxonomies and ontologies: an etymological note. **Journal of Documentation**, [S.l.], v. 59, n. 1, p. 7-18, 2003.

GÓMEZ-PÉREZ, A. Evaluation of ontologies. **International Journal of Intelligent Systems**, Alberta, Canadá, v. 16, n. 3, 2001. p. 391-409. Disponível em: <[http://onlinelibrary.wiley.com/doi/10.1002/1098-111X\(200103\)16:3%3C391::AID-INT1014%3E3.0.CO;2-2/abstract](http://onlinelibrary.wiley.com/doi/10.1002/1098-111X(200103)16:3%3C391::AID-INT1014%3E3.0.CO;2-2/abstract)>. Acesso em: 03 abr. 2014.

GOMÉZ-PÉREZ, A. PAZOS, Juristo N, J. Evaluation and assessment of the knowledge sharing technology. In: MARS, N. J. L. **Towards very large knowledge bases: knowledge building & knowledge sharing**. Amsterdam: IOS, 1995. p. 289-296. Disponível em: <https://www.researchgate.net/publication/243773900_Evaluation_and_Assessment_of_the_Knowledge_Sharing_Technology>. Acesso em: 03 abr. 2014.

GÓMEZ-PÉREZ, A.; FERNÁNDEZ, M.; VICENTE, A.J. de. Towards a method to conceptualize domain ontologies. In: 12TH EUROPEAN CONFERENCE ON ARTIFICIAL INTELLIGENCE (ECAI'96), 12., 1996, Budapest, Romênia. **Proceedings...** Budapest, Romênia: Facultad de Informática (UPM), 1996. Disponível em: <http://www.academia.edu/13104696/Towards_a_Method_to_Conceptualize_Domain_Ontologies>. Acesso em: 12 ago. 2014.

GRUBER, Thomas R. Toward principles for the Design of Ontologies Used for Knowledge Sharing. In: FORMAL ONTOLOGY IN CONCEPTUAL ANALYSIS AND KNOWLEDGE REPRESENTATION, 1993, Padova. Italy, 1993.

GRUNINGER, M.; FOX, M. S. Methodolgy for the design and evaluation of ontologies. In: PROCEEDINGS OF THE IJCAI WORKSHOP ON BASIC ONTOLOGICAL ISSUES IN KNOWELDGE SHARING, Toronto. **Proceedings...** Montreal, Canadá: [s.n], 1995. Disponível em: <<http://stl.mie.utoronto.ca/publications/gruninger-ijcai95.pdf>>. Acesso em: 12 ago. 2014.

GUARINO, N. (Ed.). Formal ontology in information systems. In: PROCEEDINGS OF THE FIRST INTERNATIONAL CONFERENCE (FOIS'98), 1998, Trento, Italia. **Proceedings...** Trento, Italia: IOS Press, 1998. p. 3-15.

HLOMANI, Hlomani; STACEY, D. A. An Extension to the Data-driven Ontology Evaluation. In: 15TH INTERNATIONAL CONFERENCE ON INFORMATION REUSE AND INTEGRATION (IRI), 15., 2014, Ontário, Canadá. **Proceedings...** Ontário, Canadá: IEEE, 2014.

HLOMANI, Hlomani; STACEY, D. A. Contributing evidence to data-driven ontology evaluation: Workflow ontologies perspective. In: 5TH INTERNATIONAL CONFERENCE ON KNOWLEDGE ENGINEERING AND ONTOLOGY DEVELOPMENT, 5., Vilamoura, Portugal. **Proceedings...** Vilamoura, Portugal: SCITEPRESS, 2013. p. 207-213.

LOZANO-TELLO, A.; GÓMEZ-PÉREZ, A. Ontometric: A method to choose the appropriate ontology. **Journal Of Database Management**, Hershey, Pensilvânia, v. 15, n. 2, 18 p. 2004. Disponível em: <<http://www.igi-global.com/article/ontometric-method-choose-appropriate-ontology/3308>>. Acesso em: 10 set. 2014.

MAEDCHE, A.; STAAB, S. Measuring similarity between ontologies. In: EUROPEAN CONFERENCE ON KNOWLEDGE ACQUISITION AND MANAGEMENT, 2002, Madri, Espanha. **Proceedings...** Madri, Espanha: [s.n], 2002. Disponível em: <<http://disi.unitn.it/~accord/RelatedWork/Matching/Staab2002.pdf>> . Acesso em: 20 set. 2014.

MOREIRA, Alexandra; ALVARENGA, Lídia; OLIVEIRA, Alcione de Paiva. O nível do conhecimento e os instrumentos de representação: tesouros e ontologias. **DataGramZero: Revista de Ciência da Informação**, Rio de Janeiro, v. 5, n. 6, 2004. Disponível em: <http://www.dgz.org.br/dez04/Art_01.htm>. Acesso em: 02 jun. 2014.

NOY, N. F.; MCGUINNESS, D. L. Ontology Development 101: A Guide to Create Your First Ontology. 2001. Disponível em: <http://protege.stanford.edu/publications/ontology_development/ontology101.pdf> . Acesso em: 11 set. 2014.

NOY, N. F., HAFNER C.D. The state of the art in ontology design: a survey and comparative review. **AI Magazine**, Palo Alto, CA, v. 18, n. 3, 1997. Disponível em: <<http://www.aaai.org/ojs/index.php/aimagazine/article/view/1306/1207>>. Acesso em: 02 jun. 2014.

ORME, A.; YAO, H.; ETZKORN, L. Indicating ontology data quality, stability, and completeness throughout ontology evolution. **Journal of Software Maintenance and Evolution: Research and Practice**, Orme, TN, v. 19, n. 1, p. 49-55, Jan./Feb. 2007. Disponível em: <<http://dx.doi.org/10.1002/smr.341>>. Acesso em: 20 jan. 2015.

OUYANG, Liubo et al. A method of ontology evaluation based on coverage, cohesion and coupling. In: EIGHTH INTERNATIONAL CONFERENCE ON FUZZY SYSTEMS AND KNOWLEDGE DISCOVERY. 8., 2011, Shanghai. **Proceedings...** Shanghai: IEEE, 2011. p. 2451-2455.

PATEL, Chintan et al. OntoKhoj: A semantic web portal for ontology searching, ranking and classification. INTERNATIONAL WORKSHOP ON WEB INFORMATION AND DATA MANAGEMENT. 5., 2003, Chintan. **Proceedings...** Chintan: ACM, 2003. p. 58–61.

PORZEL, R.; MALAKA, R. A task-based approach for ontology evaluation. In: 16TH EUROPEAN CONFERENCE ON ARTIFICIAL INTELLIGENCE AND WORKSHOP ON ONTOLOGY LEARNING AND POPULATION IN. 16., Valencia, Espanha. **Proceedings...**Valencia, Espanha: [s.n], 2004. Disponível em: <<http://olp.dfki.de/ecai04/final-porzels.pdf>>. Acesso em: 10 set. 2014.

SABOU, M. *et al.* Ontology selection: Ontology evaluation on the real semantic web. 4TH INTERNATIONAL WORKSHOP ON EVALUATION OF ONTOLOGIES FOR THEWEB AT THE 15TH INTERNATIONAL WORLD WIDE WEB, 4., 2006, Edinburgh, UK. **Proceedings...** Edinburgh, UK: Knowledge Media Institute, 2006. Disponível em: <<http://ai.ia.agh.edu.pl/wiki/media/pl:miw:2009:eon2006.pdf>>. Acesso em: 10 set. 2014.

SMITH, Barry. Beyond concepts: ontology as a reality representation. Proceedings of FOIS 2004. In: 3RD INTERNATIONAL CONFERENCE ON FORMAL ONTOLOGY IN INFORMATION SYSTEMS. 3., 2004, Torino, Italia. **Proceedings...** Torino, Italia: IOS Press, 2004.

SOERGEL, Dagobert. The rise of ontologies or the reinvention of classification. **Journal of the American Society of Information Science**, [S.l.], v. 50, n. 12, p. 1119-1120, oct. 1999. Disponível em: <<http://www.dsoergel.com/cv/B70.pdf>>. Acesso em: 15 set. 2014.

SOUZA, R. R.; ALVARENGA, L. A web semântica e suas contribuições para a ciência da informação. **Ciência da Informação**, Brasília, DF, v. 33, n. 1, p. 132-141, jan./abr. 2004. Disponível em: <<http://dx.doi.org/10.1590/S0100-19652004000100016>>. Acesso em: 15 set. 2014.

STRASUNSKAS, D.; TOMASSEN, S. L.; Empirical insights on a value of ontology quality in ontology-driven web search. In: MEERSMAN, R.; TARI, Z. (Eds.). **Lecture notes in Computer Science**. Berlin: Springer, 2008. v. 2, cap. 2, p. 1319-1337.

STUDER, R.; BENJAMINS, V.; FENSEL, D. Knowledge Engineering: Principles and Methods. **Data and Knowledge Engineering**, [S.l.] v. 25, n. 1-2, p. 161-197, 1998.

SUÁREZ-FIGUEROA, M. C.; PRADEL, C.; HERNANDEZ, N. Verifying ontology requirements with SWIP. In: 18TH INTERNATIONAL CONFERENCE ON KNOWLEDGE ENGINEERING AND KNOWLEDGE MANAGEMENT (EKAW)18., 2012, Galway, Irlanda. **Proceedings...** Galway, Irlanda: [s. n], 2012.

USCHOLD, Mike; GRUNINGER, M. Ontologies: Principles, methods and applications. **Knowledge Engineering Review**, [S.l.], v. 11, n. 2, p. 93-136, 1996. Disponível em: <<http://journals.cambridge.org/action/displayAbstract;jsessionid=273174E2C8795929808D17E3A1D7C27C.journals?fromPage=online&aid=4071848>>. Acesso em: 20 jun. 2015.

VICKERY, B. C. Ontologies. **Journal of Information Science**, [S. l.], v. 23, n. 4, p. 272-286, 1997.

VRANDECIC, Denny. Ontology evaluation. In: STAAB, S.; STUDER, R. **Handbook on Ontologies**. Springer: Springer, 2009. p. 293-313. Disponível em:<http://link.springer.com/chapter/10.1007/978-3-540-92673-3_13>. Acesso em: 12 ago. 2014.

ZAVITSANOS, Elias; PALIOURAS, Georgios; VOUIROS, George A. Gold standard evaluation of ontology learning methods through ontology transformation and alignment. **Transactions on Knowledge and Data Engineering**, [S.l.], v. 23, n. 11, p. 1635-1648, 2011.