

ESTUDO DE MODELAGEM CONCEITUAL BASEADA EM ONTOLOGIAS DE BIODIVERSIDADE

Isomar Lima da SILVA¹; José Laurindo Campos dos SANTOS²

¹Bolsista PAIC/FAPEAM-INPA. Coordenação de Tecnologia da Informação – CTIN/INPA; ²Orientador, PhD - CTN/INPA

1. Introdução

As pesquisas na área de Web Semântica têm avançado rapidamente. O objetivo é contribuir para a conclusão a próxima geração de tecnologia Web. A visão da Web Semântica é adicionar semântica ao conteúdo da Web objetivando facilitar os processos de busca e utilização para homens e máquinas (Campos 2007). Para isso é necessário utilizar os recursos de modelagem de ontologias para auxiliar no processo de exploração, transformação e visualização de dados e informações como imagens ou em outras formas sensoriais como a percepção, para se obter uma maior compreensão. Ontologias segundo Guarino (1995) “é um ramo da filosofia que estuda a organização e a natureza do mundo”, na computação ontologias são usadas para representação do conhecimento, descrever conceitos de um determinado domínio.

A modelagem de ontologias é fundamental para compreensão de dados de biodiversidades, pois facilita sua visualização, estes dados estão em diferentes formatos, o que os torna de difícil acesso. Um exemplo disto, segundo Campos *et al.* (2007), algumas árvores taxonômicas são claramente definidas – como em zoologia, para mamíferos. Entretanto, ainda existem divergências de autores em vários domínios além de muitas espécies para serem classificadas ou reclassificadas. Por esses motivos existem diversas instituições de pesquisas do Brasil como o Instituto Nacional de Pesquisas da Amazônia (INPA), o Museu Emílio Goeldi (MPEG) que vem dedicando suas atividades no estudo na área de biodiversidade brasileira.

Ao longo de anos de pesquisas, muitos dados foram gerados através de coletas e experimentos científicos, mas estes não seguem um mesmo padrão ou formato nas suas bases de dados e com a crescente procura por estes tipos de informações (biodiversidade) para fins econômicos ou de pesquisas, tornou-se de fundamental importância modelar, integrar e disseminar os diferentes tipos de dados das mais diversificadas bases de dados de biodiversidade para que se tenha um acesso e visualização clara dessas informações.

O uso de ontologias viabiliza a modelagem e a interoperabilidade semântica de sistemas distribuídos heterogêneos, isso contribui para que informações que são distribuídas de forma isolada sejam utilizadas de maneira íntegra (Campos 2007). Isto acontece por que ontologias produzem definições, axiomas que devem ser aceitos por uma comunidade no âmbito de um domínio, e tem por finalidade permitir que múltiplos agentes compartilhem conhecimento facilitando a interoperabilidade semântica.

O objetivo deste trabalho é estudar metodologias de modelagem conceituais de ontologias com a utilização da OntoUML, uma linguagem de modelagem conceitual baseada em uma (UFO) Ontologia de Fundamentação Unificada (Campos 2007), e a ferramenta de modelagem conceitual baseada em ontologias de biodiversidade OntoEditor documentando o estudo realizado.

2. Material e Métodos

Para elaboração deste trabalho realizou-se estudos bibliográficos em livros, artigos e dissertações voltadas para a área de biodiversidade e modelagem conceitual de ontologias, que serão apresentados neste capítulo.

2.1 Levantamento de trabalhos relacionados

Entre vários estudos relacionados à modelagem conceitual de ontologias, Lanzenberger *et al.* (2009), afirma que a modelagem conceitual tem um potencial interessante quando se trata de criar, explorar ou verificação complexa e grande coleção de dados, como ontologias. Por isso havendo uma grande quantidade de dados de um certo domínio as técnicas de modelagem de ontologias permitem que as mesmas possam ser compreendidas de forma simples e clara.

Em Guizzardi *et al.* (2007) em Ontologias de Fundamentação e Modelagem Conceitual, afirma que “a linguagem adequada de modelagem conceitual deveria possuir primitivas de modelagem que refletissem as categorias conceituais definidas em uma ontologia de fundamentação”, ou seja os modelos conceituais representam um parte da realidade.

Monteiro (2006) afirma que modelagem conceitual possibilita a representação do conhecimento contido nas definições de forma sistêmica, apoiando-se em alguns princípios comuns encontrados em instrumentos para a organização do conhecimento. Estes princípios ajudam nas definições dos conceitos, domínios e as relações conceituais. Com a modelagem conceitual de ontologias é possível provê uma fácil interoperabilidade entre sistemas de informação, processada de forma inteligente por agentes e permitindo o reuso de conhecimento entre sistemas (Pinto e Martins 2004).

Existem diversos grupos se dedicando ao estudo de ontologias, o grupo NEMO (Núcleo de Estudos em Modelagem Conceitual e Ontologias), por exemplo, defende explicitamente a necessidade de uma abordagem multidisciplinar para a construção de teorias de fundamentação para modelagem conceitual. O grupo NEMO desenvolve teorias que contribuem para o desenvolvimento de ontologias de fundamentação para dar suporte à modelagem conceitual com suas teorias que ajudam a compor a ontologia de fundamentação denominada *UFO (Unified Foundational Ontology)*. A UFO reúne teorias axiomáticas que explicitam sobre as principais categorias de conceitos usados em modelagem conceitual (Guizzardi *et al.* 2011).

Segundo Albuquerque (2011) ontologia captura um consenso sobre os conceitos do universo de discurso a partir da perspectiva do especialista de domínio. Na modelagem conceitual de ontologia conceitualização significa extrair regras informais que restringem a estrutura de um pedaço da realidade ou a abstração parcial da realidade, ou seja, não se busca representar todos os conceitos de um universo de discurso. A modelagem conceitual de ontologias usa definição dos conceitos dentro do domínio a ser estruturado precedendo a formulação de axiomas a partir dos conceitos que originarão outros conceitos. A modelagem conceitual de ontologias necessita de variáveis bem definidas para que uma determinada porção da realidade seja bem representada e a modelagem conceitual descreva adequadamente a realidade do domínio do problema.

Hepp (2007) estabelece seis variáveis características de um projeto ou modelagem de ontologia: expressividade, tamanho da comunidade relevante, dinâmica conceitual no domínio, número de elementos conceituais no domínio, grau de subjetividade em uma conceitualização do respectivo domínio e o tamanho médio da especificação por elemento. A Figura 1 mostra essas seis variáveis sob a forma de um gráfico de radar.

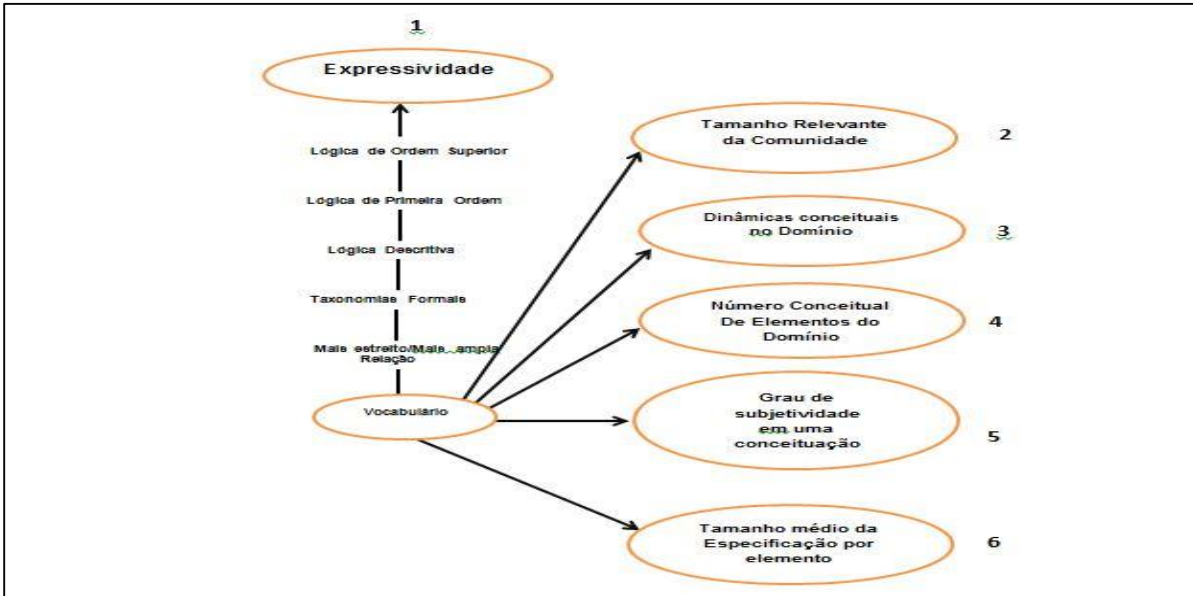


Figura 1- Variáveis características de um projeto de ontologia. Fonte: Hepp (2007).

A modelagem conceitual representa de maneira estruturada um domínio de conhecimento fazendo delimitações de suas relações, teorias e métodos. Para se ter uma modelagem conceitual de um domínio como o da biodiversidade faz-se necessário a utilização de boas ferramentas de modelagem e desenvolvimento de ontologias.

Com base na experiência prévia de alguns dos autores tais como: Oliveira *et al.* (2006), Minguine e Henrique (2006) no uso de ferramentas para a modelagem de ontologias, escolheu-se para esse estudo a ferramenta OntoEditor que segundo Sure *et al.* (2002), “é um ambiente de engenharia de ontologias, que é bastante singular no seu gênero, uma vez que combina a metodologia baseada em desenvolvimento de ontologias com capacidades de colaboração e de inferência”, ou seja, oferece as soluções possíveis após lhe serem feitas solicitações pré-determinadas. Na tela inicial do OntoEditor há quatro opções no menu: nova ontologia, abrir ontologia, excluir ontologia e visualizar ontologia.

Cada uma dessas opções do menu do ontoEditor possui telas específicas com campos de preenchimentos ou de seleção, os quais o usuário deve preencher.

O restante das opções só passa a valer ou ter suas funções ativadas quando a ontologia é criada e seu arquivo é submetido (*upload*) este arquivo que é feito o *upload* é em forma de texto.

O OntoEditor mostra a estrutura *folder-tree* (como uma estrutura de pastas) da ontologia, e seus nós que podem ser selecionados, abertos, fechados e editados nas atividades de edição. Os nós mostrados na

estrutura da ontologia o podem ser abertos pelo usuário com opções de alterar nó, excluir nó, incluir nó é criar novo link.

Como foi visto nos parágrafos anteriores a modelagem conceitual de ontologias é muito complexa, pois para a criação (modelagem) de uma ontologia segundo Monteiro (2006), é necessário seguir um roteiro para o desenvolvimento da mesma tais como: determinar os objetivos, definir o domínio, delimitar o escopo, considerar a reutilização de ontologias, considerar a reutilização da ontologia específica, modelar os conceitos, avaliar, revisar, documentar e aplicar a ontologia de acordo com seu propósito preterido.

Em relação aos trabalhos citados nos parágrafos anteriores, este estudo se diferencia pelo estudo da ferramenta e da linguagem de modelagem conceitual baseada em ontologias de biodiversidade, fazendo uma análise destas e assim, mostrar as melhores técnicas disponíveis e implícitas de informações dentro de um domínio específico (biodiversidade). A tabela 1 mostra outras ferramentas utilizadas para a modelagem de ontologia e suas principais características.

Tabela 1- Ferramentas para a modelagem, uso e edição de ontologias. Fonte: (Almeida e Bax, 2003, adaptada pelo autor).

Ferramentas	Descrição
GKB-Editor (Generic Knowledge Base Editor)	Ferramenta para navegação e edição de ontologias por meio de sistemas de representação baseados em frames. Oferece interface gráfica, em que os usuários podem editar diretamente a base de conhecimento e selecionar a parte que é de seu interesse (Paley e Karp, 1997).
JOE (JavaOntologyEditor)	Ferramenta para construção e visualização de ontologias. Proporciona gerenciamento do conhecimento em ambientes abertos, heterogêneos e com diversos usuários. As ontologias são visualizadas como um diagrama entidade-relacionamento, como o gerenciador de arquivos do MS Windows ou como uma estrutura em árvore (Mahalingam e Huhns, 1997).
OntoEditor	É um ambiente gráfico para edição de ontologias que permite inspeção, navegação, codificação e alteração de ontologias. O modelo conceitual é armazenado usando um modelo de ontologia que pode ser mapeado em diferentes linguagens de representação. As ontologias são armazenadas em bancos relacionais e podem ser implementadas em XML, FLogic, RDF(S) e DAML+OIL (Maedche <i>et al.</i> 2000)
Protégé	Ambiente interativo para projeto de ontologias, de código aberto, que oferece uma interface gráfica para edição de ontologias e uma arquitetura para a criação de ferramentas baseadas em conhecimento. A Arquitetura é modulada e permite a inserção de novos recursos (Noy, Ferguson e Musen 2000).
WebOnto	Ferramenta que possibilita a navegação, criação e edição de ontologias, representadas na linguagem de modelagem OCML. Permite o gerenciamento de ontologias por interface gráfica, inspeção de elementos, verificação da consistência da herança e trabalho cooperativo. Possui uma biblioteca com mais de cem ontologias (Domingue 1998).

3. Resultados e Discussão

Na análise e estudo dos materiais bibliográficos usados para este estudo constatou-se que a modelagem de ontologia facilita o processo de busca na Web, pois passa pelo discernimento humano baseado na interação com os demais usuários da rede isto ocorre por que os usuários estão percebendo que estabelecer comunicação entre os sistemas de seus próprios interesses facilitando o recebimento de dados, coletar e compartilhamento de informações tudo isso de forma mais rápida e simultânea.

Observou-se que o domínio do problema é complexo e a falta de uma conceitualização básica de domínio se uma grande dificuldade enfrentada pelos especialistas, pois grupos distintos podem ter diferentes visões de um mesmo domínio de um problema no caso biodiversidade.

Na análise e estudo de algumas ferramentas existentes para a modelagem de ontologias tais como: o OntoEditor, Protégé observou-se que as mesmas são fundamentais para a modelagem de ontologias no entanto, há apenas algumas para ajudar com a modelagem e representação de dados brutos para que sejam feitos de forma simples e legível, como o OntoEditor que combina metodologia baseada em desenvolvimento de ontologias, capacidades de colaboração e inferência.

Mesmo com algumas dessas ferramentas enfatizando interfaces de usuário intuitivas, a maioria destas requer conhecimento especializado em ontologia, ou seja, um especialista, (Albuquerque 2011).

4. Conclusão

O processo de desenvolvimento de ontologias é muito complexo e envolve pesquisadores da área de ontologias e pesquisadores de domínios específicos que utilizam ontologias, isso acontece especialmente nas áreas biológicas e correlatas por esses fatores ainda não existe uma ontologia consensual em biodiversidade que, é fundamental para que haja uma integração das diversas bases de dados de biodiversidade.

A modelagem conceitual de ontologias facilita o processo de busca da web semântica, pois se utiliza de relacionamento entre termos, descrição de conteúdo. Apesar de existir diversas ferramentas modelagem de ontologias como o OntoEditor, a falta de utilização modelagem e criação de ontologias são alguns dos empecilhos para a integração e visualização de dados de biodiversidade.

Contudo, ainda existem diversas atividades pertencentes ao processo de modelagem de ontologias que o OntoEditor não contempla, mas outras ferramentas de mesma funcionalidade, como: *Protégé*, *InxightStarTree*, *HyperEditor* e *TreeBolicGenerator* já fornecem. Apesar da existência de dessas diversas ferramentas fácil utilização para a criação e manipulação de ontologias como o OntoEditor e as demais ferramentas citadas os usuários e desenvolvedores não se interaço pois os mesmo enfrentam grandes barreiras como a utilização e criação de ontologias complexas e a desestruturação da web atual. Para melhorar essa situação estas ferramentas poderiam oferecer os recursos gráficos necessários às demandas do domínio complexos como biodiversidade, facilitando a interação com o especialista do domínio.

5. Referências Bibliográficas

- Albuquerque, A.C.F.; Campos, J.L. *Ontology Supported by CLOSi Data Schemas in the Semantic Web Context*. In Proceedings of ITEE 2005, Second International ICSC symposium on information technologies in environmental engineering, By Walter Leal Filho, Jorge Marx Gomez, Claus Rautenstrauch. Editors. September 25-27, 2005: Otto-von-Guericke-Universität Magdeburg, Germany: ISBN 978-3832243623.
- Ashburner, M. et al. 2000. *Gene ontology: tool for the unification of biology*. *The gene ontology consortium*. Nature Genetics.
- Almeida, M.B.; Bax, M.P. 2003. *Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção*. Ciência da informação, Brasília, DF.
- Berners-LEE, T.; Hendler, J.; Lassila, O. 2001. *The SemanticWeb*. *Scientific American*, 284(5): 34-43.
- Campos, M.L.A. 2007. Integração de Ontologias: o domínio da bioinformática. RECIIS - *Revista Eletrônica de Comunicação Informação & Inovação em Saúde*, 117 p.
- Domingue, J.1998. *Tadzebao and WebOnto: discussing, browsing, and editing ontologies on the Web*. In: Eleventh Workshop on Knowledge Acquisition, Modeling and Management, Alberta, Canada.
- Franch, X.; Guizzardi, R.; Guizzardi, G.; López, L. 2011. *Ontological Analysis of Means-End Links*. In Anais do 5th International I* Workshop (ISTAR 2011), Trento, Italy.
- Gennari, John H. et al. 2002. *The Evolution of Protégé: An Environment for Knowledge-Based Systems Development*. Stanford Medical Informatics.
- Noy, N.F.; McGuinness, D.L. 2001. *Ontology Development 101: A Guide to Creating Your First Ontology*. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05.
- Guizzardi, G.; Wagner, G. 2004. *On A Unified Foundational Ontology and some Applications of it in Business Modeling*. Open INTEROP Workshop on Enterprise Modelling and Ontologies for Interoperability, at the 16th International Conference on Advances in Information Systems Engineering (CAiSE), Latvia.
- Guizzardi, G.; Wagner, G. 2010. *Using the Unified Foundational Ontology (UFO) as a Foundation for General Conceptual Modeling Languages*. In: *Theory and Application of Ontologies*, Roberto Poli (editor), ed. Berlin: Springer-Verlag.
- Monteiro, F.S. 2006. *Modelagem conceitual: a construção de uma ontologia sobre Avaliação do ciclo de Vida (ACV) para fornecer a disseminação de seus conceitos*. Monografia, curso de biblioteconomia, Universidade de Brasília UnB.
- Minghim, R.; Henrique, L. 2006. *OntoEditor – Uma ferramenta Web para edição de ontologias*. Disciplina: Visualização Computacional.
- Oliveira, L.H.M.; Aluísio, S.M.; Almeida, G.M.B. 2007. *OntoEditor – Uma ferramenta Web para edição de ontologias*. V Workshop em tecnologia da Informação e da Linguagem Humana. Rio de Janeiro, RJ.
- Sure, Y.; Erdmann, M.; Angele, J.; Staab, S.; Stude, R.; Wenke, D. 2002. *OntoEdit: Collaborative Ontology Development for the Semantic Web*. University of Karlsruhe.
- Hepp, M. 2007. *Ontologies: state of the art, business potential, and grand challenges*. Digital Enterprise Research Institute, university of Innsbruck, Technikerstrabe 21 a, A-6020 Innsbruck, Austria.