

# **ODRM (ONTOLOGY DIAMOND RESEARCH MODEL): Um Conjunto Formal de Métodos, Processos e Ferramentais para Estruturar Bases de Conhecimento Pautadas em Ontologia**

**Sonia Limoeiro Monteiro<sup>1</sup>, Hernane Borges B. Pereira<sup>2</sup>, Eduardo M. F. Jorge<sup>3</sup>,  
Uedson Santos Reis<sup>4</sup>, Tiago S. A. Cahyba<sup>5</sup>**

<sup>1</sup>Laboratório Nacional de Computação Científica  
Laboratório Nacional de Computação Científica (LNCC) – Petrópolis, RJ – Brasil

<sup>2</sup>Departamento de Ciências Exatas  
Universidade Estadual de Feira de Santana (UEFS) – Feira de Santana, BA – Brasil

<sup>2 3 4 5</sup>Programa de Modelagem Computacional e Tecnologia Industrial  
SENAI CIMATEC – Salvador, BA – Brasil

<sup>3 4 5</sup>Departamento de Exatas  
Universidade Estadual de Feira de Santana (UEFS) – Feira de Santana, BA – Brasil

slmo@lncc.br, hernanebbpereira@gmail.com,  
emjorge, uedson, cajahyba @reconcavotecnologia.org.br

**Abstract.** *This article presents the ODRM (ONTOLOGY DIAMOND RESEARCH MODEL), a set of formal methods, processes and tools to build databases knowledge based on Ontology. This model is being developed by a group of interdisciplinary study with the aim that more people will benefit from Ontologies reducing the semantic gap between the abstractions of the real world and a Formal Ontology.*

**Keywords:** Knowledge Diffusion, Ontology, Semantic Web, Folksonomy.

**Resumo.** *O presente artigo apresenta o ODRM (ONTOLOGY DIAMOND RESEARCH MODEL), um conjunto formal de métodos, processos e ferramentais para estruturar bases de conhecimento pautadas em Ontologia. Este modelo está sendo desenvolvido por um grupo de estudo interdisciplinar com o objetivo que mais pessoas se beneficiem das Ontologias reduzindo o gap semântico entre as abstrações do mundo real e uma Ontologia Formal.*

**Palavra-Chaves:** Difusão do Conhecimento, Ontologia, Web-Semântica, Folksonomia.

## **1. Introdução**

A busca por métodos de como representar e recuperar informações que possam ser contextualizadas e transformadas em conhecimento é uma temática presente em todos os tempos, mas com enfoque especial na pós-modernidade. O cenário é intrigante, pois na pós-modernidade um homem conectado a um computador tem acesso aos mais diversos sistemas computacionais que potencializam a capacidade de processamento e memória dos seres humanos. Bases de conhecimento gigantescas estão acessíveis contendo milhares de livros, artigos, informações de museus virtuais, jornais, etc. Logo neste cenário a formalização de conceitos é algo necessário, principalmente quando um grupo de indivíduos deseja compartilhá-los e também para fornecer mecanismos de inteligência às máquinas.

Um ponto focal é entender o conhecimento como algo difícil de estruturar, de ser capturado por máquinas e de difícil transferência, já que o conhecimento está relacionado diretamente com o homem e suas experiências cotidianas e ao meio ambiente que ele está inserido. Portanto, a habilidade de representação das abstrações em signos ou conceitos é algo inerente a forma como a mente humana estrutura os seus pensamentos. Extraordinariamente em uma criança por volta de dois anos esta habilidade já pode ser notada como resultado da sua atividade mental. (VIGOTSKI, 2007)

Neste panorama brilha a demanda por formas e estratégias de como modelar as representações mentais em conceitos até transformá-las em estruturas computacionais aplicadas no contexto da representação e recuperação da informação. A pesquisa por estruturas de representação do conhecimento em conceitos varia quanto ao nível de formalidade. Existe aí um paradoxo, visto que quanto mais formal for a linguagem melhor para compreensão da máquina, porém os métodos formais se afastam da linguagem natural do homem. Observa-se que comumente a atividade de modelagem de conceitos (com o maior nível de formalização) para a representação em estruturas computacionais são áreas de pesquisa da Inteligência Artificial, da Ciência da Informação e da Ciência da Computação. O enfoque principal dessas áreas é na formalização dos conceitos objetivando a manipulação dos mesmos pelos computadores.

O foco deste projeto busca um aprofundamento nos mecanismos formais para a construção de estruturas computacionais de base de conhecimento pautados em Ontologia. Redes Semânticas, Mapas Conceituais, Árvores de Conhecimento, Ontologias são alguns exemplos de estruturas disponíveis para a representação de conceitos (JANASSEN, et al., 1993). Dentre essas estruturas, o foco do projeto está nas Ontologias devido ao seu grau de formalismo. Ontologia pode ser definida como “especificação explícita e formal de um conceito compartilhado” (GRUBER, 1993), é o estado da arte para a interoperabilidade de conceitos, retirando dos sistemas a propriedade das abstrações e garantindo uma visão mais epistemológica do domínio a ser representado. Outro ponto a ser ressaltado é que através das Ontologias existe a possibilidade da definição de axiomas sobre os elementos conceituais, determinando restrições sobre sua representação em determinado domínio.

A construção de bases de conhecimento pautadas em Ontologias transcende a junção de dados e a dificuldade de lidar com dados em estruturas indefinidas, demandando uma taxonomia unificada dos conceitos de uma organização ou área de conhecimento. Observa-se que uma estrutura de ontologia para possibilitar a difusão do conhecimento depende da compreensão do usuário de uma área de conhecimento. O objetivo da pesquisa é a especificação do ODRM (*Ontology Diamond Research Model*). O ODRM consiste em um conjunto formal de métodos, processos e ferramentas baseados em teorias das áreas da: Psicologia Cognitiva, Ciência da Informação, Ciência da Computação, Inteligência Artificial, Gestão do Conhecimento Organizacional entre outras para estruturar bases de conhecimento pautadas em Ontologia. Os padrões são aplicados para a construção de aplicações em base de conhecimento. Um princípio fundamental do ODRM é a externalização de conceitos, permitindo não só o compartilhamento dos mesmos, mas o reaproveitamento em sistemas de várias naturezas e a ampliação do espaço semântico de organizações ou grupos de indivíduos. Portanto, o ODRM visa subsidiar, por exemplo, modelagens de Ontologias Organizacionais, sistemas de mineração textual de informações que realizam consultas baseados em Ontologias, montagem de sites com buscas semânticas, portais de gerenciamento de processo de negócio fundamentados em Ontologias, entre outros. As aplicações enumeradas constituem ambientes para a Difusão do Conhecimento.

## **2. Trabalhos em Andamento**

Um dos primeiros trabalhos do grupo foi a pesquisa visando resgatar o conceito de Ontologia na Filosofia e na Epistemologia, e analisar o processo de incorporação do termo pela Ciência da Computação. Buscou-se contribuir para o esclarecimento terminológico do termo, bem como sua utilização e aplicação. Esta pesquisa gerou o artigo “Uma visão histórico-epistemológica do conceito Ontologia: Desvelando a Ontologia na Filosofia e na Ciência da Computação” publicado no 2º Workshop da Modelagem Computacional da Difusão do Conhecimento no LNCC (Laboratório Nacional de Computação Científica) localizado em Petrópolis em setembro de 2008.

Outra iniciativa foi o projeto *Ontology Tagging* desenvolvido com o intuito de simplificar a manipulação de uma Ontologia pelo usuário na classificação e recuperação de conteúdo na web. A proposta foi a conversão de uma Ontologia em uma metáfora visual, mais especificamente uma nuvem de tags (etiquetas ou rótulos), para ser utilizada dentro do contexto da Folksonomia (VANDER WAL, 2007). Cada conceito da Ontologia é utilizado como tag sugestiva para classificação, armazenamento e recuperação de qualquer tipo de conteúdo, juntamente com outras tags, que podem ser inseridas manualmente pelo usuário. A técnica sugerida visa à utilização conjunta das tags comuns e ontológicas (extraídas da ontologia) para minimizar os

pontos negativos e realçar os pontos positivos de ambas as tags. O usuário pode utilizar a tag ontológica para atribuir um conceito mais formal e compartilhado ao seu conteúdo e uma tag digitada por ele mesmo para atribuir uma semântica própria. Além disso, as tags digitadas pelo usuário podem ser úteis no momento de retroalimentação ou atualização de uma ontologia de domínio, implementando os conceitos de conhecimento colaborativo (HENDLER e GOLBECK, 2007). Esse trabalho foi publicado na ERBASE 2009 no mês de Maio através de um artigo intitulado “*Ontology Tagging – Um componente para Integração de Folksonomia com Ontologia*” (REIS, et al., 2009).

Como continuação do projeto *Ontology Tagging* surgiu o *Cumulus*. Um projeto que trata do desenvolvimento de uma solução computacional que contempla o armazenamento e a recuperação de arquivos utilizando *tags*, e aborda também modelos baseados no contexto semântico em que o conteúdo está inserido. A pesquisa discute métodos, fundamentos e técnicas de recuperação e representação de informação. Foram desenvolvidos algoritmos que integram o componente *Ontology Tagging* e os módulos de busca e de *upload*, assim como métodos que permitem ao *Cumulus* gravar *tags* derivadas de uma ontologia na base de dados, associar uma ou mais *tags* a um arquivo e transferi-lo a um disco remoto. A solução apresentada constitui-se em uma extensão da metáfora visual proposta pelo componente *Ontology Tagging* ao caracterizar cada *tag* proveniente de uma ontologia como um conceito formal e compartilhado, formando um metadado semântico.

Na área educacional o grupo publicou o artigo “*WebQuest Manássio: Uma Técnica de Ensino Baseada na Integração de WebQuest e Web Semântica*” no evento WEIBASE Workshop de Educação em Computação e Informática Bahia-Sergipe que ocorreu na UESC (Universidade Estadual de Santa Cruz) em Maio de 2009 (OLIVEIRA, et al., 2009). O projeto que gerou o artigo construiu um solução educacional denominada *WebQuest Manássio* que combina as técnicas de *WebQuest* e *Web Semântica*, permitindo que o professor construa um *WebQuest* semântico. A relevância está em apresentar ao aluno conteúdos de um determinado domínio, através da técnica *WebQuest*, onde determinados termos são descritos através de uma Ontologia. Ressalta-se que o artigo ganhou um prêmio como o melhor trabalho do evento *WEIBASE*.

O grupo também tem atuado na formação e disseminação de conteúdos envolvendo Ontologia ministrando disciplinas e minicursos com alunos das mais diversas formações. Seguem as realizações mais recentes:

- Disciplina de Sistemas de Representação do Conhecimento oferecida em no DMMDC Doutorado de Doutorado Multi-Institucional e Multidisciplinar em Difusão do Conhecimento em 2008.2
- Minicurso em Ontologias: Sistemas de Representação de Conhecimento oferecido no curso de verão em Fevereiro de 2009 no LNCC em Petrópolis.
- Mini curso sobre Modelagem de Ontologia na Escola Regional de Computação Bahia-Sergipe em Maio de 2009 na UESC.

### **3. Contribuições Esperadas**

Na sociedade pós-moderna, conhecida também pela sociedade da informação, a demanda por técnicas de como estruturar conhecimento é algo essencial. As pesquisas do grupo trabalham os fundamentos de como elaborar e materializar conceitos em estruturas de conhecimento com mecanismos similares a cognição humana, visando armazenamento e disseminação de informação semanticamente organizada. Para isso adotam-se estratégias para dar condições as máquinas a ajudar os indivíduos, possibilitando que decisões possam ser aferidas através de sistemas pautados em base de conhecimento. Os projetos desenvolvidos por este grupo visam atuar na difusão do conhecimento possibilitando que mais pessoas se beneficiem das Ontologias, reduzindo o *gap* semântico entre as abstrações do mundo real e uma Ontologia formal. A linguagem formal apesar de seus benefícios para uso em ambientes computacionais, distancia-se da cognição humana e essa característica dificulta o processo de desenvolvimento e manipulação das Ontologias. O desenvolvimento de ferramentas que aproximem os conceitos e o formalismo da Ontologia do

processo de cognição humana é abordado pelo grupo, assim como aplicações práticas que comprovam a eficiência e a eficácia do uso da Ontologia na difusão do conhecimento.

#### **4. Experiência e Perfil do Grupo**

O grupo conta com:

- Sonia Limoeiro Monteiro, Doutora em Informática e Mestre em Matemática;
- Hernane Borges de Barros Pereira, Mestre em Software, e Doutor e Mestre em Engenharia Multimídia;
- Eduardo Manuel de Freitas Jorge, Doutorando em Difusão do Conhecimento e Mestre em Ciência da Computação;
- Tiago Silveira de Andrade Cahyba, Mestrando em Modelagem Computacional, Pós-Graduado em Gestão Empresarial e Graduado em Sistemas de Informação;
- Uedson Santos Reis, Mestrando em Modelagem Computacional e Graduado em Sistemas de Informação;
- Cássio Silva Oliveira, Graduado em Sistemas de Informação;
- Manuella Almeida, Graduada em Sistemas de Informação;
- Tatiana Simas Silva, Graduada em Sistemas de Informação;
- Ronaldo Pereira Evangelista, Graduando em Sistemas de Informação.

#### **5. Direções de Pesquisa**

Neste momento o ODRM é o projeto macro que rege as pesquisas do grupo. Como já supracitado este representa um conjunto formal de métodos e processos para estruturar bases de conhecimento pautadas em Ontologia, podendo subsidiar ou apoiar, por exemplo:

- Modelagens de Ontologias Organizacionais;
- Montagem de sistemas de mineração textual de informações que realizam consultas baseados em Ontologias;
- Montagem de sites com buscas semânticas;
- Portais de gerenciamento de processo de negócio fundamentados em Ontologias;
- Construção de soluções educacionais que se beneficiam dos conceitos de Ontologia.

A modelagem de uma base de conhecimento pautada em Ontologia, que mantenha e manipule dados estruturados e semi-estruturados é uma dessas pesquisas. Dois dos integrantes do grupo, Tiago Cahyba e Uedson Reis, estão desenvolvendo suas dissertações de mestrado abordando este tema. O trabalho de Tiago Cahyba tem por objetivo a criação um componente capaz de, a partir de um modelo conceitual único (por exemplo, uma ontologia que defina o negócio de uma organização), realizar buscas em estruturas de dados armazenadas em banco de dados relacionais ou modeladas em forma de data warehouse. O desafio é realizar conversões entre o padrão de consulta em ontologia (SparQL) e o padrão de consulta em bases de dados relacionais (SQL) ou data warehouse (DMX). projeto de Uedson Reis tem por objetivo modelar uma solução computacional para armazenamento e recuperação semântica de conteúdo, onde serão incluídos dados semi-estruturados. O modelo de consulta desenvolvido no projeto de Reis será acoplado ao projeto de Cahyba para que juntos contemplem uma base de conhecimento capaz de lidar com buscas semânticas em dados com estruturas diversas. A base conhecimento resultante permitirá buscas mais assertivas não só pautadas na sintática. Assim, esses trabalhos buscam a melhor interação com a Ontologia através de metáforas visuais, baseadas em nuvens, grafos baseados em mapas conceituais ou redes semânticas.

Os dois projetos citados dependem da construção previa de Ontologias para guiar as suas relações semânticas. Porém foi identificada uma problemática no processo de formalização de conceitos durante a criação de Ontologias. Através da aplicação de cursos com conteúdos envolvendo Ontologias para alunos de diversas formações foi constatada uma dificuldade no processo de modelagem de conceitos com os métodos atuais. Assim, está se buscando definir ou identificar métodos que possam suavizar a complexidade do processo de modelagem e construção de Ontologias.

Nos principais métodos de modelagem, a construção de Ontologias tem como meta a identificação de classes e seus relacionamentos. A classe é uma forma natural de agrupar os objetos (indivíduos, coisas, etc.) que possuem as mesmas características. Assim, um conjunto de objetos do mesmo tipo é denominado de classe. As metodologias existentes (LAL, Kactus, Methontology, Método 101, Cyc, Uschold, etc.) têm diferenças, mas um princípio comum a todas é a identificação de classes, seus relacionamentos e axiomas para posteriormente identificar objetos ou indivíduos (BREITMAN, 2005).

Ciente disso, as questões que nortearam o surgimento de um novo projeto, desenvolvido por Eduardo Manuel de Freitas Jorge em sua tese de doutorado, foram: Por que, apesar da existência de uma vasta bibliografia de um conjunto de metodologias e ferramental, a modelagem de conceitos formais, como o de uma ontologia, não é algo trivial? Os métodos atuais podem ser melhorados? A identificação de objetos e suas relações é um processo mais próximo da cognição humana do que a identificação de classes e suas relações?

Analisa-se, assim, a possibilidade de inverter o método de modelagem baseado em classes, identificando primeiramente um subconjunto de relação entre objetos. Através desse subconjunto seria assim possível determinar por inferência as relações e alguns axiomas das futuras classes. O cerne é analisar como um subconjunto de relações entre instâncias poderia determina a regra de relacionamento entre todas as instâncias. A proposta é a criação do método M-MOBI (Método de Modelagem de Ontologias Baseado em Instâncias) que visa verificar se este processo aumenta o grau de simplicidade na criação, organização, gestão e difusão de ontologias. Algumas questões serão alisadas tais como: É possível reutilizar as ligações existentes entre objetos na modelagem de novas Bases de Conhecimento? Quais inferências são possíveis de realizar a partir da identificação das instâncias e suas relações?

Por fim, espera-se que estes projetos suscitem novos e frutíferos debates sobre Ontologia e conhecimento na sociedade contemporânea.

## 6. Bibliografia

**BREITMAN Karin.** Web Semântica a Internet do Futuro [Livro]. - Rio de Janeiro : LTC, 2005.

**GRUBER T.** A Translation Approach to Portable Ontology Specifications In Knowledge Acquisition. - 1993.

**HENDLER James e GOLBECK Jennifer.** Metcalfe's Law, Web 2.0, and the Semantic Web. - Novembro de 2007.

**JANASSEN David H., BEISSNER Katherine e Michael YACCI** Structural Knowledge Technique for Representing, Conveying, and Acquiring Structural Knowledge . - New Jersey : LEA, 1993.

**OLIVEIRA Cássio, DOMINIQUE Manuella e JORGE Eduardo** WebQuest Manássio: Uma Técnica de Ensino Baseada na Integração de WebQuest e Web Semântica = WebQuest Manássio: Uma Técnica de Ensino Baseada na Integração de WebQuest e Web Semântica. - Ilhéus : ERBASE - SBC, 2009.

**REIS Uedson [et al.]** Ontology Tagging - Um Componente para Integração de Ontologia com Folksonomia. - Ilhéus : [s.n.], 2009.

**VANDER WAL Thomas.** Folksonomy Coinage and Definition [Online] // Vanderwal.net. - Fevereiro de 2007. - Agosto de 2008. - <http://vanderwal.net/folksonomy.html>.

**VIGOTSKI Lev Semenovitch** A formação social da mente: o desenvolvimento dos processos psicologicos superiores / trad. Luís José Cipolla Neto. - São Paulo : [s.n.], 2007. - 7.