

**CENTRO UNIVERSITÁRIO FEEVALE**  
**INSTITUTO DE CIÊNCIAS EXATAS E TECNOLÓGICAS**  
**CURSO DE CIÊNCIA DA COMPUTAÇÃO**

**Sistema de Inferência Baseado em Ontologia**

(Título Provisório)

por

DIEGO CIRINO KERN  
diegoc\_k@yahoo.com.br

**Anteprojeto de Trabalho de Conclusão**

Rodrigo Rafael Villarreal Goulart  
rodrigo@feevale.br

Novo Hamburgo, Setembro de 2006

## Sumário

Dados de Identificação .....	3
Resumo .....	4
Motivação .....	5
Objetivos .....	10
Metodologia .....	11
Cronograma .....	12
Bibliografia .....	13

## **Dados de Identificação**

**Área de Estudo:** Inteligência Artificial.

**Título provisório do trabalho:** Sistema de inferência baseado em ontologia

**Orientador(a):** Ms. Rodrigo Rafael Villarreal Goulart

### **Identificação do aluno:**

Nome: Diego Cirino Kern

Telefones:

Celular: (51) 96476005

Residencial: (51) 3561 - 1052

E-mail: diegoc\_k@yahoo.com.br

## Resumo

A World Wide Web, ou simplesmente Web, se tornou uma excelente fonte de pesquisa, devido ao elevado número de informações que ela disponibiliza. Porém, a maioria desses dados não possui uma estrutura adequada. Frente a este cenário Tim Berners Lee propôs o termo Web Semântica. Essa seria uma extensão da Web atual com o objetivo de estruturar e dar significado a informação nela inserida, desta maneira, aumentar a capacidade de cooperação dos computadores com as pessoas. As ontologias vem se destacando como forma de estruturar de forma organizada as informações de um determinado domínio de conhecimento, objetivando um entendimento semântico de situações do mundo real. Numa visão filosófica, o termo ontologia foi proposto por Aristóteles, onde tratava da natureza e da organização do ser. Desde a década de 90 esse termo também vem sendo adotado pela inteligência artificial. Gruber propôs uma das principais conceitualizações sobre o tema, dizendo que uma ontologia é a especificação formal, explícita e compartilhada de uma conceitualização. A W3C (World Wide Web Consortium) é uma organização que trabalha para desenvolver padrões para a criação e interpretação para os conteúdos da web, e recomenda que as pessoas utilizem a linguagem OWL para estruturar suas ontologias, esperando-se que esta linguagem torne-se um padrão. Desta maneira, faz-se necessário o uso de ferramentas que sejam capazes de inferir sobre essas ontologias, extraindo conhecimento. Com este objetivo, serão pesquisados e avaliados mecanismos de inferência para que uma proposta de sistema seja modelada, desenvolvida e avaliada.

**Palavras Chaves: Ontologia, OWL, web semântica e inteligência artificial.**

## Motivação

Há alguns anos ocorre um grande aumento do número de fontes de informação disponibilizadas através de meios digitais em grandes repositórios, assim como a web. Assim, novas formas de representação e recuperação de informação são cada vez mais necessárias, fundamentando pesquisas em diversas áreas da ciência, como a ciência da computação, a ciência da informação e a lingüística.

Atualmente o termo ontologia vem se destacando no que tange as ciências anteriormente citadas. A sua utilização visa estruturar de forma organizada as informações de um determinado domínio de conhecimento e refletir um entendimento semântico de situações do mundo real.

Numa visão filosófica, o termo ontologia foi proposto por Aristóteles, onde tratava da natureza e da organização do ser, onde também eram utilizadas propriedades para diferenciar espécies do mesmo gênero. Desde o início da década de 90 esse termo também vem sendo adotado pela área da inteligência artificial, e para ele são dadas diversas definições.

Uma das principais conceitualizações sobre o tema diz que uma ontologia é a especificação formal, explícita e compartilhada de uma conceitualização (GRUBER, apud ARAUJO, 2003).

“Uma ontologia é um artefato de engenharia, constituído de um vocabulário de termos organizados em uma taxonomia, suas definições e um conjunto de axiomas formais usados para criar novas relações e para restringir as suas interpretações segundo um sentido pretendido” (NOY, HAFNER, apud GUIZZARD, 2000, p.38).

Para a construção de uma ontologia não há uma metodologia padronizada a seguir. A maneira como ela será escrita dependerá da alternativa escolhida por quem a estiver construindo.

De acordo com (GUIZZARD, 2000), existem algumas propostas de metodologias para a construção de ontologias, mas nenhuma é estruturada de forma adequada. As metodologias de Noy & McGuinness (2000); Gruninger e Fox (2005);

Uschold e King (1995); Uschold e Gruninger (1996) formam-se na mesma proposta e são as principais para a elaboração de ontologias.

Dentre essas metodologias propostas na literatura, uma metodologia é apresentada por Falbo (1998), na qual ele relata como essenciais atividades como identificar as questões que deverão ou não serem respondidas e executadas pela ontologia, capturar os conceitos e relações do domínio a ser representado, construir taxonomias e definir axiomas, avaliar se a ontologia está de acordo com os requisitos estabelecidos e documentar os processos envolvidos na metodologia de desenvolvimento, como os propósitos da ontologia, seus requisitos e seus cenários de motivação.

Um exemplo do uso de metodologias é a ontologia Ontomúsica, proposta por Boff (2005). A Ontomúsica tem como principal objetivo, viabilizar conhecimento sobre a história da música na Web, através da estrutura de ontologias. Informações sobre o domínio proposto são disponibilizadas para que, por exemplo, sistemas inteligentes apresentem questionamentos e dicas sobre o conteúdo. A metodologia utilizada foi proposta por Noy & McGuinness (2000).

Essa ontologia é composta por um relacionamento de classes, atributos e instâncias que descrevem períodos, compositores, obras, gêneros musicais entre outros assuntos sobre a história da música. As classes da ontologia descrevem determinados períodos da História da Música, a estes períodos relacionam-se compositores, aos quais ficam ligadas suas obras musicais, bem como os seus respectivos gêneros musicais. As obras possuem características conforme o período em que foram criadas.

Para modelar a ontologia foi utilizada a linguagem OWL (*Ontology Web Language*) (SMITH, WELTY & MCGUINNESS, 2003), linguagem recomendada pela W3C (*World Wide Web Consortium*), que disponibiliza mecanismos para representar explicitamente o significado dos termos e os relacionamentos entre os mesmos, descrevendo características especiais sobre os conceitos e os relacionamentos através de axiomas lógicos. É uma linguagem para ontologias Web que ao invés de apenas apresentar informações aos usuários também processa o conteúdo dessas informações, sendo utilizada quando informações de um determinado documento precisam ser processadas por aplicações. A OWL pode ser entendida como uma linguagem para

definição e instanciação de ontologias Web, sendo que estas devem incluir descrições de classes, propriedades e suas instâncias.

A construção e a edição da Ontomúsica foi realizada com a utilização da ferramenta Protégé-2000<sup>1</sup>. O desenvolvimento de ontologias no Protégé basicamente consiste na definição de um esquema com as classes (class), subclasses (subclass), propriedades e relações (slots) referentes ao domínio que se deseja modelar (FELICÍSSIMO et al, apud BOFF, 2003). Essa ferramenta, além da construção de ontologias também possibilita a entrada de dados e a criação de *Queries* para futuras consultas na ontologia. Ela também disponibiliza um plugin para conversão da ontologia em OWL.

A World Wide Web é uma excelente fonte de pesquisa, devido ao elevado número de informações que disponibiliza. No entanto, essas informações não possuem uma estrutura e nem seguem uma padronização adequada. Nesse contexto é que surge o termo Web Semântica.

Tim Berners Lee, idealizador da World Wide Web, também foi quem inicialmente propôs o termo Web Semântica. Segundo Berners, a definição para esse termo é a seguinte:

“A Web Semântica não é uma Web separada, mas uma extensão da atual, na qual a informação é utilizada com significado bem definido, aumentando a capacidade dos computadores para trabalharem em cooperação com as pessoas” (BERNERS, HENDLER & LASSILA, 2001).

A W3C é uma organização que trabalha para desenvolver padrões para a criação e interpretação para os conteúdos da web. Esses padrões visam estruturar e dar significado para o conteúdo distribuído atualmente na Internet, fazendo com que a informação que hoje é compreendida por humanos, também seja compreensível para os computadores.

Como já foi dito anteriormente, cresce cada vez mais o uso de ontologias para a organização e padronização desses dados. A W3C recomenda que os desenvolvedores

---

<sup>1</sup> <http://protege.stanford.edu>

utilizem a linguagem OWL para estruturar suas ontologias, esperando-se assim que esta linguagem torne-se um padrão.

Desta maneira surge a necessidade de mecanismos de inferência que utilizem ontologias modeladas em OWL. Para isso, será proposto um sistema que infere novas perguntas, utilizando a ontologia proposta por Boff. Com este objetivo, serão pesquisados e avaliados mecanismos de inferência para que então uma proposta de sistema seja modelada, desenvolvida e avaliada.

Dentre os mecanismos, será estudado o Jena, que é uma API para a linguagem de programação Java, desenvolvida por Brian McBride da *Hewlett-Packard* (JENA, 2004), ela permite o desenvolvimento de aplicações baseadas em Web Semântica.

Na sua primeira versão, a API Jena apenas possuía métodos para a manipulação de ontologias em DAML+OIL (*Darpa Agent Markup Language + Ontology Inference Layer*), porém na sua segunda versão foi adicionado um pacote específico para a manipulação de ontologias: *API Jena 2 Ontology*.

Nesta API uma determinada ontologia é transformada em um modelo abstrato de dados orientado a objetos, fazendo com que seus termos possam ser manipulados como objetos. Uma grande vantagem em se transformar ontologias em modelos orientados a objetos é que a programação orientada a objetos pode ser utilizada, fazendo com que as manipulações nestes modelos tornem-se comuns para os programadores Java, por exemplo.

Um dos módulos presentes na arquitetura do Jena é o *Reasoner*. Este permite fazer inferências sobre modelos OWL. O uso das inferências sobre modelos semânticos é permitir obter informação adicional (inferida) sobre as ontologias. O Jena fornece alguns motores de inferência pré-construídos e possibilita a criação de novos motores quando necessário, ou a possibilidade de estender os já existentes.

Outros sistemas de representação de conhecimento que serão pesquisados e avaliados são: *RacerPro*<sup>2</sup>, o *FaCT++*<sup>3</sup>. Esses sistemas seguem o padrão DIG

---

<sup>2</sup> <http://www.racer-systems.com/>

<sup>3</sup> <http://owl.man.ac.uk/factplusplus/>



(*Description Logic Interface*), o qual visa padronizar a maneira de interação das ferramentas clientes com os diferentes reasoners existentes (DICKINSON, 2004).

## **Objetivos**

### **Objetivo geral**

Desenvolver um sistema de inferência baseado em ontologias e realizar um estudo de caso.

### **Objetivos específicos**

- Pesquisar na literatura definições e pesquisas relacionadas com a construção de ontologias (incluindo OWL e Protegé) e mecanismos de inferência baseados em ontologias;
- Avaliar o trabalho desenvolvido por Rogeiro Boff com a ontologia Ontomusica (estudo de caso);
- Avaliar os mecanismos de inferência;
- Desenvolver a proposta de sistema em Java que utilize o trabalho proposto por Rogério Boff como estudo de caso;
- Implementar o sistema proposto;
- Desenvolver e avaliar experimentos com a utilização do sistema proposto;

## **Metodologia**

Primeiramente será feita uma pesquisa na bibliografia existente, com o objetivo de constituir um embasamento teórico para o desenvolvimento do trabalho. Procurando por definições relacionadas às ontologias e sobre os mecanismos de inferência baseados em ontologias. (1)

Numa segunda etapa será avaliado o trabalho feito por Rogério Boff, onde foi desenvolvida a Ontomusica, uma ontologia sobre a história da música. (2)

No passo seguinte serão avaliadas as ferramentas de inferência sobre ontologias, sendo elas: Jena, RacerPro, FaCT++. (3)

Após a avaliação das ferramentas de inferência será desenvolvida a proposta de sistema em Java que utilize o trabalho proposto por Rogério Boff como estudo de caso. (4)

Com a proposta do sistema desenvolvida o mesmo será implementado. (5)

Por fim, com o sistema concluído, serão desenvolvidos e avaliados experimentos com a utilização do sistema proposto. (6)

## Cronograma

### Trabalho de Conclusão I

Etapa	Meses				
	Ago	Set	Out	Nov	Dez
Elaboração e entrega do anteprojeto					
Pesquisa bibliográfica dos conceitos relacionados a ontologias. (1)					
Avaliar o trabalho de Rogério Boff. (2)					
Estudar as ferramentas de inferência. (3)					

### Trabalho de Conclusão II

Etapa	Meses				
	Mar	Abr	Mai	Jun	Jul
Desenvolver proposta do sistema em Java. (4)					
Desenvolver o sistema proposto. (5)					
Desenvolver e avaliar experimentos com o sistema. (6)					

## Bibliografia

ARAÚJO, Moysés de. **Educação à distância e a web semântica: Modelagem ontológica de materiais e objetos de aprendizagem para a plataforma COL**. São Paulo, 2003. 173 f. Tese para obtenção do título de doutor em engenharia – Escola Politécnica da Universidade de São Paulo.

BERNERS-LEE, T.; HENDLER, J. and LASSILA, O. (2001) **The semantic web**. Scientific American, May 2001. Disponível em: <http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21&pageNumber=2&catID=2>. Acesso em: 18 Ago. 2006.

BOFF, Rogério Eduardo. **Educação Musical à Distância Utilizando Ontologias**. 2005. Projeto de Diplomação (Bacharelado em Ciências da Computação) – Instituto de Ciências Exatas e Tecnológicas (ICET), Centro Universitário FEEVALE, Novo Hamburgo.

DICKINSON, Ian. 2004. **Implementation experience with the DIG 1.1 specification**. Digital Media Systems Laboratory, Bristol.

FALBO, Ricardo de Almeida. **Integração de Conhecimento em um Ambiente de Desenvolvimento de Software**. 1998. 198 f. Tese (Doutorado em Ciências em Engenharia de Sistemas e Computação) – UFRJ, Rio de Janeiro.

GRUNINGER, Michael; FOX, Mark S. **Methodology for the Design and Evaluation of Ontologies**. In: INTERNATIONAL JOINT CONFERENCE ON AI, IJCAI, 1995. **Proceedings...** Montreal: 1995. Disponível em: <http://citeseer.ist.psu.edu/grninger95methodology.html> . Acesso em: 18 Ago. 2006.

GUIZZARD, Giancarlo. **Uma Abordagem Metodológica de Desenvolvimento para e com Reuso, Baseada em Ontologias Formais de Domínio**. 2000. 148 f. Dissertação (Mestrado em Informática) – Universidade Federal do Espírito Santo, UFES, Vitória.

JENA - A Semantic Web Framework for Java (2004) “Jena 2 Ontology API”, <http://jena.sourceforge.net/ontology/>. Acesso em: 18 Ago. 2006.

NOY, Natalya F.; MCGUINNESS, Deborah L. **Ontology Development 101: A Guide to Creating Your First Ontology**. [S.l.: s.n.], 2000. Disponível em: <http://www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness-abstract.html>. Acesso em: 18 Ago. 2006.

Protégé-2000. Disponível em <<http://protege.stanford.edu/>>. Acessado em: 19 de Ago. 2006.

SMITH, M. K., WELTY, C. and MCGUINNESS, D. L. (eds.) (2003) **OWL Web Ontology Language Guide – W3C Candidate Recommendation 18 August 2003**. Disponível em: <http://www.w3.org/TR/owl-guide/>. Acesso em: 12 abr. 2006.

USCHOLD, Mike. **Building Ontologies: Towards a Unified Methodology**. In: ANNUAL CONFERENCE OF THE BRITISH COMPUTER SOCIETY SPECIALIST GROUP ON EXPERT SYSTEMS, 16., 1996. **Proceedings...** Cambridge: [s.n], 1996. Disponível em: < <http://citeseer.ist.psu.edu/uschold96building.html> >. Acesso em: 18 Ago. 2006.

USCHOLD, Mike; GRUNINGER, Michael. **Ontologies: Principles, Methods and Applications**. Knowledge Engineering review, [S.l.], v.11, n.2, p.93-155, 1996. Disponível em: <http://citeseer.ist.psu.edu/uschold96ontologie.html>. Acesso em: 18 Ago. 2006.