



**Universidade de Brasília - UnB**  
**Faculdade de Economia, Administração, Contabilidade e**  
**Ciência da Documentação e Informação - FACE**  
**Departamento de Ciência da Informação e**  
**Documentação - CID**  
**Curso de Graduação em Biblioteconomia**

## **ONTOLOGIA E WEB SEMÂNTICA**

**Eliana Carlan**

**Brasília, junho de 2006**

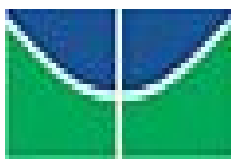
**Eliana Carlan**

## **ONTOLOGIA E WEB SEMÂNTICA**

Monografia apresentada ao Departamento de Ciência da Informação e Documentação, como requisito para obtenção do grau de Bacharel em Biblioteconomia.

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Marisa Bräscher Basilio Medeiros

**Brasília, junho de 2006**



Universidade de Brasília  
Campus Universitário Darcy Ribeiro – Asa Norte  
CEP 70910-900  
Brasília – DF – Brasil

---

Monografia apresentada ao Departamento de Ciência da  
Informação e Documentação, como requisito para obtenção  
do grau de Bacharel em Biblioteconomia.

Brasília, 20 de julho de 2006.

Aprovada por:

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Marisa Bräscher Basilio Medeiros  
(CID/UnB) - Orientadora

---

Prof<sup>a</sup>. Mestre Ciência da Informação Fabíola de Luca Coimbra  
(CID/UnB) - Membro

---

Prof. Mestre Sistemas e Computação Marcio de Carvalho Victorino  
(SI/UNIEURO) - Membro

Brasília, julho de 2006

Este trabalho é dedicado aos meus filhos  
Paulo, Pedro e Laura, por existirem.  
Com todo o meu amor!

## **Agradecimentos**

A minha orientadora, Professora Marisa, pelo incentivo e pela oportunidade de realizar este trabalho, e também pela valiosa contribuição acadêmica.

A todos os professores do CID, pelos ensinamentos.

A Deus pela força espiritual que muitas vezes me impulsionou para vencer as dificuldades.

Ao meu marido João, pelo apoio, compreensão, amor e por estar do meu lado nesta jornada. Aos meus filhos Paulo, Pedro e Laura, pela compreensão nos momentos que faltei para que este trabalho pudesse ser realizado. Aos meus pais, Mariano e Irene, pela minha vida e por tudo que me ensinaram. As minhas irmãs, Kika e Marisa pelo incentivo acadêmico e pelos bons momentos que passamos juntas.

A todos os amigos, sem citar nomes, pelas alegrias e pelos momentos de desabafo.

**“A melhor maneira de se entender o futuro é entender o presente”.**  
**J. Naisbitt**

## **Resumo**

Os estudos sobre ontologia e web semântica têm recebido cada vez mais importância em diversas áreas do conhecimento. A proposta deste trabalho é produzir um texto introdutório através de um levantamento bibliográfico, reunindo em um único documento importantes conceitos básicos sobre ontologia e web semântica e as tecnologias utilizadas para as suas implementações. Destacando a importância das ontologias dentro da web semântica e sua aplicabilidade.

### **Palavras-Chave:**

Ontologia; Web Semântica

## **Abstract**

The studies about ontology and Semantic Web have been increasing recently and have achieved more importance in many knowledge fields. The purpose of this work is produce an introductory text through a bibliographic survey including in only one document important basics concepts about ontology and Semantic Web, and the tecnologies for their implamentation. Highlighting the importance of ontologies into Semantic Web and their applicability.

### **Keywords:**

Ontology, Semantic Web,



## **Lista de figuras**

Figura 1 – Tipos de ontologias.....	23
Figura 2 – Construção de ontologias.....	32
Figura 3 – Arquitetura da Web Semântica.....	44

## Lista de abreviaturas e siglas

<b>DARPA</b>	Defense Advanced Research Projects Agency
<b>DAML</b>	DARPA Agent Markup Language
<b>DL</b>	Description Logic
<b>Flogic</b>	Frame Logic
<b>FTP</b>	File Transfer Protocol
<b>HTML</b>	HyperText Markup Language
<b>HTTP</b>	HyperText Transfer Protocol
<b>JOE</b>	Java Ontology Editor
<b>NS</b>	NameSpace
<b>OIL</b>	Ontology Inference Layer
<b>OWL</b>	Web Ontology Language
<b>RDF</b>	Resource Description Framework
<b>RDFS</b>	Resource Description Framework Schema
<b>SGML</b>	Standard Generalized Markup Language
<b>SHOE</b>	Simple HTML Ontology Extensions
<b>URI</b>	Uniform Resource Identifier
<b>URL</b>	Uniform Resource Locator
<b>W3C</b>	World Wide Web Consortium
<b>Web</b>	World Wide Web
<b>www</b>	world wide web
<b>XML</b>	eXtensible Markup Language
<b>XMLS</b>	eXtensible Markup Language Schema

# Sumário

1 Introdução.....	12
2 Problema.....	13
3 Objetivo Geral .....	14
3.1 Objetivos Específicos.....	14
4 Justificativa.....	15
5 Metodologia.....	16
6 Ontologia.....	17
6.1 O que é uma ontologia?.....	17
6.2 Componentes de uma ontologia.....	20
6.3 Tipos de ontologias.....	22
6.3.1 Quanto à função.....	24
6.3.2 Quanto ao grau de formalismo.....	24
6.3.3 Quanto à aplicação.....	25
6.3.4 Quanto ao conteúdo.....	25
6.4 Critérios para a construção de ontologias.....	26
6.5 Porque desenvolver uma ontologia.....	28
6.6 Metodologia para a construção de ontologias.....	30
6.7 Ferramentas para a construção de ontologias.....	34
6.8 Linguagens para a construção de ontologias.....	35
6.9 Vantagens no uso de ontologias.....	38
6.10 Problemas no uso de ontologias.....	39
7 Web Semântica.....	41
7.1 Histórico.....	41
7.2 Arquitetura da Web Semântica.....	42

7.2.1 A camada Unicode e URI.....	44
7.2.2 A camada XML, <i>Namespace (NS)</i> , <i>XML Schema</i> .....	45
7.2.3 A camada RDF e <i>RDF Schema</i> .....	46
7.2.4 A camada de ontologia.....	47
7.2.5 A camada lógica.....	49
7.2.6 A camada de prova.....	49
7.2.7 A camada de confiança e assinatura digital.....	50
.	
8 Agentes.....	51
9 Mapeamento de ontologia.....	52
10 Validação de ontologias.....	53
11 Aplicações da Web Semântica.....	54
12 Considerações finais.....	56
13 Referência bibliográfica.....	57

## 1 Introdução

Desde os anos noventa, as ontologias vêm sendo estudadas na área de Inteligência Artificial, representação do conhecimento e organização do conhecimento.

Recentemente, pesquisas sobre ontologias têm-se expandido para outras áreas do conhecimento como forma de integração de sistemas de informação aplicáveis em vários campos incluindo, engenharia, comércio eletrônico, educação, tecnologia, recuperação da informação entre outros, criando uma relação de interdisciplinaridade.

O motivo pelo qual as ontologias vêm ganhando popularidade é a promessa de que um domínio do conhecimento possa ser representado computacionalmente, viabilizando a comunicação entre pessoas e computadores, automaticamente, de forma inteligente.

Dentro deste contexto está inserida uma das metas da web semântica, que visa estruturar programas de buscas para melhorar o resultado das pesquisas trazendo compreensão, isto é, significado exato das palavras e as relações conceituais entre elas (BERNERS-LEE, 2001). A web semântica permite melhorar a organização através da estruturação dos dados com a utilização de linguagens que usam semântica<sup>1</sup>. A web semântica também está baseada no uso de ontologias que fornecem os vocabulários necessários para permitir a definição das relações entre os conceitos dos sistemas.

A proposta deste trabalho é trazer da literatura os principais conceitos dos componentes que envolvem as ontologias e a web semântica.

---

1 "Semântica é o estudo dos significados. Significados são idéias ou conceitos que se podem transferir da mente do falante para o ouvinte". (LYONS, 1987 p. 133)

## **2 Problema**

Atualmente, ontologias e web semântica são assuntos cada vez mais pesquisados nas mais diversas áreas do conhecimento, inclusive nas áreas de Ciência da Informação e Biblioteconomia.

Sendo assunto recente, a literatura ainda se encontra muito dispersa e em diferentes níveis de abordagem, dificultando o entendimento para alunos que estão iniciando seus estudos nestas áreas.

A escassez de documentação em português sobre conceitos e definições básicas, orientação para a construção de ontologias e sua interação com a web semântica foi a grande motivação para o desenvolvimento deste trabalho.

Coloca-se aqui, resumidamente, que uma ontologia é um conjunto de termos que podem ser usados para descrever alguma área do conhecimento ou construir sua representação. Estes termos descrevem um dado domínio e podem ser utilizados pelas máquinas para responder questões de alto nível. Neste contexto, tem-se que o papel da web semântica faz com que as informações na web sejam capturadas e entendidas por máquinas para atender aos usuários com maior precisão. As ontologias são essenciais para que o objetivo da web semântica seja alcançado.

### **3 Objetivo Geral**

Produzir um texto introdutório em língua portuguesa sobre ontologias e web semântica que forneça subsídios para a aprendizagem em nível de graduação.

#### **3.1 Objetivos Específicos**

1. Efetuar levantamento bibliográfico sobre ontologias e web semântica.
2. Identificar os conceitos básicos das áreas de ontologia e web semântica e sintetizar as abordagens dos diferentes autores.
3. Produzir um texto introdutório sobre ontologias e web semântica.
4. Disponibilizar o texto no Portal do Departamento de Ciência da Informação e Documentação para facilitar o acesso pelos alunos de graduação.

## 4 Justificativa

Ontologia e web semântica são duas áreas novas que estão sendo estudadas e ainda não dispõem de muita literatura em português. Por isso, a necessidade da produção de um texto introdutório, em língua portuguesa, com as noções básicas de conceitos, tecnologias e aplicações. Este material poderá ser utilizado como material didático e de apoio ao aprendizado para aqueles que estão iniciando estudos nesta área.

Pode-se dizer que os novos padrões que estão sendo desenhados para a web semântica trazem mudanças na arquitetura das bibliotecas digitais e dos sistemas de informação, e conseqüentemente, para os profissionais da área da Ciência da Informação e Biblioteconomia. Souza e Alvarenga (2004) apontam algumas atividades bem específicas que serão beneficiadas com esta nova Web. Projetos de novos e melhorados motores de busca; construção de novas interfaces com o usuário para sistemas de informação; construção automática de tesouros e vocabulários controlados; indexação automática de documentos; gestão do conhecimento organizacional e gestão da informação estratégica e da inteligência competitiva.

Considerando, portanto, a importância das ontologias e web semântica no contexto da Ciência da Informação, é necessário introduzir esses conceitos no ensino de graduação, no sentido de capacitar os futuros profissionais no entendimento e aplicação dessas novas tecnologias.



## 5 Metodologia

Para atingir o objetivo, foi feito um levantamento na literatura das áreas de Ciência da Informação e Ciência da Computação.

O método de coleta de dados escolhido foi o de levantamento bibliográfico com as palavras-chave *ontologia* e *web semântica*, em português e *ontology* e *semantic web*, em inglês, sem uso de qualquer tipo de filtro. As consultas foram realizadas no período de abril a julho de 2006 nas seguintes fontes de dados:

Sítios da web:

[www.scirus.com](http://www.scirus.com) – Scirus for Scientific Information

[www.w3.org](http://www.w3.org) – W3C World Wide Web Consortium

[www.semanticweb.org](http://www.semanticweb.org) – Semantic Web

[www.stanford.edu](http://www.stanford.edu) – Stanford University

[www.usp.br](http://www.usp.br) – Universidade de São Paulo

[www.cs.nyu.edu](http://www.cs.nyu.edu) – Computer Science New York University

[www.scholar.google.com](http://www.scholar.google.com) – Google Acadêmico

[www.google.com](http://www.google.com)

Base de dados:

LISA – Library and Information Science Abstracts. Acessado via IBICT.

## 6 Ontologia

Os estudos nas áreas relativas a construção de modelos conceituais<sup>2</sup>, têm sido privilegiados nos últimos anos. Os modelos conceituais permitem a construção de linguagens documentárias, sistemas computacionais, hipertextos, sistemas voltados para a construção de bases de dados de conhecimento e, mais recentemente, no âmbito da inteligência artificial, as ontologias (CAMPOS, 2004 p.22).

Campos (2004) destaca ainda que, no âmbito da Ciência da Informação, a representação do conhecimento possibilita a elaboração de linguagens documentárias, visando à recuperação da informação e à organização dos conteúdos informacionais de documentos.

Linguagem documentária é um conjunto de termos e regras para sua utilização na recuperação de informação. Linguagem artificial utilizada para o registro do assunto de documentos, dotada de vocabulário controlado e regida por uma sintaxe própria (Cavalcanti, 1978).

Assim, os sistemas de classificação, cabeçalho de assunto e tesouros são tipos de representação do conhecimento e as taxonomias e ontologias são a representação do conhecimento para acervos digitais.

### 6.1 O que é uma ontologia?

O termo “ontologia” deriva do grego “*onto*”, ser, e “*logia*”, discurso

---

<sup>2</sup> “Conhecer é modelizar, ou seja, o processo de conhecer equivale à construção de modelos de mundo/domínio a ser construído que permitem descrever e fornecer explicações sobre os fenômenos que observamos” (Teoria da Modelagem de Le Moigne (1977 apud CAMPOS, 2004 p.23)

escrito ou falado. Este termo Ontologia, com letra maiúscula, tem sido empregado ao longo da história pela Filosofia e estuda as teorias sobre a natureza da existência.

No dicionário Aurélio, ontologia está definida como a parte da filosofia que trata do ser enquanto ser, isto é, do ser concebido como tendo uma natureza comum que é inerente a todos e a cada um dos seres. Em epistemologia refere-se ao conhecimento e à sabedoria.

Na área de Inteligência Artificial, o termo ontologia foi tomado emprestado da Filosofia e para os pesquisadores da web e estudiosos em Inteligência Artificial, ontologia tem outro sentido. “É um documento que define as relações entre termos e conceitos” (BERNERS-LEE; HENDLER; LASSILA, 2001).

Uma das definições de ontologia frequentemente encontrada e estudada na literatura, é a apresentada por Gruber (1993).

Ontologia é uma especificação formal e explícita de uma conceitualização, o que existe é aquilo que pode ser representado [...] Quando o conhecimento de um domínio é apresentado num formalismo declarado, o conjunto de objetos que podem ser representados são chamados de universo do discurso. Esse conjunto de objetos, e o relacionamento descritivo entre eles, são refletidos num vocabulário representacional com o qual um programa de conhecimento de base representa o conhecimento. Mas, no contexto de Inteligência Artificial, nós podemos descrever a ontologia de um programa pela definição de um conjunto de termos representacionais. Nesse tipo de ontologia, definições associam os nomes de entidades no universo do discurso (por exemplo, classes, relações, funções ou outros objetos) com textos legíveis descrevendo o que os nomes significam, e axiomas formais que limitam a interpretação e o uso bem formado desses termos. Formalmente, uma ontologia é uma afirmação da lógica teórica (GRUBER, 1993 p. 1).

Para Borst (1997 p. 12) uma ontologia é “ Uma especificação formal e explícita de uma conceitualização compartilhada. Nessa definição, “formal” significa legível para computadores; “especificação explícita” diz respeito a conceitos, propriedades, relações, funções, restrições e axiomas explicitamente definidos e manipulados por computadores; “compartilhado” quer dizer conhecimento consensual e “conceitualização” diz respeito a um modelo abstrato de algum fenômeno do mundo real.

Uma ontologia é uma descrição formal dos conceitos e relacionamentos que existem dentro de um domínio, isso significa que uma ontologia se relaciona com um vocabulário específico e com uma linguagem específica (DAUM, 2002 *apud* ARAUJO, 2003).

A ontologia define as regras que regulam a combinação entre os termos e as relações. As relações entre os termos são criadas por especialistas e os usuários formulam consultas usando os conceitos especificados. Uma ontologia define assim uma “linguagem” (conjunto de termos) que será utilizada para formular consultas (GUARINO, 1998) .

Várias definições surgiram na literatura que enriquecem estas citadas ou sugerem outras. Para (CHANDRASEKARAM, 1999 *apud* ARAUJO, 2003) ontologia é a representação de um vocabulário, freqüentemente especializado em algum domínio ou assunto importante. Mais precisamente, não é o vocabulário que qualifica uma ontologia, mas os conceitos que os termos do vocabulário transmitem. Então, transferindo os termos de uma ontologia de uma linguagem para outra, por exemplo, do inglês para o francês, não muda o conceito ontológico.

As ontologias se apresentam como um modelo de relacionamento de entidades em um domínio particular do conhecimento. O objetivo de sua construção é a necessidade de um vocabulário compartilhado onde as

informações possam ser trocadas e também reusadas pelos usuários de uma comunidade, sejam eles humanos ou agentes inteligentes.

Destas observações pode-se concluir que a importância de uma ontologia é descrever e representar uma área do conhecimento. Determinado o domínio, sua ontologia é o sistema de representação do conhecimento deste domínio. Então, para se estruturar um sistema de representação de conhecimento eficiente, e seu vocabulário, é realizada uma análise ontológica formal do domínio.

Uma ontologia é um acordo atuando como um contrato entre parceiros, permitindo que se comuniquem com segurança dentro do contexto do domínio de informação. Por exemplo, um agente<sup>3</sup> comprometido com uma ontologia será capaz de interpretar semanticamente os itens dessa ontologia e se comunicar com outros agentes comprometidos com essa ontologia. Este compromisso, chama-se compromisso ontológico que é um acordo para usar o vocabulário compartilhado de forma coerente e consistente (GRUBER, 1993).

## 6.2 Componentes de uma ontologia

Segundo Noy e McGuinness (2001), o desenvolvimento de uma ontologia inclui a definição de classes na ontologia, estruturar as classes numa hierarquia taxonômica, definir os *slots* e descrever os valores permitidos para estes slots. As autoras apresentam os seguintes componentes de uma ontologia e citam como exemplo de uma ontologia de vinho.

1. Classes ou conceitos: descrevem conceitos do domínio. São, muitas

---

<sup>3</sup> O conceito de agentes encontra-se definido na seção 8.

vezes, o foco das ontologias. Dentro de uma classificação hierárquica as classes podem se subdividir em superclasses e subclasses, consecutivamente, conforme for a especificação. Uma subclasse herda as propriedades de sua superclasse. Por exemplo: uma classe de vinhos representa todos os vinhos. A classe dos vinhos pode se dividir em vinhos tintos, brancos e roses. Alternativamente, pode se dividir em vinhos frisantes e não frisantes.

2. *Slots* ou propriedades ou atributos ou papéis: são as várias características e atributos que descrevem cada conceito, são as propriedades das classes. Por exemplo: a classe vinho pode ter os seguintes *slots*: cor; corpo; sabor; nível de açúcar e nível de tanino.
3. Facetas<sup>4</sup> ou restrição de papéis: são as restrições das propriedades. Os *slots* podem ter diferentes facetas descrevendo os tipos de valores, valores permitidos, números de valores (cardinalidade) e outras características de valores que os *slots* podem ter. Por exemplo: no *slot* produtor da classe vinícola os valores serão os vinhos produzidos pela vinícola. O *slot* uvas de um vinho pode assumir o valor “um” ou “mais de um” se o vinho for feito por uma ou por mais de uma variedade de uva.
4. Instâncias: representam os elementos de uma ontologia, ou seja, são as ocorrências dos conceitos e relações que foram estabelecidas pela ontologia. Uma instância é um conceito que pertence a uma classe e que possui determinados valores de *slots*. Por exemplo: a instância individual Chateau-Morgon-Beaujolais representa um tipo específico de vinho da classe Beaujolais. Esta instância tem os seguintes valores de *slots* definidos:

- Corpo: leve

---

<sup>4</sup> Facetas para a Ciência da Computação tem o significado de restrição/limitação enquanto que para a Ciência da Informação faceta é a fragmentação de um assunto em diversas partes constituintes.

- Cor: vermelho
- Sabor: delicado
- Nível de tanino: baixo
- Uva: gamay (instância da classe de uva-do-vinho)
- Região: Beaujolais (instância da classe de região-do-vinho)

### **6.3 Tipos de ontologias**

As considerações a seguir são sugeridas por Guarino (1998 p. 9) para o desenvolvimento de diferentes tipos de ontologias, de acordo com o seu nível de generalidade.

1. Ontologias de alto nível: descrevem conceitos de forma bem geral como espaço, tempo, material, objeto, evento, ação, etc., os quais são independentes de um problema ou domínio particular.
2. Ontologias de domínio e ontologias de tarefas: descrevem, respectivamente, o vocabulário relacionado a um domínio genérico (como medicina ou automóveis) ou uma tarefa genérica (como diagnóstico ou vendas).
3. Ontologias de aplicação: descrevem conceitos dependendo do domínio e de tarefas particulares. Estes conceitos, freqüentemente, correspondem a papéis desempenhados por entidades do domínio, quando na realização de certas tarefas. Por exemplo, ajuda para o diagnóstico de doenças mentais.

A figura 1 mostra a relação entre estas ontologias. Os conceitos de uma ontologia de domínio ou de tarefa devem ser especializações dos

termos introduzidos por uma ontologia genérica (alto nível). Enquanto que os conceitos de uma ontologia de aplicação devem ser especializações dos termos das ontologias de domínio ou de tarefa.

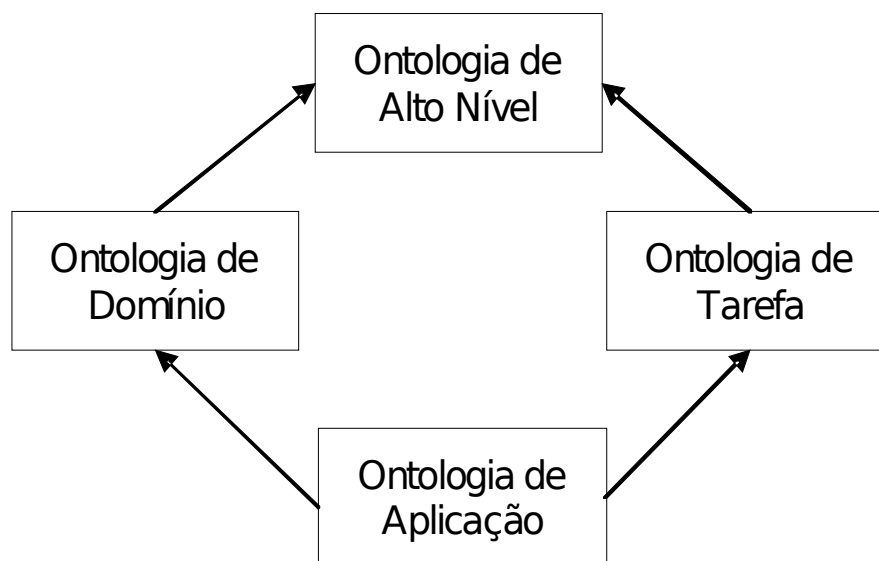


Fig. 1 Tipos de ontologias.

Fonte: Guarino, 1998

As ontologias não apresentam sempre a mesma estrutura, dependem sempre da proposta de cada uma. Mas existem algumas características e componentes básicos comuns que são encontrados em muitas ontologias, fazendo com que possuam semelhanças entre suas funções. Almeida e Bax (2003) diferenciam alguns tipos de ontologias quanto às funções, ao grau de formalismo de seu vocabulário, à aplicação, à estrutura e ao conteúdo da conceitualização. A relação a seguir coloca resumidamente essas diferenças.



### **6.3.1 Quanto à função (MIZOGUCHI, VANWELKENBUYSEN e IKEDA,1995)**

1. Ontologias de domínio: reutilizáveis no domínio, fornecem vocabulário sobre conceitos, seus relacionamentos, sobre atividades e regras que os governam.
2. Ontologias de tarefas: fornecem um vocabulário sistemático de termos, especificando tarefas que podem ou não estar no mesmo domínio.
3. Ontologias gerais: incluem um vocabulário relacionado a coisas, eventos, tempo, espaço casualidade, comportamento, funções, etc.

### **6.3.2 Quanto ao grau de formalismo (USCHOLD e GRUNINGER, 1996)**

1. Ontologias altamente informais: expressas livremente em linguagem natural.
2. Ontologias semi-informais: expressas em linguagem natural de forma restrita e estruturada.
3. Ontologias semi formais: expressas em uma linguagem artificial definida formalmente.
4. Ontologia rigorosamente formal: expressas por meio de termos definidos com semântica, teoremas e provas.

### **6.3.3 Quanto à aplicação (JASPER e USCHOLD, 1999)**

1. Ontologia de autoria neutra: um aplicativo é escrito em uma única língua e depois convertido para uso em diversos sistemas, reutilizando-se as informações.
2. Ontologias como especificação: cria-se uma ontologia para um domínio, a qual é usada para documentação e manutenção no desenvolvimento de *software*.
3. Ontologia de acesso comum à informação: quando o vocabulário é inacessível, a ontologia torna a informação inteligível, proporcionando conhecimento compartilhado dos termos.

### **6.3.4 Quanto ao conteúdo (VAN-HEIJST, SCHREIBER e WIELINGA, 2002)**

1. Ontologias terminológicas: especificam termos que serão usados para representar o conhecimento em um domínio (por exemplo os léxicos).
2. Ontologias de informação: especificam a estrutura de registros de bancos de dados (por exemplo os esquemas de bancos de dados).
3. Ontologias de modelagem do conhecimento: especificam conceitualizações do conhecimento, têm uma estrutura interna semanticamente rica e são refinadas para o uso do domínio do conhecimento que descrevem.
4. Ontologias de aplicação: contêm as definições necessárias para modelar o conhecimento em uma aplicação.

5. Ontologias de domínio: expressam conceitualizações que são específicas para um determinado domínio do conhecimento.
6. Ontologias genéricas: similares às ontologias de domínio, mas os conceitos que as definem são considerados genéricos e comuns a vários campos.
7. Ontologias de representação: explicam as conceitualizações que estão por trás dos formalismos de representação do conhecimento.

As tipologias apresentadas acima, por não ser objeto de estudo deste trabalho, foram colocadas apenas para apresentar outros tipos de ontologias e as diversidades dos autores.

#### **6.4 Critérios para a construção de ontologias**

Independente do domínio, a construção de uma ontologia requer um trabalho complexo e dispendioso.

Grüber (1993) propõe um conjunto de critérios que devem ser observados no desenvolvimento de uma ontologia, visando com isso atingir maior eficiência na proposta de construir um conhecimento compartilhado e a interoperabilidade entre os programas.

1. Clareza: uma ontologia deve, efetivamente, comunicar o significado pretendido dos termos definidos. Suas definições devem ser objetivas e independentes do contexto social ou computacional. Formalismo é um meio para esse fim. Uma definição deve ser declarada, sempre que possível, em axiomas lógicos. Onde for possível uma definição completa é preferida em relação a uma definição parcial e todas as

definições devem ser documentadas com linguagem natural, para reforçar a clareza.

2. Coerência: uma ontologia deve ser coerente, isto é, as inferências devem ser consistentes com as definições. Pelo menos as definições axiomáticas devem ser logicamente consistentes. Coerência também deve ser aplicada para os conceitos que são definidos informalmente, como aqueles descritos em documentos de linguagem natural e exemplos. Se uma sentença passível de ser inferida, a partir dos axiomas da ontologia, contradiz uma definição ou exemplo dado de maneira informal, então a ontologia é incoerente.
3. Extensibilidade: uma ontologia deve ser projetada para antecipar usos do vocabulário compartilhado, ou seja, uma ontologia deve oferecer um conceito fundamentado por uma gama de tarefas antecipadas e sua representação deve ser hábil para que possa ser estendida e especializada. Em outras palavras, uma ontologia deve ser capaz de definir novos termos para usos especiais baseado em um vocabulário já existente, de modo que não requeira a revisão de definições existentes.
4. Compromisso mínimo com implementação: a conceitualização deve ser especificada no nível do conhecimento sem depender de uma codificação particular no nível simbólico ou de codificação. Uma tendência de codificação resulta quando escolhas de representação são feitas puramente para a conveniência de notação ou implementação. Assim, essa tendência de codificação deve ser minimizada porque os agentes que compartilham conhecimento podem ser implementados em diferentes sistemas e estilos de representações.
5. Compromisso ontológico mínimo: uma ontologia requer o mínimo

compromisso ontológico, suficiente para atender à intensão da atividade compartilhada do conhecimento. Uma ontologia deve fazer poucas imposições sobre o mundo que está sendo modelado, permitindo que as partes comprometidas com a ontologia fiquem livres para especializar e instanciar a ontologia, sempre que necessário.

### **6.5 Porque desenvolver uma ontologia?**

Uma ontologia define um vocabulário comum para pesquisadores que necessitam compartilhar informações em um domínio. Isso inclui interpretações de definições de conceitos básicos do domínio e das relações entre eles, feitas por máquinas e/ou humanos. Neste contexto, Noy e McGuinness (2001) citam algumas razões para se desenvolver uma ontologia.

1. Compartilhar entendimento comum da estrutura de informação entre pessoas e agentes (*software*). Por exemplo: diferentes páginas da web contêm informações sobre Medicina. Se essas páginas compartilham e publicam a mesma ontologia de termos que usam, então os agentes de *software* podem extrair e agregar informações dessas páginas. Os agentes podem usar essas informações agregadas para responder as perguntas dos usuários ou como incluir dados para outras aplicações.
2. Facilitar o reuso de conhecimento de domínio. Por exemplo: modelos para alguns domínios diferentes precisam representar a noção de tempo. Essa representação inclui noções de intervalos de tempo, pontos no tempo, medidas relativas de tempo e outras. Se um grupo de pesquisadores desenvolvem essa ontologia em detalhes, outros

podem simplesmente reusá-la para seus domínios. E mais, se for preciso construir uma ontologia maior, é possível integrar outras ontologias já existentes descrevendo partes do domínio maior. E também pode ser reusada uma ontologia geral, assim como a ontologia UNSPSC<sup>5</sup>, e extendê-la para descrever o domínio de interesse.

3. Fazer especificações explícitas do domínio de conhecimento. Detalhar uma implementação torna possível trocar as especificações facilmente se o conhecimento sobre o domínio mudar. Uma codificação rígida sobre o mundo em um código de linguagem de programação faz com que estas especificações sejam não somente difíceis de entender, mas também difíceis de mudar, particularmente para as pessoas que não são familiarizadas com programação. E mais, especificações explícitas do conhecimento são úteis para novos usuários que precisam aprender qual o significado dos termos do domínio.
4. Analisar o conhecimento de domínio. Uma vez que os termos estejam disponíveis, fazer uma análise formal destes termos é extremamente válida quando se está reusando ontologias já existentes e quando está se extendendo uma ontologia. De forma que termos e/ou definições não se repitam ou tenham divergências.

Desenvolver uma ontologia é como definir um conjunto de dados e sua estrutura para ser usada por outros programas. Métodos de solução de problemas, aplicação de domínios independentes e agentes de *software* usam ontologias e bases de conhecimentos construídos a partir de ontologias como dados. Por exemplo, uma ontologia de vinhos, comidas e a combinação de pratos com vinhos. Esta ontologia pode ser usada como

---

5 United Nations Development Program and Dun & Bradstreet desenvolveu a ontologia UNSPSC que fornece terminologias para produtos e serviços ([www.unspsc.org](http://www.unspsc.org))

base por algumas aplicações nas ferramentas de gerenciamento de um restaurante. Uma aplicação pode criar sugestões de vinhos para o cardápio do dia ou responder questões dos garçons e dos clientes.

## **6.6 Metodologia para a construção de ontologias**

A literatura nos aponta que um dos problemas na construção das ontologias é a inexistência de um padrão definido de metodologia. A estrutura de construção de ontologias não é generalizada, devendo-se sempre considerar os propósitos e objetivos para a construção de cada ontologia. Mas, apesar disso, encontram-se ontologias desenvolvidas por vários grupos com diferentes abordagens e metodologias. Neste trabalho serão apresentadas duas metodologias de construção de ontologias.

A primeira é descrita por Uschold e King (1995) e inclui os seguintes estágios:

1. Identificação da proposta: é importante esclarecer porque a ontologia esta sendo construída, quais as intenções de uso e suas características.
2. Construindo a ontologia: é a construção da ontologia propriamente dita. Os autores sugerem que nas etapas de captura e codificação sejam observados os critérios para a construção de ontologias, propostos por Gruber (1993) e que sejam investigadas ontologias já existentes.
  - Captura: identificação dos conceitos e dos relacionamentos no domínio de interesse. Produção de textos precisos eliminando as ambigüidades e definição de cada conceito e

dos relacionamentos. Identificação de termos que se referem a cada conceito e relacionamentos.

- Codificação: é a representação da conceitualização capturada no passo anterior por meio de uma linguagem formal, ou seja, a formalização.
3. Avaliação: uma ontologia deve ser avaliada sob requisitos específicos, questões de competência e/ou o mundo real.
  4. Documentação: todas as decisões tomadas devem ser documentadas, tanto para os principais conceitos definidos na ontologia, bem como para as primitivas usadas para expressar as definições na ontologia.

A figura 2 mostra o processo de desenvolvimento de uma ontologia e suas interdependências. As linhas tracejadas indicam que há uma interação constantes, entre os estágios associados. As linhas cheias mostram o fluxo principal de trabalho no processo de construção de uma ontologia. A linha que envolve os estágios de captura e formalização da ontologia realça a forte interação entre estes estágios.



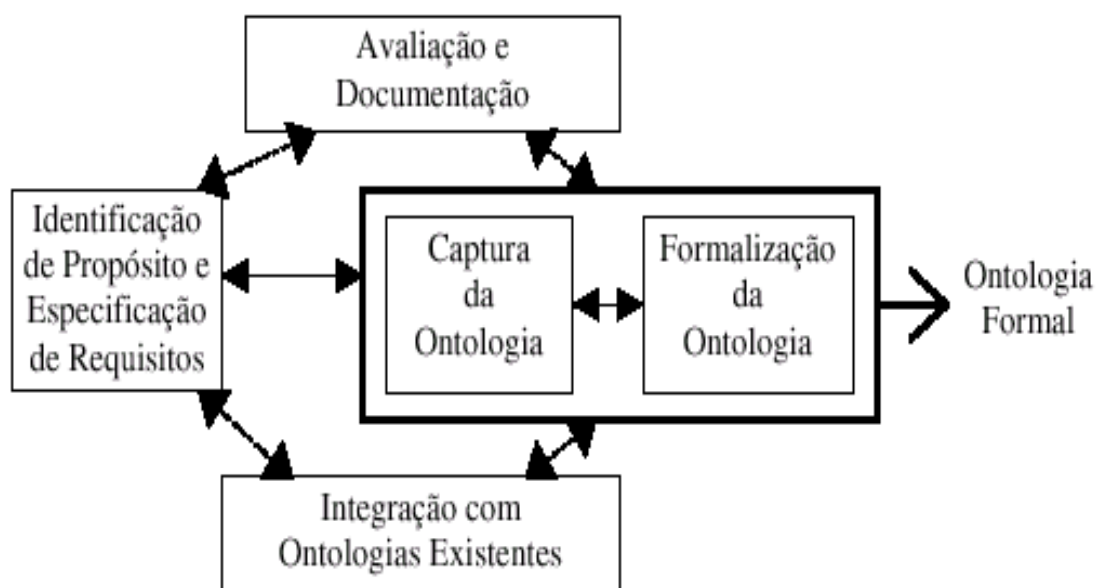


Fig. 2 Construção de ontologias.

Fonte: Falbo, 1998

A segunda metodologia é apresentada por Noy e McGuinness (2001 p.5) que sugerem um processo para modelar uma ontologia, usando Protégé 2000, Ontolingua (1997) e Chimaera (2000). Este processo é apresentado resumidamente a seguir:

1. Determinar o domínio e o escopo da ontologia: para definir o domínio e o escopo deve-se começar respondendo algumas perguntas:
  - Qual o domínio que a ontologia vai cobrir?
  - Para que será usada a ontologia?
  - As informações contidas na ontologia deverão responder quais tipos de questões?
  - Quem irá usar e manter a ontologia?

As respostas destas questões podem mudar durante o processo de desenvolvimento da ontologia, mas a qualquer momento elas ajudarão a

limitar o escopo do modelo.

2. Considerar o reuso de ontologias existentes: muitas ontologias já estão disponíveis em forma eletrônica e podem ser importadas<sup>6</sup> num ambiente de desenvolvimento de ontologias que esteja sendo usado.
3. Enumerar os termos importantes da ontologia: fazer a lista de todos os termos e enunciados sobre eles e esclarecer pontos importantes com os especialistas do domínio, para saber quais termos são importantes e quais propriedades possuem.
4. Definir as classes e as hierarquias: existem várias possibilidades para se desenvolver uma hierarquia. A que será apresentada é a descrita por Uschold e Gruninger (1996).
  - De cima para baixo (*top-down*) começa com as definições mais gerais, passando-se posteriormente para as mais específicas.
  - De baixo para cima (*bottom-up*) começa com definições de classes mais específicas e vai-se agrupando essas em classes mais gerais.
  - Combinação: é uma combinação dos dois métodos acima. A partir de um conceito destacado identificam-se os mais gerais e os mais específicos.
5. Definir as propriedades das classes (*slots*): uma vez definidas algumas classes, deve-se descrever a estrutura interna dos conceitos. Determinar que classe cada propriedade descreve. Essas

---

<sup>6</sup> Existem bibliotecas de ontologias reusáveis na web. Por ex.: Ontololingua - [www.ksl.stanford.edu/software/ontolingua/](http://www.ksl.stanford.edu/software/ontolingua/); DAML - [www.daml.org/ontologies/](http://www.daml.org/ontologies/).

propriedades serão os *slots* e ficarão vinculadas às classes.

6. Definir as facetas dos *slots*: *slots* podem ter facetas diferentes descrevendo os tipos de valores e a cardinalidade. A cardinalidade define os valores de um *slot*.
7. Criar instâncias: o último passo é criar as instâncias individuais das classes. Definir uma instância individual de uma classe requer: a escolha da classe; a criação de uma instância individual dessa classe e o preenchimento dos valores de *slots*.

## **6.7 Ferramentas para a construção de ontologias**

Algumas ferramentas têm sido utilizadas para auxiliar o desenvolvimento e a construção de ontologias. “Critérios devem ser definidos para que as ferramentas de construção de ontologias possam ser comparáveis” (ALMEIDA e BAX, 2003 p. 15). E como afirma CAMPOS (2004, p. 31), um modelo conceitual deve ser visto, também, como um espaço comunicacional em que transpomos o mundo fenomenal para um espaço de representação. E a ontologia utiliza-se da representação gráfica, como ferramenta para garantir um projeto lógico mais bem estruturado de um sistema. Em geral, as ferramentas utilizam linguagens de representação para a construção das ontologias. A seguir algumas ferramentas para a construção, uso e edição de ontologias e suas descrições, baseado num trabalho de Almeida e Bax (2003).

1. Ontolingua: conjunto de serviços que possibilitam a construção de ontologias compartilhadas entre grupos. Permite acesso a uma biblioteca de ontologias, tradutores para linguagens e um editor para criar e navegar pela ontologia (FARQUHAR, FIKES e RICE, 1996).

2. JOE (*Java Ontology Editor*): 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).
3. OntoEdit: 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 et alii, 2000).
4. Protégé 3.2: é um 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 e MUSEN, 2000).

## **6.8 Linguagens para a construção de ontologias**

Como nas outras etapas da construção de uma ontologia, critérios devem ser definidos para que se possa comparar as linguagens para a construção de ontologias. As linguagens usadas para a construção de ontologias são, na verdade, códigos para serem lidos em páginas da *web* e transformados para serem entendidos por agentes eletrônicos que pesquisam informações. Existem diversos padrões e linguagens para a

construção e compartilhamento de ontologias na web, todos baseados em XML<sup>7</sup> (eXtensible Markup Language). Algumas das linguagens encontradas na literatura serão apresentadas a seguir, resumidamente, baseadas no trabalho de Almeida e Bax (2003).

1. CycL: linguagem formal que expressa conhecimento por meio de um vocabulário de termos (constantes semânticas, variáveis, números, sequências de caracteres, etc.) os quais são combinados em expressões, sentenças e finalmente bases de conhecimento (LENAT e GUHA, 1990).
2. Flogic (*Frame Logic*): integra *frames* e lógica de primeira ordem. Trata de uma forma declarativa os aspectos estruturais das linguagens baseadas em *frames* e orientadas a objeto (identificação de objetos, herança, tipos polimórficos, métodos de consulta, encapsulamento, etc.). Permite a representação de conceitos, taxonomias, relações binárias, funções, instâncias, axiomas e regras (KIFER, LAUSEN e WU, 1990).
3. LOOM: descendente da família KL-ONE (*Knowledge Language One*), é baseada em lógica descritiva e regras de produção. Permite a representação de conceitos, taxonomias, relações n-árias, funções, axiomas e regras de produção (BRILL, 1993).
4. Ontolingua: combina paradigmas das linguagens baseadas em *frames* e lógica de primeira ordem. Permite a recuperação de conceitos, taxonomias de conceitos, relações n-árias, axiomas, instâncias e procedimentos. Sua alta expressividade causa problemas na construção de mecanismos de inferência (CHAUDHRI et alii, 1996).
5. RDF (*Resource Description Framework*) / RDFS (*RDFSchema*):

---

<sup>7</sup> O conceito de XML encontra-se definido na seção 7.2.2.

desenvolvidos pelo W3C, têm por objetivo a representação de conhecimento por meio da idéia de redes semânticas. São linguagens que permitem a representação de conceitos, taxonomias de conceitos e relações binárias (LASSILA e SWICK, 1999).

6. SHOE (*Simple HTML Ontology Extensions*): utiliza extensões ao HTML, adicionando marcações para inserir metadados em páginas Web. As marcações podem ser utilizadas para a construção de ontologias e para anotações em documentos da web (HEFLIN e HENDLER, 2000).
7. OIL (*Ontology Inference layer*): antecede o DAML+OIL, é base para uma linguagem para a web semântica. Combina primitivas de modelagem das linguagens baseadas em *frames* com a semântica formal e serviços de inferência da lógica descritiva. Pode verificar classificação e taxonomias de conceitos (FENSEL et alii, 2000).
8. DAML+OIL (*DARPA Agent Markup Language + Ontology Inference Layer*): é uma linguagem de marcação semântica para a web que apresenta extensões a linguagem como o DAML, RDF e RDFS, por meio de primitivas de modelagem baseadas em linguagem lógica (HORROCKS et alii, 2001).
9. OWL (*Web Ontology Language*): é uma linguagem para definir e instanciar ontologias na web (w3c.org, 2004). É uma linguagem que permite que semântica seja explicitamente associada ao conteúdo dos dados na web e formalmente especificada através de ontologias, compartilhadas na Internet. Atualmente o W3C utiliza a OWL como padrão para construção de ontologias (SMITH et alii., 2002).

A escolha de uma dessas linguagens deve ser feita de acordo com os propósitos de representação da ontologia.

## 6.9 Vantagens no uso de ontologias

As ontologias constituem uma ferramenta poderosa para suportar a especificação e a implementação de sistemas computacionais de qualquer complexidade. Viera e Chaves (2001) citados por Falbo (1998) apresentam três principais benefícios no emprego de ontologias.

1. Comunicação: são ferramentas úteis para ajudar as pessoas a se comunicarem, sob várias formas, acerca de um determinado conhecimento. Elas podem ajudar as pessoas a raciocinar e a entender o domínio do conhecimento e, portanto, atuam como uma referência para a obtenção do consenso numa comunidade profissional sobre o vocabulário técnico a ser usado nas suas interações.
2. Formalização: pela sua natureza formal, a especificação do domínio elimina contradições e inconsistências envolvendo as restrições, resultando, em uma especificação não ambígua. Mais ainda, como é usada uma notação formal, a especificação formalizada pode ser automaticamente verificada e validada, por um provador automático de teoremas. Com um mecanismo de inferência, é também possível derivar novos conhecimentos de forma automática, a partir da base de conhecimento já presente na ontologia.
3. Representação do conhecimento e reuso: a ontologia forma um vocabulário de consenso e representa o conhecimento de domínio de forma explícita no seu alto nível de abstração, possuindo um potencial enorme de reuso. O conhecimento formalizado na camada de domínio pode ser especializado em diferentes aplicações, servindo diferentes propósitos, por diferentes equipes de desenvolvimento, em diferentes espaços do tempo.

## 6.10 Problemas no uso de ontologias

Alguns problemas são apontados no uso de ontologias identificados por O'Leary (1997) e citados por Falbo (1998).

1. A escolha de uma ontologia é um processo político, já que nenhuma ontologia pode ser totalmente adequada a todos os indivíduos ou grupos.
2. As ontologias necessitam evoluir, não são estáticas.
3. Estender ontologias não é um processo direto. Ontologias são, geralmente, estruturadas de maneira precisa e, como resultado, são particularmente vulneráveis a questões de extensão, dado o forte relacionamento entre complexidade e precisão das definições.
4. A noção de bibliotecas de ontologias sugere uma relativa independência entre diferentes ontologias. A interface entre elas constitui, portanto, um impedimento, especialmente porque cada uma delas é desenvolvida no contexto de um processo político.

A principal aplicação e benefício das ontologias consiste em prover semântica à Internet. A visão da Internet semântica é uma idealização de Tim Berners-Lee. A primeira geração da rede, a Internet, permitia apenas a troca de dados entre máquinas distintas. A segunda geração, a *World Wide Web* (rede de alcance mundial), ou simplesmente *www* ou *web*, provocou uma revolução por disponibilizar uma vasta gama de aplicativos e informações para as pessoas. Porém, as páginas só possuem informação léxica.

A próxima geração da rede será a web semântica, em que sua maior



motivação é transformar os dados e aplicativos em elementos legíveis e compreensíveis por *software* ou agentes inteligentes, facilitando a comunicação. Neste contexto, as ontologias desempenham um papel fundamental para a web semântica que será apresentada a seguir.

## 7 Web Semântica

Nesta seção será apresentada a web semântica, um pouco da sua história, definições de termos, as camadas de esquema da sua arquitetura e qual a sua função e, por fim, onde a web semântica está sendo aplicada. A fonte de recurso bibliográfico foi baseada, principalmente, na documentação encontrada nas páginas do sítio do *World Wide Web Consortium*<sup>8</sup>, [www.w3.org](http://www.w3.org).

### 7.1 Histórico

Com o fenômeno atual da explosão de documentos em formato digital, vimos surgir alguns problemas principalmente na organização e na recuperação de informações. Para que se tenha uma representação apropriada de documentos, faz-se necessário criar sistemas de indexação eficientes, capazes de recuperar as informações neles contidas satisfazendo as necessidades dos usuários, o mais precisamente possível.

Embora a web tivesse sido planejada para facilitar o acesso, o intercâmbio e a recuperação de informações, não é isso o que podemos observar hoje. O que vemos é um crescimento exponencial de forma desordenada gerando uma situação caótica com um imenso repositório de documentos que deixa muito a desejar quando precisamos recuperar aquilo que temos necessidade (SOUZA e ALVARENGA, 2004 p.133).

Neste contexto surge a proposta da web semântica. Projetada pelo

---

<sup>8</sup> *World Wide Web Consortium*, em uma tradução literal, “Consórcio da Teia de Alcance Mundial”. É um consórcio de empresas, profissionais, cientistas e instituições acadêmicas que é responsável pela criação de padrões tecnológicos recomendados para serem usados na *World Wide Web*.

*World Wide Web Consortium*, ou simplesmente W3C, tendo como diretor e um dos fundadores Tim Berners-Lee . Berners-Lee é considerado o “pai da Web” por ser o criador da *World Wide Web* (www) em 1989. Nas palavras de Berners-Lee (2001) a “web semântica não é uma web separada, mas uma extensão da atual. Nela a informação é dada com um significado bem definido, permitindo melhor interação entre os computadores e pessoas”. Dessa forma, programas especializados e personalizados, chamados de agentes podem interagir por meio de uma infra-estrutura de dados, trocando informações entre si, automatizando tarefas de rotina dos usuários.

A web semântica é a criação e implantação de padrões tecnológicos que permitem que dados sejam compartilhados e reusados. É um esforço colaborativo liderado pela W3C com participação de um grande número de pesquisadores e empresas. Está baseado na RDF (*Resource Description Framework*), que integra uma variedade de aplicações usando XML (*eXtensible Markup Language*) para sintaxe e URIs (*Uniform Resource Identifier*) para identificação. Cada um destes itens serão definidos no decorrer do trabalho.

Portanto, o objetivo da web semântica é transformar o conteúdo atual da web num formato no qual não só humanos, mas também computadores entendam o significado das informações, de forma a permitir que agentes inteligentes recuperem e manipulem informações pertinentes.

## **7.2 Arquitetura da Web Semântica**

A web tradicional (www) é baseada em documentos escritos em *HyperText Markup Language* (HTML), uma linguagem usada, principalmente, para a criação das páginas que estão na rede. A HTML não

especifica nenhuma semântica quanto ao conteúdo da página, preocupando-se apenas com a apresentação dos documentos, portanto não detém um formalismo adequado para a representação de conceitos manipuláveis por máquinas.

Para que a web semântica pudesse ser realizada foram estabelecidos alguns padrões e a utilização de novas tecnologias. Algumas já são usadas, enquanto outras ainda precisam ser desenvolvidas como: determinação de linguagens formais para a representação do conhecimento no contexto da web semântica, disponibilização de informações em linguagens baseadas em ontologias, criação de serviços de inferência e criação de políticas e privacidade na web semântica. Berners-Lee (2001) comenta que quando programas que coletam conteúdos na Web, de fontes diversas, processarem as informações e trocarem resultados com outros programas, o poder real da web semântica será atingido. Ressaltando assim, o importante papel dos agentes nas aplicações da web semântica.

A arquitetura da web semântica foi dividida em camadas de linguagens, segundo a visão do W3C (2001), como mostra a figura 3.

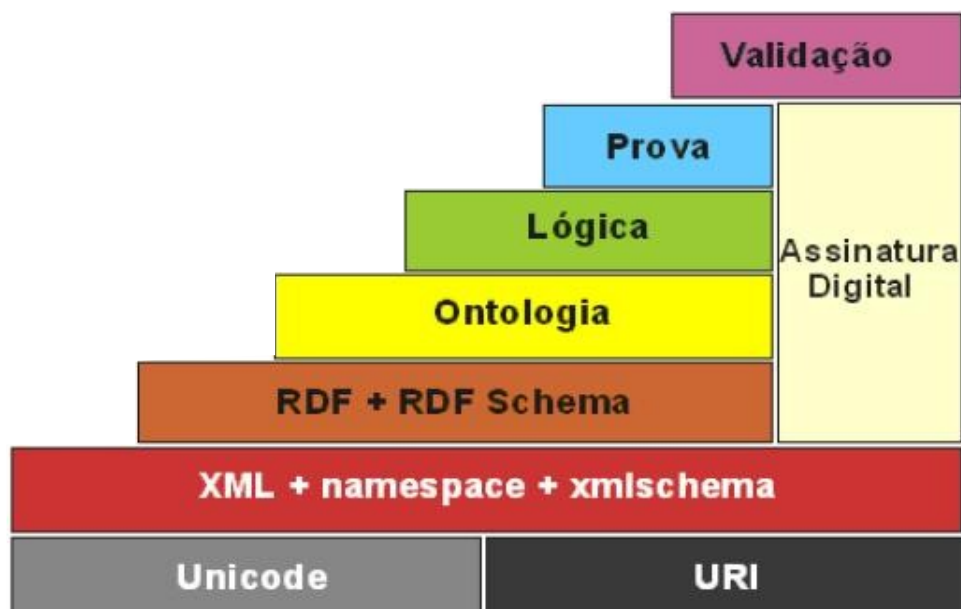


Fig. 3 Arquitetura da web semântica.

Fonte: W3C, 2001

Esta arquitetura define as tecnologias necessárias para que os conteúdos das páginas possam ser compreendidos por computadores.

### 7.2.1 A camada Unicode e URI

A camada Unicode e URI estabelecem um conjunto de caracteres e fornecem meios para a identificação de objetos na web semântica.

Unicode permite que textos e imagens possam ser lidos pelos computadores em qualquer lugar.

O URI (*Uniform Resource Identifier*) fornece um endereço global único dos recursos disponíveis na Web, seja este uma página de texto, vídeo, imagem, som, uma aplicação, etc. O tipo mais comum de URI é a URL

(*Uniform Resource Locator*), que descreve o endereço de uma página na web utilizando protocolos específicos da Internet como HTTP (*HyperText Transfer Protocol*) e FTP (*File Transfer Protocol*), por exemplo: <http://www.google.com.br>

### **7.2.2 A camada XML, Namespace (NS), XML Schema**

*EXtensible Markup Language* (XML) é uma linguagem de marcação muito simples e flexível. Flexível no sentido de que podem ser acrescentadas novas *tags* (marcações) à medida da necessidade. Em alguns aspectos, o XML assemelha-se ao HTML (*HyperText Markup Language*), ambas são derivadas do SGML (*Standard Generalized Markup Language*) e contêm *tags* para descrever o conteúdo de um documento. O XML está desempenhando um papel importante e crescente de uma grande variedade de dados na web. O XML fornece uma sintaxe básica para estruturar documentos, sem impor nenhuma restrição semântica a eles.

O W3C define um XML *NameSpace* (NS) como sendo uma coleção de nomes, identificados pela referência URI, que é usada em documentos XML como tipos de elementos e nomes de atributos. É um padrão para prover um nome único de elemento e de atributo em uma instância XML. Uma instância XML pode conter nomes de elementos e de atributos de mais de um vocabulário XML. Se em cada vocabulário é dado um *namespace* então a ambiguidade entre nomes idênticos de elementos ou atributos pode ser solucionada, onde usuários e programas de validação de código XML podem consultar a sintaxe e propriedades semânticas dos conceitos cobertos. Cada nome de elemento em um *namespace* deve ser único.

O XML *Schema* é uma linguagem de definição que oferece facilidades para descrever a estrutura de conteúdos de documentos XML. Através

dessa linguagem, define-se o formato válido de documentos XML, incluindo quais os elementos e atributos são permitidos ou não, quais são as suas localizações, o número de ocorrências de cada elemento e outras características. Ou seja, proporciona mecanismos para a definição de gramáticas para correção de documentos XML.

### **7.2.3 A camada RDF e RDF Schema**

A *Resource Description Framework* (RDF) é uma linguagem, desenvolvida pelo W3C, para representar informação na Web. RDF é uma estrutura para processar metadados (dados sobre dados) e seu principal objetivo é o de facilitar o intercâmbio de informações, que podem ser entendidas por máquinas, entre aplicativos via web. Basicamente, RDF define um modelo de dados para descrever estes dados semanticamente, de maneira que possam ser “entendidos” por computadores.

O RDF estabelece um padrão de metadados para ser embutido na codificação XML e sua implementação é exemplificada pelo *RDF Schema* (RDFS), que faz parte da especificação do padrão. A descrição dos dados e metadados é feita pelo esquema de “triplas” recurso-propriedade-enunciado são, respectivamente, sujeito-predicado-objeto.

1. Recurso (*resources*): um recurso pode ser uma página da web, uma parte de uma página da web, um grupo de páginas ou um objeto que não é acessível diretamente via web, como um livro impresso.
2. Propriedades (*properties*): Uma propriedade é um aspecto específico, uma característica, um atributo ou uma relação usada para descrever um recurso.

3. Enunciados (*statements*): cada tipo de propriedade possui valores e expressa o relacionamento com algum outro recurso. Esses valores podem ser primitivos como textos e números ou podem referenciar outros recursos. Nas palavras de Araújo (2003) o objeto é o valor da propriedade, que pode ser outro recurso, uma cadeia de caracteres, um endereço da web ou outro tipo de dado definido pela XML.

O padrão RDF, as ontologias e os *namespaces* compartilhados vão permitir que qualquer indivíduo ou organização publique informações em páginas web de forma que produtos de software ou agentes possam interpretar a informação de forma inteligente (SOUZA; ALVARENGA, 2004 p.136).

RDF *Schema* é uma linguagem que define a estrutura válida para os documentos RDF. É semelhante à linguagem XML *Schema* e junto com o RDF são recomendações do W3C que definem o padrão para a representação de metadados e são a base das linguagens que expressam semântica na web semântica.

#### **7.2.4 A camada de ontologia**

A camada de ontologia fornece um vocabulário compartilhado, que pode ser usado para modelar determinado domínio (área do conhecimento) seus conceitos, propriedades e relações. Isto faz com que máquinas “pensem” a respeito dos significados dos dados e possam inferir novos fatos. Como foi visto na seção 6 deste trabalho, onde foi dado um enfoque maior sobre as ontologias.

As ontologias usam linguagens que representam a semântica das informações na web e que possibilitam a troca de dados entre ambientes diferentes. Atualmente, o consórcio W3C recomenda o uso da *Web*



*Ontology Language* ou OWL para a representação de ontologias e metadados.

A OWL é uma linguagem de marcação semântica e é uma revisão da DAML+OIL linguagem de ontologias na web, também recomendada pelo W3C. É utilizada quando a informação contida num documento precisa ser processada por uma aplicação. Pode ser usada para representar o significado dos termos num vocabulário e as relações entre os termos. OWL tem mais facilidade para expressar significado e semântica do que XML, RDF e RDFS e mais, OWL vai além dessas linguagens na habilidade de representar conteúdos interpretados por máquinas na web. OWL adiciona mais vocabulário para descrever propriedades e classes: relações entre classes (ex.: disjoint), cardinalidade (ex.: “exatamente um”), igualdade, mais riqueza na descrição de propriedade, características de propriedades (ex.: simetria) e classes enumeradas. A OWL prove três sublinguagens:

1. OWL *Lite* – permite definir uma ontologia de classes e propriedades e as instâncias dessas classes e propriedades, por exemplo, a cardinalidade máxima ou mínima assume apenas os valores 0 ou 1. OWL *Lite* provê uma rápida migração de tesouros e taxonomias.
2. OWL DL - ( *Description Logic*<sup>9</sup> ou Lógica de Descrição) estende a OWL *Lite* permitindo restrições de cardinalidade que não são limitadas em 0 ou 1. Também define classes baseadas em valores específicos de propriedades usando constructos como *hasValue*, *class expression* usando combinações booleanas como *unionOf*, *intersectionOf* e *complementOf*.
3. OWL *Full* – estende OWL DL, é a OWL completa. Permitindo que

---

<sup>9</sup> *Description Logic* é uma linguagem de representação de conhecimento que pode ser usada para representar informações estruturadas e bem definidas. Com ela podemos descrever conceitos em um domínio e usar semântica baseada em lógica de predicados de primeira ordem. (W3C.org, 2003)

classes sejam tratadas simultaneamente como coleções e instâncias. OWL *Full* permite uma ontologia aumentar o significado de vocabulários pré-definidos em RDF ou OWL.

Cada uma destas sublinguagens é uma extensão da sua precedente. As linguagens menos expressivas (OWL *Lite* e DL) estão contidas nas mais expressivas (OWL DL e *Full*), de maneira que uma ontologia definida numa linguagem menos expressiva é aceita por uma linguagem mais expressiva: a recíproca não é verdadeira.

### **7.2.5 A camada lógica**

A camada de lógica tem como objetivo especificar linguagens de lógica de primeira ordem, também chamadas de regras, para facilitar a construção de inferências, as quais os agentes poderão utilizar para relacionar e processar informações, executando serviços inteligentes.

### **7.2.6 A camada de prova**

A camada de prova executa as regras de inferência da camada de lógica onde os agentes têm mais poder para raciocinar sobre conceitos e relacioná-los na camada de ontologia e provar que as informações trocadas são verdadeiras. É nesta camada que verifica-se a consistência dos dados acessados na web semântica.

### **7.2.7 A camada de confiança e assinatura digital**

A camada de confiança avalia, através de assinatura digital, se a prova está correta ou não. Pode-se criar um agente que trabalhe apenas com dados confiáveis na web semântica. A assinatura digital visa garantir a procedência de um documento, o que é fundamental para decidir se a informação é confiável ou não. Trata-se de uma característica importante, tanto para a web semântica quanto para o ambiente da Internet, na qual blocos de dados criptografados são utilizados para garantir a autenticidade das fontes e a confiabilidade das informações consultadas pelos agentes.

As camadas de lógica, prova e confiança, ainda estão em estudo pelo consórcio W3C, e portanto, não estão completas. E para que estas camadas entrem em operação, as camadas inferiores devem estar bem sedimentadas, o que ainda está acontecendo.

## 8 Agentes

Os agentes exercem um papel muito importante na web semântica. “São entidades de *software* que empregam técnicas de inteligência artificial com o objetivo de auxiliar o usuário na realização de uma determinada tarefa, agindo de forma autônoma e utilizando a metáfora de um assistente pessoal” (SOUZA e ALVARENGA, 2004 p. 137).

Os agentes são programas com a função de coletar conteúdos de várias fontes na Web, processar estas informações e compartilhar os resultados com outros programas. O poder total da web semântica será atingido quando estes programas executarem sua função. Segundo Wooldridge e Jennings (1995 apud SOUZA e ALVARENGA, 2004 p. 137) as propriedades desejáveis dos agentes são:

1. Autonomia: agir sem qualquer tipo de intervenção, possuindo controle sobre suas ações e estado interno.
2. Sociabilidade: interagir com outros agentes (humanos ou artificiais) por meio de algum tipo de linguagem de comunicação.
3. Reatividade – perceber alterações em seu ambiente, reagindo a tempo.
4. Proatividade: estar apto a tomar iniciativas, em vez de simplesmente atuar em resposta ao ambiente.
5. Continuidade temporal: estar sendo executado continuamente, possivelmente captando informações sobre o usuário e sobre o ambiente, para melhor desempenhar suas funções.

## 9 Mapeamento de ontologias

Com o crescente número de ontologias publicadas e disponibilizadas ao acesso na web surgiu a necessidade de estruturar estas ontologias para facilitar o seu acesso e aplicação. Uma ontologia simples não é suficiente para suportar as tarefas num ambiente de distribuição, ou seja, o mapeamento de ontologias é uma tecnologia fundamental onde a troca e a partilha de informações são essenciais, por exemplo, a web semântica, comércio eletrónico, interoperabilidade entre sistemas de informação, etc. Múltiplas ontologias precisam ser acessadas para muitas aplicações. Um mapeamento pode prover uma camada comum de várias ontologias que podem ser acessadas e fazer a troca de informações semanticamente. Portanto, o mapeamento ontológico é visto como uma solução para o campo da pesquisa de ontologias (KALFOGLOU e SCHORLEMMER, 2003 p.1)

Então, um mapeamento de ontologias é um processo em que duas ontologias são relacionadas semanticamente em nível conceitual, onde as instâncias da ontologia de origem são transformadas em alvo das ontologias de destino de acordo com a relação de equivalência semântica (SILVA, 2003). Trata-se de uma tarefa que relaciona o vocabulário de duas ontologias que compartilham o mesmo domínio do conhecimento, respeitando as suas estruturas ontológicas originais (KALFOGLOU e SCHORLEMMER, 2003 p.3).

O mapeamento de ontologias é um processo lento, complexo e subjetivo, em que sua completa automação nem sempre é possível, sendo parte desse trabalho feita manualmente. É necessário, portanto, adotar um sistema de apoio, auxiliando na evolução e na adaptação das diferenças entre a interligação e as ontologias. O sistema de apoio deve fornecer mecanismos sofisticados permitindo o controle e promovendo interações longo do processo.

## **10 Validação de ontologias**

“A validação de ontologias pode ser definida como a garantia de qualidade e a adequação de uma ontologia para ser usada ou reusada num contexto específico, para um objetivo específico”(FERNÁNDEZ, CANTADOR e CASTELLS, 2006 p.1).

A validação de ontologias serve como uma prova para a usabilidade da ontologia desenvolvida em seus ambientes e softwares: checar os pedidos, testar o alvo das aplicações e analisar o uso com parceiros (FENSEL et alii, 2003 p. 328).

Segundo Kalfoglou e Hu (2006) a perspectiva de acesso, recuperação e reuso de uma variedade de ontologias no ambiente da web ou da web semântica necessitam de estratégias de avaliação que são: serem abertas aos usuários, transparentes por natureza e com referências para os padrões que irão aderir; flexível para a emenda, fácil de trocar e de adaptar à diferentes usos e num domínio específico, refletir opiniões de vários especialistas. Os autores também acreditam que a certificação e padronização de ontologias para fins comerciais, bem como o licenciamento proverão certas garantias na validade.

## 11 Aplicações da Web Semântica

A literatura nos aponta alguns projetos para aplicações da web semântica, embora a maioria ainda restritos ao ambiente acadêmico. Como as camadas da web semântica ainda não estão totalmente consolidadas, não existe nenhuma indicação oficial do W3C. Dentre esses projetos, destaca-se o concurso “Desafio Web Semântica” ou *Semantic Web Challenge* (<http://challenge.semanticweb.org/>). Este concurso existe desde 2003 e oferece aos participantes a chance de mostrar o melhor da web semântica com a seguinte proposta,

- Ajudar a mostrar para a sociedade o que a web semântica pode prover.
- Proporcionar aos pesquisadores uma oportunidade para demonstrar seus trabalhos e compará-los com os de outros participantes.
- Estimular os pesquisadores a uma meta final mais elevada mostrando o estado da arte anualmente.

A seguir algumas aplicações da web semântica:

1. <http://www.confoto.org> - vencedor do *Semantic Web Challenge* 2005. É um sítio de serviço de consultas e anotações em fotos de conferências.
2. <http://prauw.cs.vu.nl:8080/flink/> – vencedor do *Semantic Web Challenge* 2004. Flink é uma apresentação de trabalhos científicos e conectividade social de pesquisadores da web semântica, em particular, da comunidade de pesquisadores que têm colaborado com

seus trabalhos nas séries da Conferência Internacional de Web Semântica (*International Semantic Web Conference – ISWC*).

3. <http://triplestor.aktors.org/semanticwebchallenge/> - vencedor *Semantic Web Challenge* 2003. Explora o domínio de pesquisas da Ciência da Computação do Reino Unido.
4. [Http://swoogle.umbc.edu](http://swoogle.umbc.edu) – É um sítio de busca de ontologias.
5. <http://del.icio.us/> - Conecta pessoas com seus favoritos, artigos, blogs música, etc. E pode ser acessado de qualquer computador na Web.
6. <http://www.foaf-project.org> – O amigo de um amigo (*Friend of a friend – FOAF*) um projeto que cria páginas que descrevem pessoas no ambiente acadêmico.
7. <http://www.seco.semanticweb.org/> - É um sítio de busca que usa aplicações da web semântica.



## **12 Considerações Finais**

Este trabalho proporcionou uma breve visão sobre ontologias e a web semântica. Apresentaram-se definições, tipos, metodologias, ferramentas e linguagens para a construção de ontologias. Também foi apresentada a web semântica suas camadas e sua aplicação.

Pode-se observar também que existem pendências em relação a implementação total dos objetivos da web semântica proposto por Berners-Lee, Lassila e Hendler (2001). Algumas tecnologias permanecem em estudos, como por exemplo, as camadas de lógica, prova e confiança. Bem como, não existe ainda um consenso para a padronização de metodologias e validação de ontologias.

Sendo os estudos sobre ontologia e web semântica um assunto recente, a literatura encontra-se dispersa e em diferentes níveis de abordagem. Daí a importância de reunir os conceitos básicos, bem como, sintetizar as abordagens dos diferentes autores sobre o tema num único documento no nível de graduação e em língua portuguesa. Uma vez que a literatura é, basicamente, em língua inglesa.

Espera-se, com isto, que este trabalho sirva de base para que estudos futuros aprofundem algumas das abordagens aqui descritas.

### 13 Referência Bibliográfica

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

ARAUJO, Moyses de. **Educação a distância e a Web Semântica:** modelagem ontológica de materiais e objetos de aprendizagem para a plataforma COL. 2003 Tese (Doutorado). Escola Politécnica da Universidade de São Paulo. São Paulo. Disponível em: < [www.teses.usp.br](http://www.teses.usp.br) > Acessado em: maio 2006.

BERNERS-LEE, Tim; HENDLER, James; LASSILA, Ora. The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. **Scientific American**. 2001. Disponível em: <<http://www.cs.nyu.edu/rgrimm/teaching/reading/semantic-web.pdf>> Acessado em: maio 2006.

BORST, W. N. **Construction of engineering ontologies**. 1997. 243 f. Tese (Doutorado). – University of Twente, Enschede, 1997. Disponível em: < <http://doc.utwente.nl/fid/1392> > Acesso em: abril 2006.

CAMPOS, M.L. DE A. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. **Ciência da Informação**, Brasília, v.33, n.1, p.22-32, jan./abril 2004.

CAVALCANTI, C.R. *Indexação e tesouro: metodologia e técnicas*. Brasília, ABDF, 1978.

DACONTA, Michael C.; OBRST, Leo J.; SMITH, Kevin T. **The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management**. Indiana: Wiley, 2003. 281 p.

DAHLBERG, I. Teoria do conceito. **Ciência da Informação**, Rio de Janeiro, v.7, n.2, p.101-107, 1978.

FALBO, Ricardo de A. **Interação de conhecimento em um ambiente de desenvolvimento de software**. 1998. f.188. Tese (Doutorado).

Universidade Federal do Rio de Janeiro. Rio de Janeiro.

FENSEL, Dieter et al. (Editado por). **Spinning the Semantic Web: Bringing the World Wide Web to its Full Potencial**. Cambridge-USA: MIT Press, 2003. 477p.

FERNÁNDEZ, Miriam; CANTADOR, Iván; CASTELLS, Pablo. CORE: A Tool for Collaborative Ontology Reuse and Evaluation. In: PROCEEDINGS OF THE 4<sup>th</sup> INTERNACIONAL EON WORKSHOP (EON'06). Maio 2006. Edinburgo. Reino Unido.

FERREIRA, Aurélio Buarque de Holanda. Novo dicionário da língua portuguesa, 1<sup>a</sup> ed., 15<sup>a</sup> imp., Rio de Janeiro: Nova Fronteira, [1986], p. 999: ontologia.

GUARINO, Nicola. Formal Ontology and Information Systems. In: PROCEEDING OF FOIS'98, TERNTO, ITALY, 6-8 JUNE 1998. Amsterdam, IOS Press, pp. 3-15.

GUARINO, Nicola. Formal Ontology, Conceptual Analysis and Knowledge Representation. In: FORMAL ONTOLOGY IN CONCEPTUAL ANALYSIS AND KNOWLEDGE REPRESENTATION. March 1993. Padova. Italy. Available as Technical Report KSL 93-04, Stanford University.

GRUBER, T. (1996). What is an ontology? [S. l. : s. n.], 1996. Disponível em: <<http://ksl.stanford.edu/kst/what-is-an-ontology.html>> Acessado em: abril 2005.

GRUBER, Thomas R. Toward principles for the Design of Ontologies Used for Knowledge Sharing. In: FORMAL ONTOLOGY IN CONCEPTUAL ANALYSIS AND KNOWLEDGE REPRESENTATION. March 1993. Padova. Italy. Available as Technical Report KSL 93-04, Stanford University.

KALFOGLOU, Yannis; SCHORLEMMER, Marco. Ontology mapping: the state of art. In: THE KNOWLEDGE ENGINEERING REVIEW. Vol. 18:1, 1-31. 2003. Cambridge University Press. United Kingdon. Disponível em: <http://www.ecs.soton.ac.uk/~yk1/publications.html#KEROntoMap>. Acessado em: junho 2006.

KALFOGLOU, Yannis; HU, Bo. Issues with evaluating and using publicly available ontologies. In: PROCEEDINGS OF THE 4<sup>th</sup> INTERNACIONAL EON WORKSHOP (EON'06). Maio 2006. Edinburgo. Reino Unido. Disponível em: <http://www.ecs.soton.ac.uk/~yk1/publications.html#KEROntoMap>. Acessado em: junho 2006.

LIMA, Fernanda. **Modelagem Semântica de Aplicação na WWW**. Rio de Janeiro: PUC-Rio, 2003. f. 155. Tese (Doutorado) – Programa de Pós-Graduação em Informática da PUC-Rio, Rio de Janeiro, 2003.

LIMA, Vânia M. Alves. **Da classificação do conhecimento científico aos sistemas de recuperação de informação**: enunciação de codificação e enunciação de decodificação da informação documentária. São Paulo: Escola de Comunicações e Artes da Universidade de São Paulo. f. 156. Tese (Doutorado) - Universidade de São Paulo, São Paulo, 2004. Disponível em: < <http://www.teses.usp.br/teses/disponiveis/27/27143/tde-06032006-150120/>>. Acessado em: maio 2006.

LYONS, John. **Linguagem e Linguística**, uma introdução. 1<sup>o</sup> ed. Rio de Janeiro: Editora Guanabara Koogan. 1987. 322p.

MAO, Ming. A Overview of Ontology Mapping. Fev. 2005. Disponível em: <http://lms.exp.sis.pitt.edu/~kes/KP/ResOnt-ming.pdf>. Acessado em Junho 2006.

MARTINS, José Junior. **Classificação de páginas na Internet**. 2003. Tese (Mestrado em Ciências da Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e Computação – Universidade de São Paulo, São Paulo. Disponível em: <[http://www.teses.usp.br/teses/disponiveis/55/55134/tde-12092003-1013581/publico/Martins\\_Dissertação.pdf](http://www.teses.usp.br/teses/disponiveis/55/55134/tde-12092003-1013581/publico/Martins_Dissertação.pdf)>. Acessado em: 20 maio 2005.

MURAKAMI, Tiago Rodrigo. **Tesauros e a world wide web**. São Paulo: Escola de Comunicações e Artes da Universidade de São Paulo, 2005. f. 75. Monografia (Graduação) – Universidade de São Paulo, São Paulo, 2005.

NOY, Natalya F.; McGUINNESS, Deborah L. **Ontology Development 101: A Guide to Creating Your First Ontology**. Stanford University: Stanford, 2001. Disponível em: <<http://www.ksl.stanford.edu/people/dlm/papers/ontology101/ontology101->

noy-mcguinness.html> Acessado em: maio 2006.

SABOU, Marta; LOPEZ, Enrico Motta V.; UREN, Victoria. Ontology Selection: Ontology Evaluation on the Real Semantic Web. In: PROCEEDINGS OF THE 4<sup>th</sup> INTERNACIONAL EON WORKSHOP (EON'06). Maio 2006. Edinburgo. Reino Unido. Disponível em: <http://scholar.google.com/scholar?hl=ptBR&lr=&client=firefox-a&q=marta+sabou+enrico+motta+victoria+uren&btnG=Pesquisar&lr=>. Acessado em junho 2006.

SILVA, Nuno A. Pinto da. **Multi-Dimensional Service-Oriented Ontology Mapping**. Vila Real, 2004. f. 279. Tese (Doutorado). Universidade de Trás-os-Montes e Alto Douro. Vila Real. Portugal, 2004. Disponível em: <http://www.dei.isep.ipp.pt/~nsilva/R&D/PhD/PhD.htm>. Acessado em junho 2006.

SMITH, Michael K.; WELTY, Chris; McGUINNESS, Deborah. **OWL Web Ontology Language Guide**. W3C, 2004. Disponível em: <http://www.w3.org/TR/owl-guide/> Acessado em maio 2006.

SOUZA, R. S. ; ALVARENGA, L. A web semântica e suas contribuições para a ciência da informação. **Ciência da Informação**, Brasília, v.33, n.1, p.132-141, jan./abril 2004.

TOSCANO, Wagner; **Minerador Web: Um estudo sobre mecanismos de descoberta de informações na Web**. 2003. Tese (Mestrado em Engenharia) – Escola Politécnica – Universidade de São Paulo, São Paulo. Disponível em : <<http://www.teses.usp.br/teses/disponiveis/3/3141/tde-17122003-150851/> > Acessado em: 22 maio 2005.

TRISTÃO, A.M.D. ; FACHIN G.R.B. ; ALARCON, O.E. Sistema de classificação facetada e tesouros: instrumentos para organização do conhecimento. **Ciência da Informação**, Brasília, v.33, n.2, p.161-171, maio/ago. 2004.

USCHOLD, Mike; KING, Martin. Towards a Methodology for Building Ontologies. In: "WORKSHOP ON BASIC ONTOLOGICAL ISSUES IN KNOWLEDGE SHARING". University of Edinburgh: Edinburgh, 1995.