

# Uma Introdução à Engenharia de Ontologias no contexto da Web Semântica

Anarosa Alves Franco Brandão, Carlos José Pereira de Lucena

Departamento de Informática  
Pontifícia Universidade Católica do Rio de Janeiro - PUC-Rio  
R. Mq São Vicente, 225 - Gávea - Rio de Janeiro - 22453 900 - Brasil

{anarosa, lucena}@inf.puc-rio.br

PUC-RioInf. MCC29/02 Novembro, 2002

## ABSTRACT

The Semantic Web is an evolving research area which strongly depends on ontologies. When developing applications the semantic web, one needs to pay attention to the ontologies they will use. Some ontology description languages are not fitted to some kinds of applications and the interoperability is not a common feature of most of the tools to develop ontologies, which may prevent the importation of existing ontologies. Because of that, several research groups to develop their own ones. In this work we survey several definitions of ontology for the semantic web, describe some ontology development methodologies and relate a case study about developing an ontology for a semantic web-based application, which will generate academic reports using semantic web content from a web site.

**Keywords:** ontologies, semantic web, ontology engineering

## RESUMO

A pesquisa em web semântica tem evoluído nos últimos anos, e depende fortemente de ontologias. Ao desenvolvermos aplicações para a web semântica, precisamos dedicar especial atenção às ontologias que iremos utilizar. Algumas linguagens de descrição de ontologias não são adequadas para alimentar alguns tipos de aplicações e a maioria das ferramentas de desenvolvimento não apresenta interoperabilidade, o que pode impedir a importação de ontologias de uma ferramenta para a outra. Por este motivo, muitos desenvolvedores de aplicações estão optando por desenvolver suas próprias ontologias. Neste trabalho damos uma visão geral das definições para ontologias e web semântica, descrevemos algumas metodologias de desenvolvimento de ontologias e relatamos um estudo de caso sobre o desenvolvimento de uma ontologia para criar conteúdo que vai alimentar uma aplicação para a web semântica, aplicação esta que deve gerar relatórios acadêmicos a partir de conteúdo semântico de um *web site*.

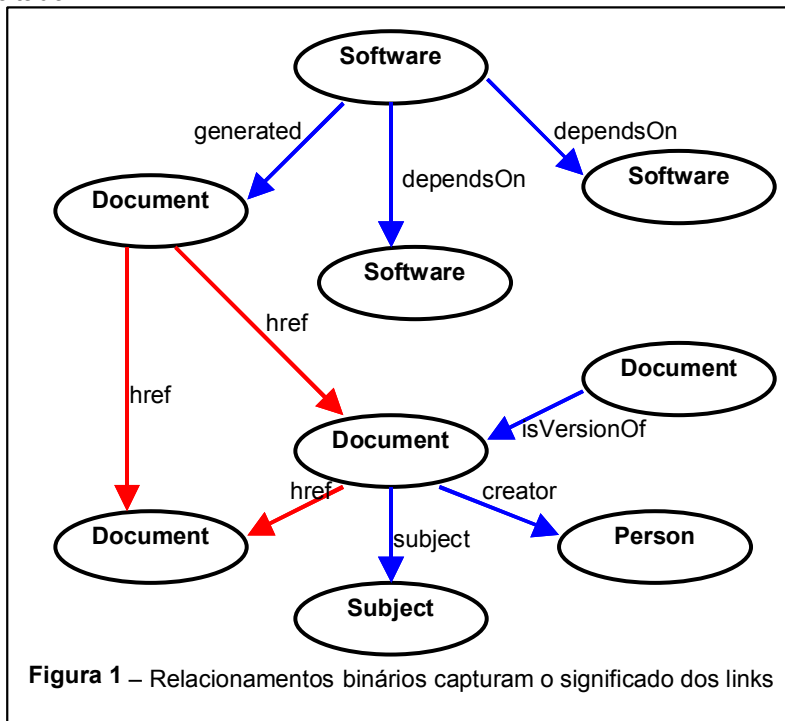
**Palavras-chave:** ontologias, web semântica, engenharia de ontologias

## 1. Introdução

No final do século XX assistimos ao início de uma nova era: a era da informação. O surgimento de redes de comunicação e disseminação de informação com alcance global, como *Internet* e *World Wide Web* (www ou, simplesmente, web), fez com que uma quantidade muito grande de informação estivesse ao alcance de qualquer um, em qualquer lugar, bastando possuir uma linha telefônica e um computador com acesso à internet. Tal fato gerou para o usuário da web o problema de excesso de informação, acompanhado da desorientação e consequente dificuldade de acesso à informação.

As primeiras soluções para o problema de gerenciamento do excesso de informação foram dadas pelas ferramentas de busca (altavista, yahoo, google, etc), baseadas em filtragem da informação através de meta-informações, como palavras-chave, além de heurísticas para categorização. Estas soluções já se mostram ineficientes, dada a quantidade de informação indesejada que o usuário comum recebe ao fazer uma simples busca.

Uma solução para o problema foi proposta por Tim Berners-Lee, o criador da www, ainda no seu documento seminal [8]. Neste documento ele propunha a utilização da idéia de hipertexto semântico, onde os *hiperlinks* também poderiam expressar relacionamentos entre documentos, como por exemplo *isVersionOf*, *dependsOn*, *generated*, etc., esquematizado na figura 1, traduzida do documento citado.



A linguagem para web disponível quando da proposição [8], HTML, não provia suporte para este tipo de construção. O surgimento de XML (eXtensible Markup Language) mostrou um caminho a ser seguido a fim de tornar possível a materialização da proposta de Berners-Lee de uma "web semântica". A idéia básica de XML, o uso de *tags* específicas para descrever tipos de dados contidos num documento, registradas em DTDs (Data Type Description), serviu de inspiração para

a primeira prova de conceito de uma Web Semântica, feita pelo grupo liderado pelo professor Jim Hendler<sup>1</sup>, usando a linguagem de marcação SHOE<sup>2</sup> (Simple HTML Ontology Extensions).

A chave do sucesso desta prova de conceito foi o uso de ontologias, o que possibilitou não só a definição de conceitos num domínio de conhecimento (neste caso, um departamento de Ciência de Computação) de forma a evitar ambiguidades (possíveis em DTDs usadas em XML), como também permitiu a definição de relacionamentos entre estes conceitos e inferências básicas envolvendo-os.

A possibilidade de representação de conteúdo web de forma inteligível e processável por agentes de *software* abriu caminho para o surgimento de novas tecnologias, como RDF<sup>3</sup>, RDFS, DAML+OIL<sup>4</sup>, e mais recentemente, OWL<sup>5</sup>. Estas tecnologias podem orquestrar uma mudança de paradigma na web, qual seja, da recuperação de informação para o gerenciamento e aquisição do conhecimento nela disponibilizado.

Neste contexto, onde sistemas de *software* baseados na web são cada vez mais complexos e a web semântica já se mostrou possível, incorporar possibilidades de uso racional da web semântica às aplicações web é uma tendência natural. Assim, direcionarmos esforços na pesquisa relacionada à engenharia de ontologias no contexto da web semântica nos pareceu uma tarefa de grande interesse e motivou este trabalho.

Neste trabalho apresentamos as principais definições para ontologias e web semântica, descrevemos algumas metodologias e alguns métodos de desenvolvimento de ontologias assim como algumas ferramentas destinadas para edição de ontologias, comparando-as no que diz respeito à criação de ontologias para uso em aplicações web. Finalizamos com um estudo de caso seguido das conclusões finais.

## 2. Ontologias e Web Semântica

O termo Ontologia tem origem na Filosofia, com Aristóteles[24] e está relacionado ao estudo da existência. Na comunidade de Ciência da Computação, mais especificamente na área de Inteligência Artificial, a primeira referência ao termo foi feita em 1991 [26] pelo grupo de pesquisa *DARPA Knowledge Sharing Effort*. Partindo da idéia de componentes reutilizáveis de conhecimento como uma forma de facilitar a construção de sistemas baseados em conhecimento, eles chegam ao termo ontologia como uma forma de descrever o que seriam os tais componentes reutilizáveis.

*"Building Knowledge-based systems today usually entails constructing new knowledge-bases from scratch. It could be done by assembling reusable components. Systems developers would then only need to worry about creating the specialized knowledge and reasoners new to specific task of their system, using then to perform some of its reasoning. In this way, declarative knowledge, problem solving techniques and reasoning services would all be shared among systems. This approach would facilitate building bigger and better systems cheaply... "*

*"An ontology defines the basic terms and relations comprising the vocabulary of a topic area, as well as the rules for combining terms and relations to define extensions to the vocabulary".*

---

<sup>1</sup> <http://www.cs.umd.edu/~hendler>

<sup>2</sup> <http://www.cs.umd.edu/projects/plus/SHOE>

<sup>3</sup> <http://www.w3.org/rdf>

<sup>4</sup> <http://www.daml.org> e <http://www.w3.org/TR/daml+oil-walkthru/>

<sup>5</sup> <http://www.w3.org/TR/owl-ref/>

Desde então o estudo de ontologias na comunidade de Ciência da Computação tem sido bastante fecundo, como se pode ver numa simples busca por palavra-chave na web, com o termo *ontology*. Em 04/09/2002, usando a ferramenta Google, obtivemos 352.000 páginas contendo-o, sendo que nas primeiras páginas as citações eram relacionadas à Ciência da Computação.

Em 1993, Gruber[19] deu a definição mais citada na literatura:

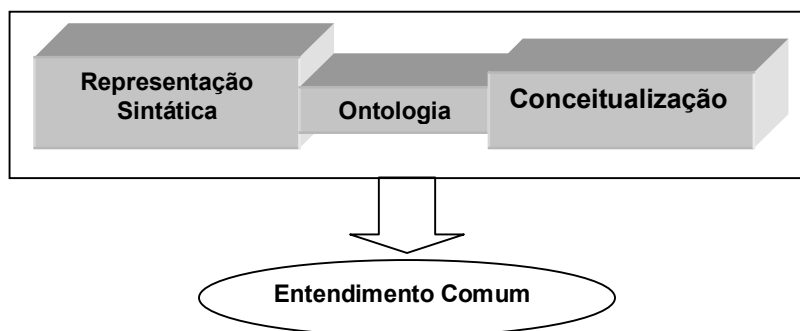
*An ontology is an explicit specification of a conceptualization.*

Nesta definição, o termo conceitualização corresponde aos conceitos, objetos e outras entidades que devem (ou que se presume) existir em uma área de interesse, assim como os relacionamentos entre eles, referenciado por Gruber para [15]. Em 1997, Borst [9] fez uma leve modificação na definição de Gruber, tornando-a, a nosso ver, mais adequada:

*Ontologies are defined as a formal specification of a shared conceptualization.*

A partir da definição acima podemos deduzir que ontologias são cruciais para sistemas de *software* que têm por finalidade a busca ou a combinação/integração de informações provenientes de diversas comunidades. Este é o caso das informações contidas na web, mais precisamente na web semântica, pois ontologias podem representar a semântica dos documentos e permitir que ela seja usada por aplicações web e por agentes de *software*. Ao usarmos ontologias em aplicações web ou habilitarmos agentes de *software* a entendê-las e processá-las, estamos abrindo caminho para que as aplicações que venham a surgir no futuro sejam mais inteligentes, no sentido de considerarmos uma capacidade maior de execução de tarefas num nível conceitual mais próximo do humano [30].

Assim, ontologias possibilitam o preenchimento do "vazio" semântico entre a representação sintática da informação e sua conceitualização, como esquematizado na figura 2.



**Figura 2** – Ontologia como chave para um entendimento comum

O consórcio W3C, a fim de viabilizar a web semântica, iniciou um esforço conjunto a fim de padronizar o modo de se representar formalmente a informação contida na web, definindo o *RDF (Resource Description Framework)*. Assim, na página do W3C referente à web semântica<sup>6</sup>, encontramos a seguinte definição:

*The Semantic Web is the abstract representation of data on the World Wide Web, based on the RDF standards and other standards to be defined. It is being developed by the W3C, in collaboration with a large number of researchers and industrial partners.*

Uma tradução livre para o termo web semântica, dada por Berners-Lee [5] seria:

<sup>6</sup> <http://www.w3.org/2001/sw/>

*Web semântica é a extensão da web obtida via adição de semântica ao formato atual de representação de dados.*

Para o desenvolvimento de uma web semântica, o significado da informação que alimentará a máquina deve ser inteligível através da definição de regras a serem aplicadas aos (meta)dados assim como regras que definam como elas podem ser transformadas em outros (meta)dados.

Pelo que foi descrito podemos observar que ontologias e web semântica convivem em simbiose, dado que até o momento ainda não há outra possibilidade de implementação de solução para a web semântica sem a utilização de ontologias.

### **3. Principais metodologias de desenvolvimento de ontologias**

Como a Engenharia de Ontologias é uma área de pesquisa que está dando seus primeiros passos, ainda não existem metodologias para o desenvolvimento de ontologias que sejam largamente utilizadas e aceitas pela comunidade científica. O que existe são propostas de metodologias, algumas delas mais testadas que outras.

A primeira referência ao termo Engenharia de Ontologias como uma área de pesquisa foi feita em 1996 por Mizoguchi e Ikeda [25]. Como área de pesquisa, podemos dizer que sua base é composta pelas primeiras propostas de metodologias de desenvolvimento de ontologias em 1995, através do relato da experiência obtida durante o desenvolvimento da *Enterprise Ontology* [29] e com o projeto TOVE (TOronto Virtual Enterprise) [20]. Desde então, várias outras propostas surgiram, como o método de desenvolvimento do projeto Esprit KACTUS[3], para o domínio de circuitos elétricos, o projeto METHONTOLOGY [16], [17], um framework para construção de ontologias, dentre outros.

Apesar das várias tentativas de se criar metodologias para o desenvolvimento de ontologias, a prática mostra que a maioria dos grupos de pesquisa cria o seu próprio método de desenvolvimento, dependendo das características da aplicação que pretendem desenvolver usando a ontologia.

A seguir, descrevemos brevemente algumas metodologias:

#### **3.1. Metodologia de Uschol&King**

Esta metodologia teve como base a experiência de desenvolvimento da ontologia *Enterprise*[29], concebida para dar suporte à modelagem de processos empresariais. Os passos de desenvolvimento da metodologia são:

1. Identificação de propósito	Definição do porque construir a ontologia e para que ela será utilizada;
2. Construção da ontologia	2.1 Identificação dos conceitos-chave e dos relacionamentos no domínio de interesse; definição textual e não ambígua dos conceitos e relacionamentos; 2.2 Codificação da ontologia, através da representação dos conceitos e relacionamentos acima em uma linguagem formal; 2.3 Questionamento sobre reutilização de ontologias existentes (deve ser feito em paralelo aos dois passos anteriores);
3. Avaliação da ontologia	Através de critérios técnicos como verificação da especificação de requisitos, validação das questões de competência,

	comparação com o mundo real, etc;
4. Documentação	Deve conter toda a descrição do processo, podendo ter formato diferente para tipos distintos de ontologias, mas que será determinante para o reuso da ontologia desenvolvida.

### 3.2. Metodologia de Grüninger&Fox

Esta metodologia foi baseada na experiência de desenvolvimento da ontologia do projeto TOVE[20], no domínio de processos de negócios e modelagem de atividades. Podemos descrevê-la como segue:

1. Descrição de cenários motivacionais	Os cenários motivacionais são descrições de problemas ou exemplos que não são cobertos adequadamente por ontologias existentes. A partir destes cenários-problema se chega a um conjunto de soluções possíveis que carregam a semântica informal dos objetos e relações que posteriormente serão incluídos na ontologia;
2. Formulação informal das questões de competência	Baseados nos cenários, são elaboradas questões de competência, com a intenção de que seja possível representá-las e respondê-las usando a ontologia a ser desenvolvida;
3. Especificação dos termos da ontologia numa linguagem formal	3.1 Definição de um conjunto de termos/conceitos, a partir das questões de competência. Estes conceitos servirão de base para a especificação numa linguagem formal; 3.2 Especificação formal da ontologia usando uma linguagem de representação de conhecimento, como por exemplo KIF;
4. Descrição formal das questões de competência	Descrição das questões de competência usando uma linguagem formal;
5. Especificação formal dos axiomas	Criação das regras, descritas em linguagem formal, a fim de definir a semântica dos termos e relacionamentos da ontologia;
6. Verificação da completude da ontologia	Estabelecimento de condições que caracterizem a ontologia como completa através das questões de competência formalmente descritas.

### 3.3. METHONTOLOGY

Esta metodologia foi desenvolvida no laboratório de Inteligência Artificial da Universidade de Madri, e é na verdade um framework que, dentre outras funcionalidades, dá suporte à construção de ontologias no nível do conhecimento [16][17]. Associado a este framework existe um ambiente de desenvolvimento de ontologias: ODE (Ontology Development Environment) [17]. Diferentemente das demais, esta metodologia descreve a identificação do processo de desenvolvimento da ontologia dividindo-o em tipos de atividades a serem desenvolvidas, descreve o ciclo de vida de uma ontologia, a partir da evolução de protótipos assim como técnicas específicas para cada atividade executada.

1. Atividades de gerenciamento do projeto	1.1 Planejamento: identificação de tarefas a serem desempenhadas, como estas tarefas devem ser organizadas, quanto tempo e quais recursos elas devem consumir até serem completadas. Esta atividade é essencial quando se pretende fazer reuso de ontologias existentes; 1.2 Controle: atividade que garante que as tarefas
---	--

	planejadas na fase anterior sejam executadas completamente; 1.3 Garantia de qualidade: atividade que assegura que os produtos resultantes das atividades (ontologia, software, documentação) sejam satisfatórios;
2. Atividades orientadas ao desenvolvimento	2.1 Especificação: atividades que definem porque a ontologia será construída, que uso será feito dela e quem serão seus usuários finais; 2.2 Conceituação: atividades de estruturação do domínio de conhecimento da ontologia usando modelos de significado no nível do conhecimento; 2.3 Formalização: atividades de transformação do modelo conceitual da atividade anterior num modelo formal ou semi-computável; 2.4 Implementação: atividades de construção de modelos computáveis numa linguagem computacional; 2.5 Manutenção: atividades de atualização e correção da ontologia.
3. Atividades de suporte - desempenhadas em paralelo ao desenvolvimento	3.1 Aquisição de conhecimento: atividades de aquisição de conhecimento sobre um determinado domínio; 3.2 Avaliação: atividades de julgamento técnico das ontologias, dos ambientes de software associados e da documentação produzida, usando frames de referência; 3.3 Integração: atividades essenciais quando há reuso de ontologias existentes; 3.4 Documentação: atividades de detalhamento claro e exaustivo das fases de desenvolvimento.

Ao pensarmos no desenvolvimento de uma ontologia cuja finalidade é a representação de conhecimento para a web, ou seja, a criação de conteúdo para a web semântica, contamos com algumas dificuldades adicionais, além da opção por este ou aquele método de desenvolvimento. Dentre elas, podemos citar a escolha da linguagem de descrição e do editor de ontologias.

#### 4. Ferramentas de apoio ao desenvolvimento de ontologias

Nesta seção descrevemos algumas ferramentas de apoio ao desenvolvimento de ontologias, as quais testamos durante o nosso estudo de caso. São elas: Protégé 2000, Ontoedit e OilEd. O leitor interessado em outras ferramentas, encontra uma lista delas em [28].

##### 4.1. Protégé 2000<sup>7</sup>

Protégé 2000 é um *framework* Java, *open source*, cuja arquitetura extensível permite gerar e “customizar” ferramentas de criação de bases de conhecimento. A extensão do *framework* pode ser feita através de *plugins*. Ele também pode ser definido como:

- um ambiente integrado de edição de bases de conhecimento, que dá suporte a:
  - a construção de ontologias de domínio;
  - a combinação/integração de ontologias existentes;

<sup>7</sup> <http://protege.stanford.edu/index.html>

- a “customização” de formas de aquisição de conhecimento;
- o armazenamento de uma base de conhecimento sobre determinado domínio;
- uma plataforma que permite a extensão de objetos gráficos de interface para tabelas, diagramas e componentes de animação a fim de acessar outros sistemas baseados em conhecimento embutidos em aplicações;
- uma biblioteca que pode ser usada por outras aplicações a fim de acessar e visualizar bases de conhecimento.

Como principais características, o Protégé 2000 tem um banco de dados de bases de conhecimento com grande escalabilidade (já foi possível observar que não houve perda de desempenho ao se aumentar em potência de 10 a magnitude do número de *frames* (bases) armazenados) e possui uma interface gráfica de fácil uso.

A construção de ontologias usando o Protégé 2000 é feita a partir da criação de uma hierarquia de conceitos (ou classes), que podem ser concretas ou abstratas, dependendo da permissão de instanciação ou não dos mesmos. Além da definição dos conceitos, é possível definir propriedades e relações entre eles.

A principal limitação da ferramenta, quando a intenção for desenvolver ontologias que serão usadas em aplicações web, é a impossibilidade de exportar a ontologia em DAML+OIL. Atualmente existe um *back end* para OIL e está em desenvolvimento um para DAML.

## 4.2. OntoEdit + Ontobroker<sup>8</sup>

OntoEdit é um ambiente de desenvolvimento e edição de ontologias que segue os padrões do W3C, ao oferecer várias interfaces de exportação de ontologias em linguagens como RDF(S), XML, DAML+OIL ou F-Logic. Trata-se de uma ferramenta comercial, cuja versão *free*<sup>9</sup> suporta ontologias com um número limitado de conceitos (máximo de 50).

De forma análoga aos ambientes de desenvolvimento de *software*, também o OntoEdit dá suporte a um método de desenvolvimento de ontologias, quebrando-o em três fases distintas: especificação de requisitos, refinamento e avaliação. Cada uma destas fases usa ferramentas integradas ao ambiente, dadas as características específicas de cada fase.

É possível verificar a consistência da ontologia gerada quando integramos ao ambiente a ferramenta Ontobroker, na versão profissional, e SiRLI na versão *freeware*. Ontobroker é uma máquina de inferência e consulta que processa declarações feitas em F-lógica (subconjunto de lógica de primeira ordem). SiRLI (Simple Logic-based RDF Interpreter) é uma máquina de inferência que define o *core* do Ontobroker. As duas máquinas possibilitam ao usuário fazer inferências sobre fatos, atributos e relacionamentos de conceitos definidos na ontologia, sendo que o Ontobroker possui algumas funcionalidades adicionais. Esta ferramenta é referenciada pelo W3C para a implementação de aplicações para web semântica.

## 4.3. OilEd + FaCT<sup>10</sup>

OilEd é um editor de ontologias *freeware* criado com a intenção de prover uma ferramenta simples e gratuita que demonstrasse e estimulasse o interesse na linguagem DAML+OIL. OilEd não pode ser considerado um ambiente de suporte para o desenvolvimento de ontologias, mas

<sup>8</sup> <http://www.ontoprise.de/com/product.htm>

<sup>9</sup> [http://www.ontoprise.de/com/technology\\_downloads\\_ontoedit.htm](http://www.ontoprise.de/com/technology_downloads_ontoedit.htm)

<sup>10</sup> <http://oiled.man.ac.uk/index.shtml>



sim um bloco de notas. Assim, o suporte para a criação de ontologias de larga escala é bastante limitado, não sendo possível a importação e integração de ontologias existentes.

Apesar destas limitações, a ferramenta tem acoplado a ela um *reasoner*, FaCT, que verifica a consistência da ontologia usando um classificador SHIQ. As inferências possíveis estão relacionadas à disjunção e equivalência entre conceitos, além da verificação hierárquica.

Dada a larga aceitação da ferramenta (estatísticas dos desenvolvedores contam mais de 2000 *downloads* da ferramenta), num futuro próximo OilEd estará disponível como um projeto *open-source*, com licença do tipo GPL, tendo em vista a melhoria da ferramenta pela comunidade científica.

## 5. Estudo de caso

O problema proposto foi a criação de uma ontologia para grupos e projetos de pesquisa em Ciência da Computação, com a finalidade de criar uma prova de conceito para a Web Semântica anotando as páginas do TecComm<sup>11</sup> e de todos os alunos, professores e pesquisadores ligados a ele. Estas páginas anotadas vão alimentar uma aplicação web baseada em componentes e agentes de *software*, para gerar relatórios acadêmicos. O TecComm é um laboratório temático, originado a partir do Laboratório de Engenharia de Software (LES) do Departamento de Informática da PUC-Rio. Este estudo de caso foi executado em conjunto com Leonardo M. Cunha.

### 5.1. Escolha da metodologia de desenvolvimento

A escolha da metodologia a ser usada para o desenvolvimento da ontologia foi feita baseando-se em leituras de artigos e num tutorial do grupo responsável pelo Protégé2000 [27]. Podemos dizer que seguimos um método de desenvolvimento que mistura a metodologia proposta por Uschold e King (sub-seção 3.1) e o método de desenvolvimento proposto pelo grupo do Protégé2000.

Dado que sabíamos o porquê e para quê construir a ontologia, os passos seguidos para definir os conceitos e relacionamentos da ontologia para projetos e grupos de pesquisa em Ciência da Computação, foram os seguintes:

- busca por ontologias que pudessemos reutilizar;
- definição das questões de competência;
- definição dos conceitos básicos da nossa ontologia;
- definição de conceitos mais refinados a partir dos conceitos básicos;
- definição dos relacionamentos entre os conceitos;

Ainda tendo em vista o método de desenvolvimento utilizado, a questão da avaliação da ontologia foi feita a partir das questões de competência e a documentação foi feita durante o desenvolvimento da ontologia.

### 5.2. Escolha da ferramenta de edição da ontologia

A questão da escolha da ferramenta de edição da ontologia é de extrema importância quando o propósito do desenvolvimento for criar conteúdo para a web semântica, pois isto representa uma limitação de linguagem de representação da ontologia. Como já foi dito anteriormente, o consórcio W3C publicou a linguagem DAML como padrão de representação de conteúdo para Web Semântica. Como nosso interesse também era fazer inferência usando os conceitos,

---

<sup>11</sup> <http://www.teccomm.les.inf.puc-rio.br>

relacionamentos e axiomas da ontologia, optamos por utilizar DAML+OIL, também publicada no W3C.

Dada a restrição da linguagem de representação da ontologia, uma ferramenta *freeware* para edição de ontologias conhecida e estável, como **Protégé 2000**, que não exporta ontologias em DAML, apesar de possuir um *backend* para OIL, foi descartada. Esta ferramenta foi testada quando da leitura do método proposto pelo grupo responsável pela ferramenta [27].

Começamos a fazer uma primeira versão da ontologia usando **Ontoedit**, em sua versão *freeware*, mas apesar de dar suporte para DAML+OIL e de possuir uma ferramenta de inferência (SiLRI, que é o *core* do OntoBroker), ela não dava suporte para ontologias que contivessem mais que cinquenta conceitos, exatamente o nosso caso.

Finalmente optamos pela ferramenta **OilEd** com o *reasoner* FaCT, pois era uma ferramenta *freeware*, com suporte para DAML+OIL e que já vinha com um *reasoner* embutido, para verificação da consistência da ontologia. Trata-se de uma ferramenta bastante simples, que não oferece facilidades adicionais para ontologias de larga escala, ou ainda para combinar ontologias existentes a fim de gerar uma ontologia mais geral, como o Protégé 2000, mas que não limita o número de conceitos a serem descritos, assim como os relacionamentos.

### 5.3. Questões de competência

Uma parte importante do desenvolvimento foi a elaboração das questões de competência. Questões de competência são perguntas que se pretende responder a partir de inferências feitas na ontologia.

Fixados alguns conceitos-chave para a nossa ontologia, a saber, projeto, pessoas, recursos físicos, produção científica e tecnológica, empresas parceiras e organizações patrocinadoras, começamos a refinar dentro de cada conceito-chave outros conceitos de interesse no nosso domínio de conhecimento.

As questões de competência foram decisivas na determinação de relacionamentos entre conceitos assim como de novos conceitos. Colocamos abaixo as questões de competência relacionadas ao conceito Projeto e ao conceito Pessoa, onde sublinhamos conceitos e colocamos em *itálico* relacionamentos:

Dado um Projeto, pergunta-se:

- que recursos (*físicos+humanos*) estão *associados* a este projeto?
- quais são os *coordenadores* deste projeto?
- quais são os recursos físicos de *acesso livre e restrito* deste projeto?
- qual a produção deste projeto?
- quais os *patrocinadores* deste projeto?
- quais os *parceiros* deste projeto?

Dada uma Pessoa, pergunta-se:

- a que projetos ela está *associada*?
- qual seu *papel* num projeto?
- qual sua produção num projeto?
- quais são os *papéis* (em geral) que esta pessoa *assume*?
- qual a produção geral desta pessoa?
- quais pessoas estão *subordinadas* a esta pessoa?
- quais pessoas são *chefes* desta pessoa?
- quais pessoas *assumem o mesmo papel* que ela num projeto?
- quais são os *orientandos* desta pessoa?
- quais são os *orientadores* desta pessoa?

- quais produções esta pessoa *tem* numa determinada área de pesquisa?
- quais são os patrocinadores desta pessoa?

A partir destas questões novos conceitos surgiram, como podemos ver na taxonomia que define nossa ontologia, em sua última versão (v0\_6) na figura 3. Nesta figura expandimos apenas o nó referente ao conceito Recurso. Observamos que a classe ConceitoExterno é composta de conceitos usados em nossa ontologia, especialmente nos relacionamentos, e que não são diretamente ligados ao domínio de conhecimento da ontologia.

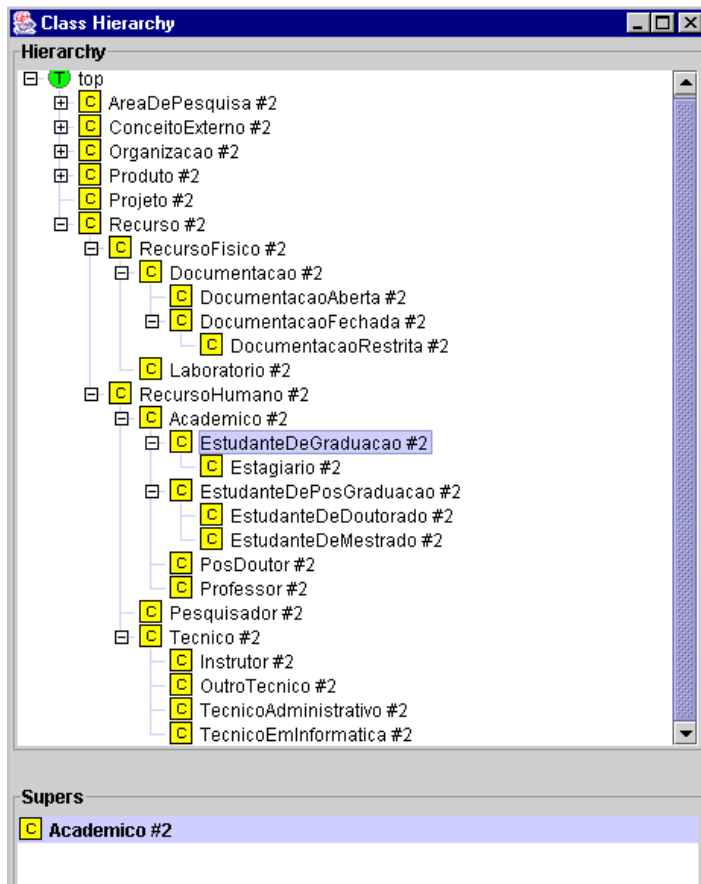


figura 3 – Taxonomia da Ontologia de Grupos de Pesquisa

Quanto aos relacionamentos (chamados propriedades no OIEd), são apenas listados pela ferramenta, não sendo possível discriminar a hierarquia entre eles. Por este motivo, decidimos optar por relacionamentos mais genéricos, como podemos ver na figura 4. Por outro lado, o que seria caracterizado como herança entre propriedades foi descrito como restrição ao associarmos as propriedades aos conceitos, figura 5.

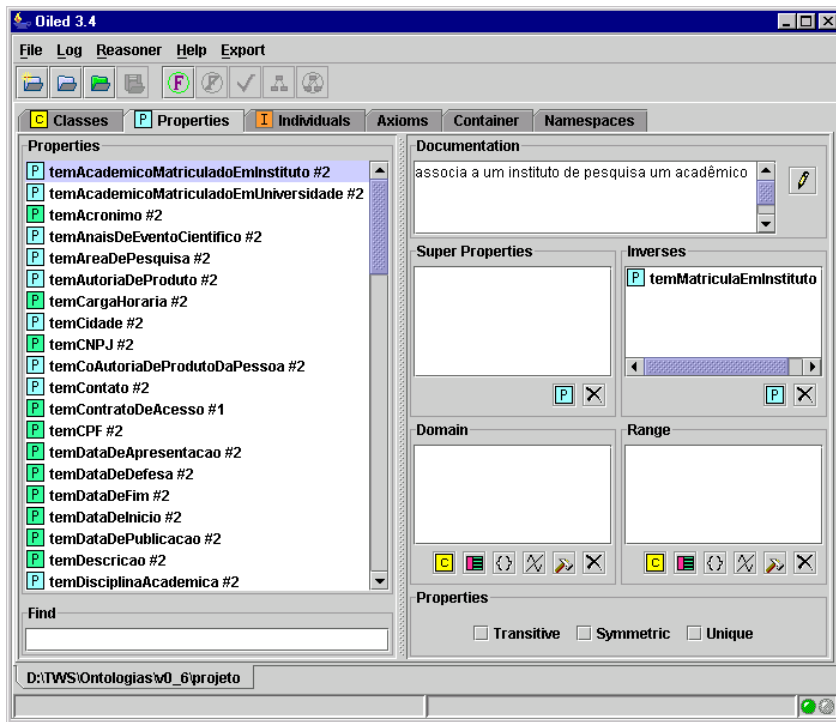


figura 4: Relacionamentos/Propriedades da Ontologia de Projetos e Grupos de Pesquisa

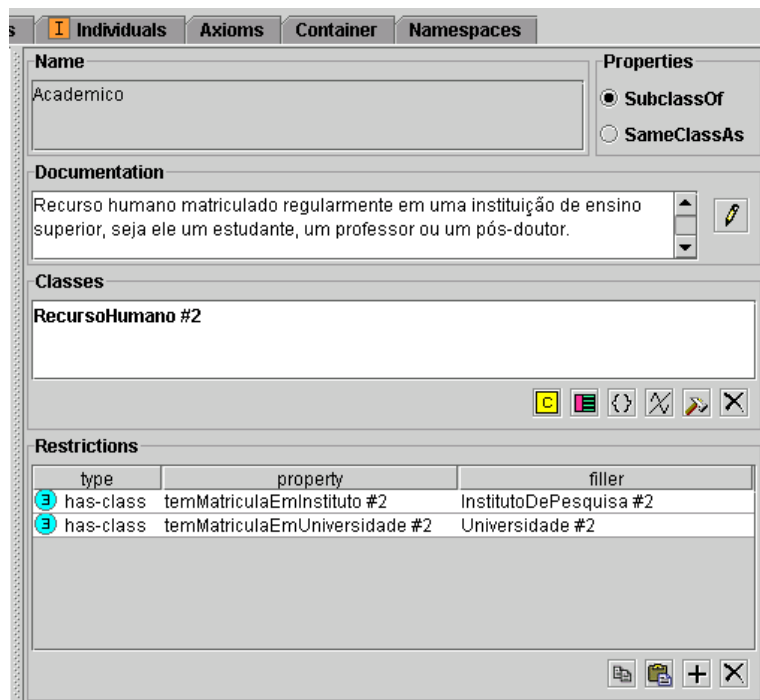


figura 5: Restrições das propriedades do conceito Acadêmico

A ontologia desenvolvida possui um total de 80 classes (conceitos), sendo que delas 77 são primitivas da ontologia e 3 utilizam axiomas para redefiní-las, 81 propriedades (relacionamentos) e 10 indivíduos (instâncias básicas), como vemos na estatística do modelo, figura 6.

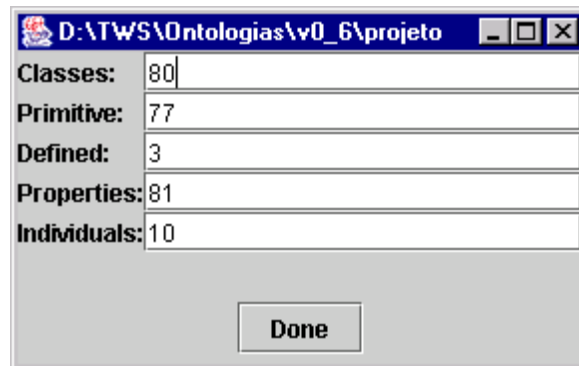


figura 6 – Estatística do modelo da ontologia

Cabe observar que nossa ontologia teve seus conceitos e relacionamentos descritos em língua portuguesa, sendo que alguns deles (Pessoa, Dissertação de Mestrado, Tese de Doutorado, etc) já possuem *links* para conceitos equivalentes em ontologias existentes, descritas em língua inglesa.

#### 5.4. Verificação da consistência da ontologia

A verificação da consistência da ontologia foi feita usando a ferramenta acoplada ao editor, FaCT4.3. As primeiras verificações indicaram algumas inconsistências, que foram sanadas ao redefinirmos os campos de domínio e imagem das propriedades, como podemos ver nas figuras 7a e 7b.

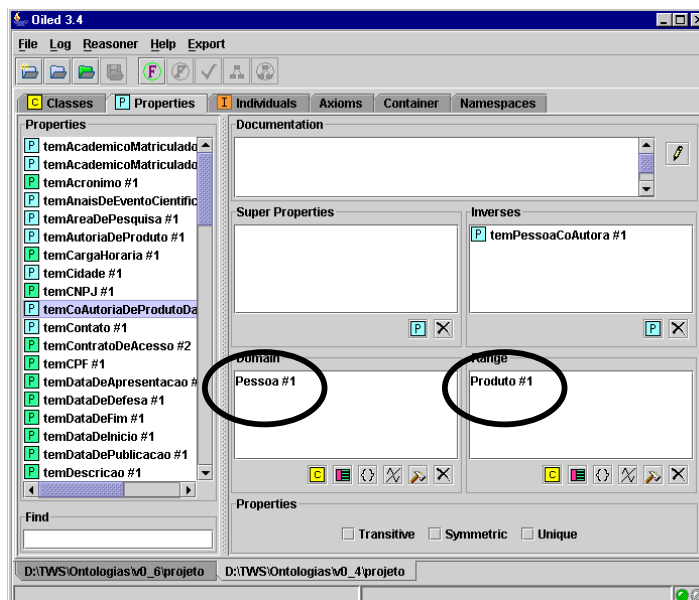


figura 7a: definição que gerou inconsistência

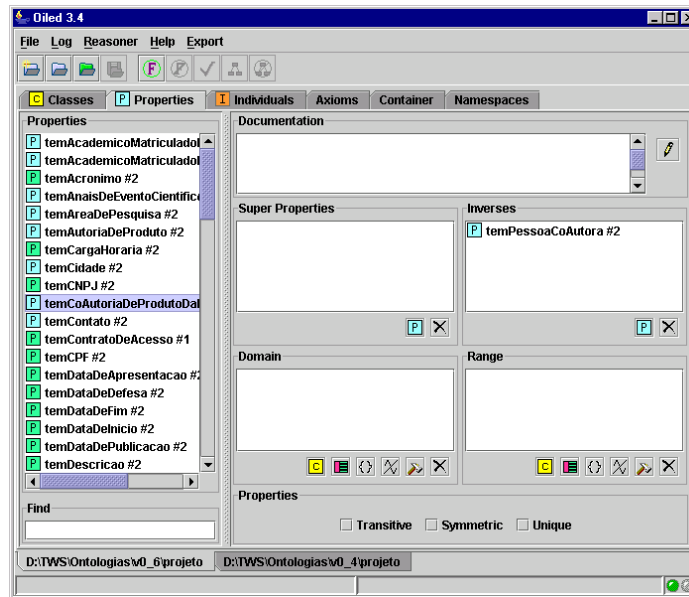


figura 7b: definição final das propriedades

## 6. Considerações finais

A experiência que adquirimos durante o desenvolvimento da ontologia para projetos e grupos de pesquisa nos mostrou que a área de engenharia de ontologias tem avançado, mas nos pareceu que o termo “engenharia de ontologias” ainda não está totalmente consolidado na comunidade científica.

Ao direcionarmos esforços para o desenvolvimento de ontologias para serem usadas na web semântica, juntamos um assunto novo a outro novíssimo e dependente do anterior. Esta junção nos trouxe limitações de desenvolvimento naturais quando se usa novas tecnologias, como por exemplo a restrição da linguagem de descrição da ontologia e a escassez de ferramentas para sua edição e verificação.

Nossa escolha pela linguagem DAML+OIL para descrição da ontologia foi centrada na recomendação do W3C para descrição de conteúdo para a web semântica. Dada esta escolha, o número de ferramentas (gratuitas) para o desenvolvimento de ontologias é bem restrito e, muitas vezes, as ferramentas não dão o suporte necessário para o que se pretende fazer, como foi o caso do Ontoedit em sua versão *freeware*.

O uso de ontologias se coloca como um requisito necessário para que a web semântica se torne tão popular quanto a web atual, pois elas proporcionarão a universalização do entendimento do conteúdo disponibilizado na web. Para que isso ocorra, é preciso que os desenvolvedores de ontologias façam o máximo de reuso das ontologias existentes, seja através de composição de ontologias ou de equivalência entre conceitos.

Finalmente, pudemos observar que o interesse comercial na área de web semântica é muito grande pois quem primeiro dominar a tecnologia de desenvolvimento de aplicações para a web semântica será o grande “player” nas áreas de comércio eletrônico, gerenciamento de conhecimento e outras áreas correlatas.

## 7. Referências:

- [1] Bechhofer, S., Horrocks, I, Goble, C. & Stevens, R: *OilEd: a Reason-able Ontology Editor for the Semantic Web*, Proceedings of KI2001, Joint German/Austrian Conference on Artificial Intelligence, September 19-21, Vienna. Springer-Verlag LNAI Vol. 2174, pp 396--408. 2001
- [2] Benjamins, V.R, et al: *Six Challenges for the Semantic Web*, KR2002, Workshop on Formal Ontology, Knowledge Representation and Intelligent Systems for the World Wide Web, april 2002.
- [3] Bernaras, A., Laresgoiti, I., Corera, J.: *Building and Reusing Ontologies for Electrical Network Applications*, Proceedings of the European Conference on Artificial Intelligence, 1996.
- [4] Berners-Lee, T: *Semantic Web Road map*, <http://www.w3.org/DesignIssues/Semantic.html>, última visita 05/09/2002
- [5] Berners-Lee, T: *The World Wide Web - Past Present and Future*, <http://www.w3.org/2002/04/Japan/Lecture.html>, última visita 20/10/2002
- [6] Berners-Lee, T: *The Semantic Web*, <http://www.w3.org/2002/Talks/04-sweb/slide1-0.html>, última visita em 20/10/2002
- [7] Berners-Lee, T; Hendler, J. & Lassila O.: *The Semantic Web*, *Scientific American*, May 2001
- [8] Berners-Lee, T: *Information Management: A Proposal*, CERN, March 1989, May 1990
- [9] Borst, W.N.: *Construction of Engineering Ontologies*, University of Tweenty, Enschede, NL- Center for Telematica and Information Technology, 1997.
- [10] Bunge, M.: *Tratado de Filosofia Básica - Semântica: Sentido e Referência*, vol 1, EDUSP, 1976.
- [11] Chandrasekaran, B., Jesehpson, J.R. & Benjamins, V.R.: *What Are Ontologies, and Why Do We Need Them?*, IEEE Intelligent Systems, jan-feb 1999, 20-26.
- [12] Devedzić, V.: *Understanding Ontological Engineering*, Comm. of ACM, April-2002, vol. 45, n. 4, 136-144.
- [13] Fernández López, M.: *Overview of Methodologies for Building Ontologies*, Proceedings of IJCAI-99, Workshop on Ontologies and Problem-Solving Methods (KRR5), Stockholm, Sweden, 1999.
- [14] Fernández López, M., Gómez-Pérez, A. & Juristo, N: *METHONTOLOGY: From Ontological Art Towards Ontological Engineering*, Symposium on Ontological Engineering of AAAI, Stanford, California, 1997.
- [15] Genesereth, M.R. & Nilsson, N.J.: *Logical Foundations for Artificial Intelligence*, San Mateo, CA, Morgan Kauffman Publishers, 1987.
- [16] Giles, T.R.: *Introdução à Filosofia*, EDUSP, 1979.
- [17] Gómez-Perez, A.: *A Framework to Verify Knowledge Sharing Technology*, Expert Systems with Application, vol.11, n.4, 519-529, 1996.
- [18] Gómez-Pérez, A.: *Ontological Engineering: A State of the Art*, Expert Update, British Computer Society, Autumn, vol2, n.3, 33-43, 1999.
- [19] Gruber, T.R.: *A Translation Approach to Portable Ontology Specification*, Knowledge Acquisition 5: 199-220, 1993
- [20] Gruninger, M. & Lee, J.: *Ontology: Applications and Design*, Comm. of ACM, February-2002, vol. 45, num. 2, 39-41.
- [21] Guarino, N. & Welty, C.: *Towards a Methodology for Ontology Based Model Engineering*, Proceedings of ECOOP-2000, Workshop on Model Engineering, Cannes, France, 2000.

- [22] Holsapple, C.W. & Joshi, K.D.: *A Collaborative Approach to Ontology Design*, Comm. of ACM, February-2002, vol. 45, n. 2, 42 -47.
- [23] Kim, H.: *Predicting How Ontologies for the Semantic Web will Evolve*, Comm. of ACM, February-2002, vol. 45, n. 2, 48-54.
- [24] Maedche, A.: *Ontology Learning for the Semantic Web*, Kluwer Academic Publishers, 2002.
- [25] Mizoguchi, R. & Ikeda, M.: *Towards Ontology Engineering*, Technical Report AI-TR-96-1, I.S.I.R., Osaka University, 1996.
- [26] Neches, R., Fikes, R.E., et al: *Enabling Technology for Knowledge Sharing*, AI Magazine, 12(3), 36-56, 1991.
- [27] Noy, N. & McGuinness, D.L.: *Ontology Development 101: A Guide to Create Your First Ontology*, 2000
- [http://protege.stanford.edu/publications/ontology\\_development/ontology101.html](http://protege.stanford.edu/publications/ontology_development/ontology101.html), último acesso em set/2002.
- [28] OntoWeb Group, *Deliverable 1.3: A survey on ontology tools*, [http://babage.dia.fi.upm.es/ontoweb/wp1/OntoRoadMap/documents/D13\\_v1\\_0.pdf](http://babage.dia.fi.upm.es/ontoweb/wp1/OntoRoadMap/documents/D13_v1_0.pdf)
- [29] Uschold, M & King, M.: *Towards a Methodology for Building Ontologies*, Workshop on Basic Ontological Issues in Knowledge Sharing, 1995.
- [30] W3C Working Group: *Requirements for a Web Ontology Language*, W3C Working Draft 08 July 2002, <http://www.w3.org/TR/webont-req/>, último acesso em out/2002.