Marcelo Alvim Jorge

PPGCI - Programa de Pós-Graduação em Ciência da Informação
ECI - Escola de Ciência da Informação
UFMG -Universidade Federal de Minas Gerais

ONTOLOGIAS NO SUPORTE A

PORTAIS SEMÂNTICOS

Orientador :

Prof. Marcello Peixoto Bax

BELO HORIZONTE
JUNHO DE 2005

ONTOLOGIAS NO SUPORTE A

PORTAIS SEMÂNTICOS

Marcelo Alvim Jorge

Dissertação apresentada ao Programa de Pós-Graduação da Escola de Ciência da Informação da Universidade Federal de Minas Gerais como requisito parcial para obtenção do grau de Mestre em Ciência da Informação

BELO HORIZONTE
JUNHO DE 2005

RESUMO

Portais semânticos aparecem como uma evolução natural dos tradicionais portais web e podem ser entendidos como portais de informações que utilizam os padrões de representação de informação propostos pelo W3C para a chamada Web Semântica. As ontologias são a base desses novos portais web, propiciando melhor organização das informações, além de outras vantagens. O objetivo deste estudo é compreender como o uso de ontologias promove a organização da informação e agrega valor a esses portais. Para atingir este objetivo, o trabalho investiga as principais ferramentas semânticas baseadas em ontologias existentes na literatura científica e elabora uma ontologia que serve de sustentação para a proposição de um portal semântico a ser utilizado por uma comunidade universitária. Após uma pré-especificação do portal semântico, sua arquitetura funcional é sugerida e as funções semânticas são apresentadas e detalhadas através do uso da ontologia construída anteriormente. As ferramentas semânticas investigadas são utilizadas no contexto de especificação do portal semântico proposto e exemplos ilustrativos, baseados na ontologia elaborada, são extraídos. Com o trabalho espera-se contribuir para a discussão acerca dos benefícios que uma ontologia pode trazer para a organização de conceitos, promovendo a reutilização e o compartilhamento de informações em um determinado domínio do conhecimento.

ABSTRACT

Semantic portals emerged like a natural evolution of the traditional web portals and can be seen like information portals which use information representation patterns proposed by W3C for the so called Semantic Web. The ontologies are the basis of this new web portals and they provide better information organization, besides other advantages. The aim of this study is to understand how the use of ontologies promotes the information organization and aggregates value to these portals. To reach this goal, this study investigates the main ontology-based semantic tools existing in cientific literature and develops an ontology wich is a basis for the semantic portal proposal to be used by a university community. After the semantic portal pre-specification, its functional framework is suggested and the semantic functions are presented in details through the use of the previously developed ontology. The investigated semantic tools are used in the proposed portal specification, and illustrative examples, based on the developed ontology, are extracted. With this study it is hoped to contribute for the discussion about the benefits an ontology can bring to a concept organization, promoting the reuse and information sharing in a specific knowledge domain.

AGRADECIMENTOS

A Deus, uma força maior que me presenteou com o desejo de estudar e avançar sempre em direção ao conhecimento;

A minha família, principalmente meus pais Gláucia e Humberto que sempre mostraram que o mais importante na vida de uma criança é o estudo;

A minhas esposa Daniela, por ter "segurado a barra" de um marido trabalhador, professor e pesquisador;

A minhas filhas Lara e Luana, razão da minha existência;

Aos colegas professores, que muito ajudam na hora em que a gente pensa em desistir da jornada;

Ao meu orientador, que mostrou e iluminou o caminho a seguir;

Aos colegas e funcionários da ECI / UFMG pelo excelente tempo de convívio;

A UNIPAC, por permitir usar o seu nome e seu exemplo neste trabalho.

SUMÁRIO

Assunto **Página**

Resumo i

Abstract ii

Agradecimentos iii

Lista de Tabelas vi

Lista de Figuras viii

CAPÍTULO 1 – Introdução 1

CAPÍTULO 2 - Referencial Teórico 4

2.1) Portais Web 4
2.2) Portais de Comunidades e Portais Educacionais 8
2.3) Portais Semânticos 9
2.4) Ontologias 10
2.4.1) Desenvolvimento de ontologias 12
2.4.2) As ontologias e a representação do conhecimento na perspectiva da Ciência da Informação 13
2.5) A Web Semântica 14
2.6) XML – eXtensible Markup Language 19
2.7) RDF - Resource Description Framework 20
2.8) RDFS - Resource Description Framework Schema 22
2.9) RDQL - Resource Description Framework Data Query Language 23
2.10) OWL – Ontology Web Language 24

CAPÍTULO 3 - Ferramentas semânticas 28

3.1) Ferramentas para criação e manutenção de ontologias 29
3.2) Ferramentas de autoria e marcação semântica 30
3.3) Ferramentas de navegação semântica 32
3.4) Ferramentas de busca semântica 34
3.5) Ferramentas de gestão de conteúdo e repositório de metadados 37
3.6) Utilização de ontologias em ferramentas de colaboração 38

CAPÍTULO 4 - Desenvolvimento da ontologia do portal semântico 40

4.1) A ontologia ONTO-POINT 40
4.2) A construção da ontologia ONTO-POINT 41
4.2.1) Passo 1 - Determinação do domínio e escopo da ontologia 41
4.2.2) Passo 2 - Considerar a reutilização de ontologias existentes 42
4.2.3) Passo 3 - Enumerar termos importantes da ontologia 43
4.2.4) Passo 4 - Definir as classes e a hierarquia de classes 43
4.2.5) Passo 5 - Definir as propriedades das classes 48
4.2.6) Passo 6 - Definir as restrições das propriedades 51
4.2.7) Passo 7 - Criar as instâncias de classes 51
4.3) A ontologia ONTO-POINT em RDF e RDFS 52
4.4) Consultas nas triplas RDF utilizando RDQL 55

CAPÍTULO 5 - O portal UNIPAC 56

5.1) O Portal tradicional da UNIPAC 56
5.2) A proposta do portal semântico da UNIPAC 57
5.2.1) UNI-POINT - O "Ponto de Encontro" da comunidade educacional da UNIPAC 57
5.2.2) Serviços a serem disponibilizados pelo UNI-POINT 59

CAPÍTULO 6 - Funções semânticas do portal 61

6.1) Arquitetura proposta para o portal semântico UNI-POINT 61
6.2) Função de Personalização 63
6.3) Função de Busca 70
6.4) Função de Publicação 74
6.5) Função de Colaboração 82
6.6) Função de Navegação Semântica 83
6.7) Função de Navegação Taxonômica 85

CAPÍTULO 7 - Considerações finais e trabalhos futuros 89

ANEXOS 93

ANEXO I - Classes e Relacionamentos da Ontologia ONTO-POINT 93
ANEXO II - Micro-especificação do Serviço 1 : "montar perfil de usuário e grupo de usuários" 95
ANEXO III - Micro-especificação do Serviço 2 : "encontrar usuário e grupo de usuários" 99
ANEXO IV - Micro-especificação do Serviço 3 : "permitir colaboração entre usuários" 102
ANEXO V - Micro-especificação do Serviço 4 : "publicar conteúdo" 104
ANEXO VI - Micro-especificação do Serviço 5 : “recuperar conteúdo" 107
ANEXO VII - Micro-especificação do Serviço 6: "navegar semanticamente" 110

Referências Bibliográficas 113

LISTA DE TABELAS

Tabela **Página**

Tabela 2.1 - Principais diferenças entre um Portal Tradicional e um Portal Semântico 10
Tabela 2.2 – Exemplo de tripla RDF a partir da ontologia ONTO-POINT 21
Tabela 2.3 – Exemplo de tripla RDF utilizando recursos de outra ontologia 21
Tabela 5.1 – Serviços propostos para o portal semântico UNI-POINT 60
Tabela 6.1 – Serviços do portal semântico e funções semânticas utilizadas 61
Tabela 6.2 – Classe ALUNO e relacionamentos da ontologia, que formam o seu perfil 64
Tabela 6.3 – Classe ALUNO e atributos herdados da superclasse ATOR, que também formam o seu perfil 64
Tabela 6.4 – Classe GRUPO e alguns de seus relacionamentos existentes na ontologia ONTO-POINT, que formam o seu perfil 65
Tabela 6.5 – Classe GRUPO e alguns de seus atributos que também formam o perfil 66
Tabela 6.6 – Relacionamento existente entre as classes ALUNO e DISCIPLINA, registrando instâncias para a propriedade *"estuda_disciplina"* 67
Tabela 6.7 – Realização de inferência entre as classes ALUNO E DISCPLINA, através da utilização da propriedade inversa *"é_estudada_por"* 67
Tabela 6.8 – Utilização da propriedade funcional *"tem_professor_titular"* que permite concluir por inferência que Kênia e Dra. Martins Pedrosa são a mesma pessoa 68
Tabela 6.9 – Utilização da propriedade funcional inversa *"coordena_curso"* que permite concluir por inferência que Dilson e Sr. Campos Júnior são a mesma pessoa 68
Tabela 6.10 – Utilização da propriedade transitiva *"faz_parte"* que permite concluir por inferência que "Bibliotecas Digitais" faz parte da disciplina "Gestão do Conhecimento" 69
Tabela 6.11 – Utilização da propriedade simétrica "é_irmão_de" que permite concluir por inferência que "Pedro de Alcantara" e "Simone de Alcantara" são irmãos 69
Tabela 6.12 – Classes envolvidas no problema "grupos que se interessam em discutir o assunto 'Requisitos de Software'" 70
Tabela 6.13 – Instâncias de classes resultado da pesquisa "grupos que se interessam em discutir o assunto 'Requisitos de Software'" 70
Tabela 6.14 – Novas classes envolvidas no problema "grupos que se interessam em discutir o assunto Requisitos de Software'" 71
Tabelas 6.15 – Novas instâncias de classes resultado da pesquisa "grupos que interessam em discutir o assunto 'Requisitos de Software'" 72
Tabela 6.16 – Exemplo de busca para se encontrar coordenadores de curso das unidades que têm curso com a disciplina "Gestão da Informação" 73
Tabela 6.17 – Consulta original e consulta expandida auxiliada pela Ontologia 73

Tabela 6.18 – Classe DOCUMENTO e alguns de seus relacionamentos existentes na Ontologia ONTO-POINT 76
Tabela 6.19 – Classe PESQUISADOR e alguns de seus relacionamentos existentes na Ontologia ONTO-POINT, que permitem a função de navegação taxonômica 86
Tabela 6.20 – Classes GRUPO e ATOR e alguns de seus relacionamentos existentes na Ontologia ONTO-POINT, que permitem a função de navegação taxonômica 86
Tabela 6.21 – Classes da ontologia ONTO-POINT e seus relacionamentos que permitem a função de navegação taxonômica 87
Tabela I.1 - Resumo das classes e relacionamentos da Ontologia ONTO-POINT 94

LISTA DE FIGURAS

Figura **Página**

Figura 2.1 - Arquitetura de cinco camadas proposta para a Web Semântica por BERNERS-LEE (2001) 18
Figura 2.2 – Exemplo de consulta RDQL, com sintaxe semelhante ao SQL 24
Figura 2.3 – Exemplo de Propriedade Inversa 25
Figura 2.4 – Exemplo de Propriedade Funcional 26
Figura 2.5 – Exemplo de Propriedade Funcional Inversa 26
Figura 2.6 – Exemplo de Propriedade Transitiva 27
Figura 2.7 – Exemplo de Propriedade Simétrica 27
Figura 4.1 - Sugestão de áreas e sub-áreas do conhecimento referente a cursos da Área da Computação. 44
Figura 4.2 – Primeiro fragmento da ontologia ONTO-POINT, montada com a ferramenta Protégé 2000 com extensão OWL 45
Figura 4.3 – Segundo fragmento da ontologia ONTO-POINT, montada com a ferramenta Protégé 2000 com extensão OWL 47
Figura 4.4 - Ontologia ONTO-POINT com detalhe para a classe Aluno (A), seus relacionamentos (B) e suas propriedades (C) 49
Figura 4.5 - Ontologia ONTO-POINT, com detalhe para a classe DOCUMENTO (A), seus relacionamentos (B) e suas propriedades (C) .. 50
Figura 4.6 – Código fonte da ontologia ONTO-POINT elaborado pelo Protégé 2000 em OWL / RDF 52
Figura 4.7 - Modelo de dados RDF, com detalhe para o sujeito CORDENADORCURSO, o predicado *coordena* e o objeto CURSO 53
Figura 4.8 - Modelo de dados RDF incorporado de um Schema 54
Figura 4.9 - Exemplo de consulta em RDQL 55
Figura 4.10 - Possível resultado da consulta em RDQL 55
Figura 5.1 – Um exemplo do que pode ser a página principal do portal semântico UNIPAC 58
Figura 5.2 – Diagrama de Contexto mostrando os serviços disponíveis no portal 59
Figura 6.1 – Arquitetura proposta para o portal semântico UNI-POINT 62
Figura 6.2 – Ontologia ONTO-POINT com detalhe para a classe ALUNO (A), seus relacionamentos (B) e atributos herdados da superclasse ATOR (C) 65
Figura 6.3 – Ontologia ONTO-POINT com detalhe para a classe GRUPO (A), seus relacionamentos (B) e seus atributos (C) 66
Figura 6.4 – Exemplo de busca utilizando a linguagem RDQL para "documentos publicados por membros dos grupos que discutem o assunto Inteligência Artificial" 74
Figura 6.5 – Resultado possível da busca para "documentos publicados por membros dos grupos que discutem o assunto Inteligência Artificial" 74
Figura 6.6 – Sistema de Gestão de Conteúdo Plone com documentos armazenados (A) e a ontologia ONTO-POINT (B). 75
Figura 6.7 – Documento a ser publicado no portal (A) e a ontologia ONTO-POINT . 77
Figura 6.8 – Ontologia ONTO-POINT no centro da tela (A) com seus termos associados ao documento a ser publicado (B). 78

Figura 6.9 – Documento apresentado como resultado da navegação (A) através do conceito Bibliografia (B) escolhido a partir da ontologia ONTO-POINT 79
Figura 6.10 – Marcação Semântica de documento dentro da ferramenta SMORE, com destaque para criação de instância da classe COORDENADORCURSO 80
Figura 6.11 – Marcação Semântica de documento com montagem de triplas RDF (A) e com instâncias criadas para as classes (B) 81
Figura 6.12 – Instâncias de classes representadas em linguagem RDF/XML (A) e relacionamentos da classe DISCIPLINA (B) 82
Figura 6.13 – Classes da ontologia na barra de ferramentas (A) e instâncias marcadas semanticamente para as classes Unidade (B) e Curso (C) ... 84
Figura 6.14 – Informações adicionais em menu dinâmico propostas para a classe Curso 85
Figura 6.15 – Ementas das disciplinas cadastradas no portal (A), encontrados através de navegação na ontologia ONTO-POINT (B) 88
Figura II.1 – Diagrama de Atividades relativo ao serviço 1: "montar perfil de usuário e grupo de usuários" 97
Figura II.2 – Um exemplo do que pode ser a Interface do serviço 1: "montar perfil de usuário e grupo de usuários" 98
Figura III.1 – Diagrama de Atividades relativo ao serviço 2: "encontrar usuário e grupo de usuários" 100
Figura III.2 – Um exemplo do que pode ser a Interface do serviço 2: "encontrar usuário e grupo de usuários" 101
Figura IV.1 – Diagrama de Atividades relativo ao serviço 3: "permitir colaboração entre usuários" 102
Figura IV.2 – Um exemplo do que pode vir a ser a Interface do serviço 3: "permitir colaboração entre usuários" 103
Figura V.1 – Diagrama de Atividades relativo ao serviço 4: "publicar conteúdo" 105
Figura V.2 – Um exemplo do que pode vir a ser a Interface do serviço 4: "publicar conteúdo 106
Figura VI.1 – Diagrama de Atividades relativo ao serviço 5: “recuperar conteúdo" 108
Figura VI.2 – Um exemplo do que pode ser a Interface do serviço 5: “recuperar conteúdo" 109
Figura VII.1 – Diagrama de Atividades relativo ao serviço 6: "navegar semanticamente" 111
Figura VII.2 – Um exemplo do que pode ser a Interface do serviço 6: "navegar semanticamente" 112

CAPÍTULO 1 – Introdução

Comunidades se fortalecem na proporção em que as pessoas envolvidas aumentam seu grau de participação, responsabilidade e passam a se posicionar criando um senso de propósito, objetivos comuns e compromissos de longa duração, estabelecendo um sentimento de identidade e pertencimento à comunidade.

Portais Web se tornaram populares, dando suporte a grupos de indivíduos pela troca de conteúdo entre seus integrantes, suprindo suas necessidades de informação. Surgem, neste contexto, oportunidades para que instituições, especialmente as educacionais, ampliem seu alcance e fortaleçam suas comunidades.

Mais recentemente, portais ditos 'semânticos' aparecem como uma extensão natural dos portais tradicionais e têm atraído a atenção como uma forma de promover um nível mais avançado de organização de informação. Segundo vários autores, portais semânticos proveriam um corpus mais bem estruturado de informações em determinado domínio de conhecimento. Potencializados pelas tecnologias da chamada Web Semântica, atendem melhor às necessidades de compartilhamento de informação de comunidades de usuários (STOJANOVIC et al, 2001), (SPYNS et al, 2002), (REYNOLDS et al, 2004), (ZHANG et al, 2005).

As ontologias (GRUBER, 1993) constituem a base formal dos portais semânticos, categorizando e relacionando conteúdo e permitindo maior interoperabilidade entre sistemas de informação. A noção de ontologia origina-se de disciplinas da Filosofia e passa, no contexto da Ciência da Informação e da Computação, a se referir a um artefato concebido para representar formalmente determinado significado consensual sobre um conjunto de recursos de informação em um sistema de informação. Uma ontologia constituiria uma 'coluna vertebral' que organiza informações e contribui para seu compartilhamento e reuso. Além disso, segundo a literatura da área, o uso de ontologias permitiria arquiteturas de navegação mais sofisticadas, bem como respostas a consultas mais elaboradas, estabelecendo inferências que explicitam informações implícitas em um corpus modelado formalmente.

O objetivo geral desta pesquisa é investigar e compreender, na prática, se o uso de uma ontologia, cuja construção também é proposta, apóia a organização da informação em portais (hipótese da pesquisa) e de que forma se dá tal apoio. Como objetivo específico, espera-se determinar as principais vantagens e dificuldades do uso da ontologia em portais. Buscando realizar o objetivo acima, após levantar e analisar as principais ferramentas baseadas em ontologias apresentadas na literatura científica contemporânea sobre portais Web, propôs-se a arquitetura funcional de um portal educacional para uma comunidade universitária[1]. Tal arquitetura serviu de estudo de caso e forneceu contexto didático para se argumentar a respeito da hipótese de pesquisa.

A pesquisa também sugere que o uso de ferramentas baseadas em ontologias nos portais pode melhorar o acesso às informações e minimizar o problema de sobrecarga de informações. Tais ferramentas são usadas, por exemplo, na personalização do acesso e do trabalho dentro do portal; na melhoria da busca de conteúdo; no apoio à categorização e colocação de metadados; no apoio à navegação com sugestão de fontes de informação; na codificação e publicação de conteúdo de maneira estruturada; além do apoio à comunidade em seus processos de colaboração.

Como não foram encontrados trabalhos que mostram, de maneira geral, como e onde as ontologias podem ser usadas em portais, em especial no campo da educação superior, entende-se aí a contribuição desta pesquisa.

Ressalta-se que, em muitos dos momentos da construção dos enunciados constantes do presente texto, os termos documento, conteúdo, informação e conhecimento foram muitas vezes tratados indiscriminadamente.

Na classificação de WILSON (1999), o método adotado na pesquisa é qualitativo com observação direta[2] e indireta[3] com estrutura emergente; a estrutura emerge a partir do processo de pesquisa, realizado sobre ontologias e ferramentas baseadas em ontologias. Tal estrutura constituiu o contexto para o estudo de caso que, segundo MYERS (1997), é uma investigação empírica de um fenômeno contemporâneo em seu contexto na vida real. Assim, o estudo de caso serviu de contexto para a argumentação, através de exemplos, da hipótese pesquisada.

---

[1] alunos, professores, coordenadores de curso, pesquisadores e visitantes, entre outros.
[2] análise pelo uso direto das ferramentas.
[3] análise da documentação (artigos, tutoriais e textos em geral) das ferramentas.

A pesquisa segmentou-se em quatro fases. Na primeira, foi realizado um estudo das ferramentas semânticas mais citadas na literatura. As etapas que se sucederam nesta fase foram a investigação, na literatura científica, de ferramentas baseadas em ontologia; a seleção das ferramentas úteis ao objetivo do estudo; e a consideração do uso dessas ferramentas no portal proposto. Em uma segunda fase, uma ontologia foi desenvolvida em quatro etapas: estudo do domínio do conhecimento; estudo de metodologias para desenvolvimento de ontologias; investigação de ontologias já existentes para o referido domínio; e finalmente, a construção propriamente dita da ontologia. Posteriormente a essa construção, as ferramentas analisadas na primeira fase foram instanciadas fazendo uso da referida ontologia e foram extraídos exemplos ilustrativos. A terceira fase refere-se à proposta de construção do portal, com a ontologia elaborada servindo de base para a argumentação da hipótese de trabalho inicial. Esta fase pode ser dividida em três etapas que são: a proposição do portal; a especificação de interfaces e funções semânticas; e uma exemplificação de uso da ontologia no portal semântico especificado. Finalmente, uma quarta fase procurou mostrar as conclusões do estudo.

Seguindo a metodologia acima, o documento se organiza da seguinte maneira: no Capítulo 2 serão apontados os referenciais teóricos que norteiam a pesquisa: tecnologia dos portais Web, portais educacionais, portais semânticos, ontologias, Web Semântica e tecnologias que podem ser úteis para a Gestão do Conhecimento em um ambiente educacional universitário. O Capítulo 3 apresenta ferramentas baseadas em ontologias úteis para uma possível implementação do portal proposto. No Capítulo 4 é apresentado o processo de construção de uma ontologia denominada ONTO-POINT, que servirá de base para o portal semântico sugerido em seguida. No Capítulo 5 será apresentado o portal tradicional da Universidade Presidente Antônio Carlos (UNIPAC), e suas limitações no que diz respeito à reutilização e compartilhamento de informações e conhecimento, e será lançada a proposta de confecção do portal UNI-POINT que poderá vir a minimizar esses problemas. Esse capítulo apresenta os serviços que serão propostos para o funcionamento do portal. O Capítulo 6 detalha as funções semânticas propostas, com exemplos de utilização efetiva, dentro do portal, da ontologia construída, pelas ferramentas selecionadas. O Capítulo 7 apresenta as considerações finais do estudo e trabalhos futuros e em seguida são apresentados os anexos e as referências bibliográficas.

CAPÍTULO 2 - Referencial Teórico

Este capítulo contempla os fundamentos teóricos que norteiam a pesquisa, sendo apresentados os portais Web, portais educacionais e portais semânticos. São definidas as ontologias e as metodologias para a construção das mesmas, além de destacar a importância destes artefatos na representação do conhecimento no contexto da Ciência da Informação. O capítulo termina com conceitos da Web Semântica e suas tecnologias, que podem ser úteis para a Gestão do Conhecimento em um ambiente educacional universitário.

2.1) Portais Web

A World Wide Web ou www (chamada a partir de agora neste trabalho apenas de Web) apresenta um número crescente de conteúdo, de usuários e de fontes de informação, com os portais Web se tornando muito populares, suprindo as necessidades de informação dos seus usuários.

De acordo com PINHEIRO e MOURA (2004), portal é um conceito recente, que tinha um significado bem diferente há alguns anos atrás, “pois era simplesmente atribuído a mecanismos de busca, cujo propósito era facilitar o acesso à informação na Web”.

Nos últimos tempos, os portais se tornaram populares na Internet. "A capacidade dos portais em capturar, organizar e compartilhar informação é muito importante para as organizações intensivas em conhecimento dos tempos modernos", afirmam TERRA e GORDON (2002).

Os autores também afirmam que

> "um portal de informação na Web agiliza e melhora os ciclos de decisão complexos de trabalhadores do conhecimento e pode incentivar níveis mais profundos de colaboração entre os envolvidos.(...) Trata-se de um importante instrumento de gestão de informação e conhecimento.”

Segundo REYNOLDS et al (2004),

> ″portais de informação baseados na Web provêm um ponto de acesso a um corpo de informações integrado e estruturado sobre determinado domínio. (...) Eles podem variar de domínios amplos a domínios pequenos e específicos″.

BAWDEN (2001) define sobrecarga informacional[4] como "um estado no qual a informação disponível e potencialmente útil torna-se um obstáculo, ou atraso, ao invés de uma ajuda". O excesso de informação está associado à perda de controle sobre a informação e à inabilidade em usar efetivamente a informação. Como resultado, tem-se trabalho ineficiente e, eventualmente, até risco para a saúde.

TERRA e BAX (2003) afirmam que

> "embora o problema da sobrecarga informacional não seja novo, podendo ser encontrado de forma recorrente ao longo da história, suas maiores conseqüências são recentes. Os portais vêm apresentando uma tentativa de reduzir o excesso de informações que foi exacerbado pela generalização do uso de aplicações *desktop*[5] nos últimos tempos, principalmente programas de e-mail e Web".

Segundo os autores, "parece não existir uma técnica ou ferramenta que poderá resolver o problema do excesso de informações como um todo". Entretanto, consideram que as soluções propostas se enquadram, de modo geral, em duas categorias: a gerencial e a tecnológica. Sob o ponto de vista tecnológico, os portais oferecem algumas funcionalidades importantes, descritas a seguir, e que podem contribuir para a minimização desta sobrecarga.

A. Personalização

Conforme TERRA e BAX (2003), personalização, no contexto de portais corporativos, significa "adaptação das ferramentas às necessidades dos usuários em acessar rapidamente a informação mais relevante para a execução das atividades profissionais do seu dia-a-dia". Um exemplo de personalização é a montagem de perfis dentro dos portais, que possibilita a predefinição das principais aplicações ou quais classes de informações, ou assuntos, mais interessam aos usuários.

Reconhecendo as vantagens da personalização no ambiente dos portais de comunidades, KOCH e LACHER (2000) dizem que

> "um fluxo aberto de conhecimento poderia ser amplamente simplificado, se:

[4] do termo inglês *information overload*
[5] aplicações possíveis de serem utilizadas por computadores pessoais de mesa.

- Existisse um perfil de usuário para cada pessoa membro de uma comunidade, que pudesse ser usado em portais de outras comunidades;
- Fontes de Informações das comunidades pudessem ser utilizadas em diversos outros portais de outras comunidades;
- Os usuários pudessem personalizar o modo de visualização da informação das diferentes fontes".

B. Mecanismos de Busca

Conforme TERRA e BAX (2003), "um dos objetivos centrais dos portais é auxiliar os usuários a encontrar facilmente as fontes de informação mais relevantes no momento exato em que mais necessitem delas". Assim, segundo os autores, "os mecanismos de busca são ferramentas extremamente úteis, cujo maior objetivo é fornecer os resultados mais relevantes para a consulta no menor espaço de tempo".

Como resposta a circunstâncias individuais e ao crescimento exponencial da quantidade de informação e tipos de fontes de dados, muitos avanços têm ocorrido nos mecanismos de busca. Segundo BROADER (2002), os mecanismos de busca podem ser organizados em três gerações :

a) Mecanismos de busca de 1a. geração:

a.1) Busca por palavra-chave ou frase exata;

a.2) Busca booleana (aplicação de operadores lógicos AND, OR, NOT);

a.3) Busca com filtros colaborativos (consideram documentos similares relevantes)

b) Mecanismos de busca de 2a. geração:

b.1) Busca baseada em popularidade (*ranking* dos *link*s que serão mostrados);

b.2) Busca por conceitos (resultados podem não conter as palavras de busca).

c) Mecanismos de busca de 3a. geração:

c.1) Busca contextualizada;

c.2) Busca em linguagem natural;

c.3) Busca em bases de conhecimentos;

c.4) Busca por afinidades;

c.5) Buscas por mapeamento visual.

Os motores de busca atuais, normalmente de segunda geração, levam em conta a análise hipertextual das páginas como *links*, estrutura, tamanho de fontes, entre outros, para classificar a relevância dos resultados.

Segundo PINHEIRO e MOURA (2004),

> "o maior problema dos mecanismos de busca é a grande quantidade de documentos que retornam para o usuário. (...) Neste cenário, as ontologias aparecem como uma boa solução para ajudar na contextualização dos termos de pesquisa, já que ela possibilita a associação de conceitos e propriedades de um domínio específico".

C. Categorização de Assuntos

Com o objetivo de simplificar a navegação e a procura por informações e conhecimento, as organizações, se quiserem atender aos requisitos da chamada Web Semântica, precisam desenvolver categorias que sejam facilmente compreendidas pelos usuários da comunidade, além de serem representativas para as máquinas que buscarão pelas informações sob a forma de agentes de software.

A categorização adiciona informação fundamental para a classificação dos documentos, para que esses sejam organizados de acordo com uma taxonomia [6], ou uma ontologia, como no caso desse estudo, e sejam mais facilmente encontrados posteriormente.

D. Publicação e distribuição de conteúdo

Segundo TERRA e GORDON (2002),

> "um objetivo importante de qualquer solução de portal deveria ser a descentralização e delegação de poder para que os usuários pudessem facilmente incluir informação e conhecimento no sistema e ter essas entradas eficientemente e rapidamente distribuídas para grupos e para toda a comunidade".

No cenário dos portais semânticos, qualquer publicação deve ser realizada semanticamente, ou seja, devem ser feitas anotações em seu conteúdo, denominadas marcações, que são armazenadas em um repositório de metadados.

Portais semânticos privilegiam a publicação de conteúdo de forma que o próprio autor o descreva inserindo anotações no material. Segundo GOLBECK et al (2003), "a marcação semântica pode ser utilizada como estrutura para a informação nos sítios Web".

Segundo as autoras, enquanto existem muitas ferramentas para criação e manipulação de ontologias, poucas existem com o objetivo de facilitar o trabalho de marcação semântica (anotação) e publicação dos autores na Web.

GOLBECK et al (2003) relatam que "para o usuário final, não há diferenças entre um sítio normal e outro sítio marcado semanticamente". Isto significa que as pessoas que estão vendo uma página, por exemplo, não têm necessidade de conhecer a tecnologia que está sendo utilizada por detrás da aparência da página, sendo que a usabilidade não pode e não deve ser definitivamente afetada.

2.2) Portais de Comunidades e Portais Educacionais

Em termos gerais, uma comunidade é um grupo de pessoas que compartilham algum interesse ou pertencem a um mesmo contexto. MYNATT et al (1997) afirmam que "uma comunidade é um grupo social que apresenta várias características, entre elas, relações espaciais compartilhadas, convenções sociais, um senso de pertencimento e participação à comunidade e um ritmo crescente de interação social".

Comunidades são boas fontes de conhecimento em diversas áreas de interesse. Para promover o fluxo do conhecimento, sistemas de apoio a comunidades são necessários (BORGHOFF e PARESCHI, 1998). Em contraste com aplicações de *groupware*[7], ferramentas de apoio a comunidades enfocam grupos grandes e não-homogêneos, sendo que a maioria de seus membros não trabalham em projetos comuns.

LACHER et al (2001) afirmam que "portais de comunidades Web são coleções de páginas Web, dinâmicas ou estáticas, através das quais os membros da comunidade podem contribuir e obter informações de outros membros, assim como de outras fontes de interesse dos membros da comunidade".

Os portais educacionais herdam todas as funcionalidades de um portal de informações de comunidades na Web, servindo de apoio a uma comunidade educacional, como no caso estudado, oferecendo facilidades de navegação e pesquisa, bem como acesso a serviços e aplicações para usuários em particular ou grupos de usuários. Neste cenário, os

---

[6] classificação científica, classificação de vocábulos em grupos ou categorias

[7] tecnologia projetada para facilitar o trabalho em grupos. Pode ser utilizada para comunicar, cooperar, coordenar, resolver problemas, competir ou negociar.

portais educacionais são portais Web que servem como ponto de entrada para a Internet para professores, pais, estudantes, pesquisadores e as demais pessoas que fazem parte de uma comunidade educacional, favorecendo a reutilização e um efetivo compartilhamento das informações e do conhecimento entre os interessados.

A partir do momento que uma instituição educacional é uma comunidade e os portais são comunidades virtuais na Internet, eles estão sendo vistos como uma oportunidade para as instituições de ensino estenderem o seu alcance.

2.3) Portais Semânticos

Mais recentemente, portais semânticos surgiram como uma extensão natural dos portais tradicionais e têm atraído a atenção de pesquisadores e das organizações. Um portal semântico pode ser entendido como um portal de informações que usa padrões da Web Semântica que definem novas formas sobre como as informações devem ser representadas.

REYNOLDS et al (2004) afirmam que "existem muitas vantagens no uso de padrões da Web Semântica para o projeto de portais de informação". Estas vantagens estão resumidas na Tabela 2.1.

Portal tradicional	Portal Semântico
Pesquisa por texto livre e hierarquia estável de classificação	Pesquisa multidimensional por significado de uma ontologia rica de domínio
Informações organizadas em registros estruturados, encoraja projeto top-down e manutenção centralizada	Informações semi-estruturadas e extensíveis, permitem evolução bottom-up e alterações descentralizadas
Comunidade pode adicionar informações e anotações dentro da estrutura definida no portal	Comunidade pode adicionar nova classificação e esquemas organizacionais e estender a estrutura da informação
O conteúdo do portal é armazenado e manipulado de forma centralizada	O conteúdo do portal é armazenado e manipulado por uma rede descentralizada de organizações provedoras e indivíduos. Múltiplas agregações e visões do mesmo dado são possíveis
Provedores fornecem dados através de formulários específicos dos portais. Cada portal é suprido e mantido separadamente	Provedores publicam dados em formulários reutilizados pela organização em múltiplos portais. Modificações estão sob o controle deles

Portal preparado para o acesso humano. Mecanismos separados são necessários para compartilhar conteúdo com uma organização parceira	Estrutura da informação é diretamente acessível a máquinas para facilitar a integração entre portais.

Tabela 2.1 - Principais diferenças entre um Portal Tradicional e um Portal Semântico

REYNOLDS et al (2004) afirmam que "os padrões da Web Semântica permitem novas abordagens para o projeto de portais. (...) Juntos, esses padrões permitem uma abordagem mais descentralizada para a arquitetura de portais". Os padrões aos quais o autor se refere foram desenvolvidos pelo W3C (*World Wide Web Consortium*) [8] e são: XML, RDF e OWL.

XML *(eXtensible Markup Language)* permite adicionar estrutura a documentos, através da criação de tags, que são rótulos ocultos que marcam os documentos. A linguagem XML será detalhada na Seção 2.6.

RDF *(Resource Description Framework)* provê um formato flexível e extensível para representar metadados[9]. Será detalhado na Seção 2.7.

OWL *(Ontology Web Language)* é uma representação explícita de ontologias usadas para classificar e estruturar informação. A linguagem se baseia em RDF e RDFS (RDF *Schema*), que por sua vez, utilizam a sintaxe de XML. As ontologias serão discutidas na Seção 2.4, a RDF, na Seção 2.7 e a RDFS, na Seção 2.8, enquanto OWL será detalhada na Seção 2.10 deste mesmo capítulo.

2.4) Ontologias

STAAB et al (2000) afirmam que

> "uma ontologia serve como uma coluna vertebral semântica para acessar o conhecimento no portal, para contribuir com informações, tanto quanto para desenvolver e manter o portal. (...) Em particular, a ontologia permite consultas mais flexíveis e inferências sobre o conhecimento".

Em sua origem, na Filosofia, o termo ontologia é usado para designar a teoria a respeito da natureza da existência humana. Foram os pesquisadores de Inteligência Artificial (IA) e de Web Semântica que incorporaram o termo ao seu jargão. No começo

[8] http://www.w3c.org
[9] dados que descrevem dados mais complexos

dos anos 90 as ontologias se tornaram um campo popular de pesquisa entre as comunidades de pesquisadores de Inteligência Artificial, incluindo Engenharia do Conhecimento, Processamento de Linguagens Naturais e Representação do Conhecimento. Os autores de IA adotaram então o termo para descrever os modelos do mundo, ou seja, descrever aquilo que pode ser computacionalmente representado sobre o mundo. Mais recentemente, a noção de ontologia espalhou-se e o estudo das ontologias está presente em campos como Integração Inteligente de Informações, Recuperação de Informações na Internet, Compartilhamento e Reutilização do Conhecimento.

Para GRUBER (1993), "uma ontologia é uma especificação formal e explícita de uma conceitualização compartilhada". FENSEL (1999) descreve esse conceito em partes afirmando que uma 'conceitualização' refere-se a um modelo abstrato de algum fenômeno no mundo que identifica conceitos relevantes daquele fenômeno. GUARINO (1997) ainda comenta que uma 'conceitualização' explica o significado pretendido dos termos usados para indicar relações relevantes. 'Explícito' significa que os tipos de conceitos usados e os relacionamentos entre os conceitos são definidos explicitamente. 'Formal' refere-se ao fato de que uma ontologia deve ser legível para as máquinas. 'Compartilhada' reflete a noção de que uma ontologia captura o conhecimento consensual, isto é, o conhecimento não é restrito a algum indivíduo, mas aceito por um grupo.

A razão para as ontologias se tornarem tão populares parece ser o fato de que elas prometem uma compreensão comum e compartilhada de um certo domínio, que pode ser comunicada entre pessoas e máquinas, permitindo a reutilização do conhecimento e a organização da informação.

Segundo GUARINO (1998), as ontologias podem ser classificadas de acordo com o seu nível de generalidade em ontologias de alto nível, ontologias de domínio, ontologias de tarefas e ontologias de aplicações. Segundo o autor, as ontologias de alto nível descrevem conceitos gerais como espaço, tempo, assuntos, objetos, eventos, ações, entre outros, todos independentes de um problema ou domínio particular. Por sua vez, as ontologias de domínio e de tarefas descrevem, respectivamente, o vocabulário relacionado a domínios genéricos como medicina ou automóveis, por exemplo, ou a tarefas e atividades genéricas, como por exemplo, diagnósticos ou vendas, através da especialização de termos introduzidos nas ontologias de alto nível. As ontologias de aplicação descrevem conceitos

que são dependentes de um domínio particular ou de tarefa, normalmente especializações de ambas as ontologias, de domínio e de tarefas.

2.4.1) Desenvolvimento de ontologias

Segundo NOY e McGUINESS (2001), compartilhar entendimentos comuns da estrutura da informação entre pessoas e agentes de software é uma das razões para se desenvolver uma ontologia. Elas também afirmam que

> "não existe uma metodologia ou uma maneira correta para o desenho de ontologias (...) Decidir para que vamos usar a ontologia e quão detalhada ou generalista ela deve ser é fundamental para a tomada de decisões de modelagem em todo o processo".

As autoras destacam algumas regras fundamentais que se deve ter sempre em mente quando do projeto e construção de uma ontologia:

1) Não existe uma maneira correta de se modelar um domínio. Existem sempre várias alternativas e a melhor solução quase sempre depende da aplicação que se tem em mente e os acréscimos que são possíveis de serem previstos;
2) O desenvolvimento de uma ontologia é necessariamente um processo iterativo. Depois de elaborada uma versão inicial de uma ontologia, ela deve evoluir, sendo certa a necessidade de revisão da ontologia inicial. Esse processo iterativo deve continuar durante todo o ciclo de vida da ontologia;
3) Os conceitos da ontologia devem ser próximos aos objetos (físicos ou lógicos) e aos relacionamentos do domínio de interesse. É necessário sempre lembrar que uma ontologia é um modelo da realidade do mundo e os conceitos da ontologia devem refletir essa realidade.

Os blocos básicos de construção das ontologias são os conceitos, que são organizados tipicamente em uma hierarquia de classes. Estes conceitos podem ter propriedades que estabelecem relações com outros conceitos.

Ainda segundo NOY e McGUINESS (2001),

> "em termos práticos, desenvolver uma ontologia inclui definir as classes da ontologia; arranjar as classes em uma taxonomia (hierarquia

> de superclasses e subclasses); definir propriedades e descrever valores permitidos para essas propriedades; e preencher os valores para as propriedades das instâncias".

Muitas áreas do conhecimento estão desenvolvendo ontologias padronizadas que possibilitam aos especialistas de domínio a anotação e o compartilhamento de conhecimento em suas respectivas áreas. Uma ontologia permite a definição de um vocabulário comum para usuários que precisam de compartilhar informações em um domínio de conhecimento. Isto também inclui definições de conceitos básicos do domínio e relacionamentos entre eles, que também podem ser interpretados por máquinas, atendendo às demandas e às necessidades da chamada Web Semântica, a ser detalhada na Seção 2.5. Segundo MAEDCHE et al (2002), "a principal idéia da Web Semântica é tornar a informação acessível a humanos e agentes de software em uma base semântica".

2.4.2) As ontologias e a representação do conhecimento na perspectiva da Ciência da Informação

ALVARENGA (2003) afirma que

> "ao se refletir sobre o advento e uso intensivo das novas tecnologias, na perspectiva da Ciência da Informação, constata-se positiva turbulência no campo de conhecimento, especialmente no que tange à representação, à armazenagem e recuperação de informações, áreas intensamente relacionadas à cognição humana".

Segundo a mesma autora,

> "à medida que as tecnologias da informação foram sendo criadas, disponibilizadas e aperfeiçoadas, os sistemas de representação e recuperação de informações documentais assistiram a uma extrapolação dos limites dos tradicionais catálogos referenciais primeiramente em fichas, alcançando posteriormente as bases de dados online".

Partindo-se da definição de que representar significa o 'ato de colocar algo no lugar de' (VICKERY, 1986), o nível primário de representação é classificado por este autor como sendo :

> "a representação feita pelos autores no momento da expressão dos resultados de seus pensamentos, estes derivados de observações metódicas da natureza e dos fatos sociais, utilizando-se das linguagens disponíveis no contexto da produção e comunicação de conhecimentos".

Ao definir representação secundária, o mesmo autor diz que

> "após produzidos, os registros de conhecimentos constantes de documentos passam a integrar acervos de arquivos, bibliotecas, serviços ou centros de documentação/informações, sendo então novamente representados, visando-se à sua inclusão em sistemas documentais referenciais".

CAPURRO (1991) afirma que,

> "de acordo com o paradigma da representação, os seres humanos são conhecedores e observadores da realidade externa. O processo de conhecimento consistiria na assimilação de coisas por meio de suas representações na mente/cérebro do sujeito cognoscente".

Essas representações, uma vez processadas e codificadas no cérebro, seriam comunicadas a outras mentes e/ou armazenadas e processadas por máquinas. Segundo o autor, "em nossos dias, tornar-se-ia cada vez mais aceitável a hipótese de que também as máquinas seriam capazes de alcançar níveis de interpretação e uso da informação, além do simples processamento".

VICKERY (1986) também escreveu sobre como o advento do computador e a conseqüente mudança de paradigma quanto à própria materialidade dos documentos mudaria o trabalho de tratamento da informação. O autor cita o surgimento de novas situações, fatores e elementos que ampliariam o campo de pesquisa relativo à biblioteca, como por exemplo, as estruturas semânticas e sintáticas da linguagem natural, e a representação do conhecimento utilizando-se da inteligência artificial.

2.5) A Web Semântica

Segundo a enciclopédia Wikipedia[10], o termo semântica origina-se no grego clássico da palavra σημαντικ.ος, ου (sēmantikos = semântico = do signo) derivada de σημα, σημαντος (sēma = signo). Em sua primeira acepção, o termo semântica designa os procedimentos de interpretação que permitem partir de um texto sintaticamente analisado e

[10] http://en.wikipedia.org/wiki/semantics

obter um conhecimento. Em sua segunda acepção, o termo semântica designa a Lingüística, disciplina do campo de conhecimento que estuda esses procedimentos de interpretação. Em sua terceira acepção, o termo é usado para designar o estudo desses procedimentos de interpretação. Em sua quarta acepção, o mesmo termo significa o produto da interpretação humana, como, por exemplo, na expressão 'dar uma semântica para o texto'. Em sua quinta acepção, que interessa a este estudo, o termo é usado como sinônimo de ontologia.

Segundo BERNERS-LEE (2001),

> "grande parte do conteúdo da Web de hoje é projetada para 'humanos' lerem, não para programas de computador manipulá-los significativamente, pois as máquinas ainda não têm um método seguro para processar a semântica".

De acordo com ele, o objetivo da Web Semântica é

> "trazer estrutura para o conteúdo das páginas Web, aumentando-lhes o significado, criando um ambiente onde agentes de software podem fluir de página a página, cumprindo sofisticadas tarefas para os seus usuários".

GOLBECK et al (2003) dizem que

> "a Web Semântica é uma visão de futuro da Web que irá dar significado a todos os dados, assim como tornar as informações públicas e acessíveis a qualquer um que tenha interesse nelas. Enquanto muitos sítios da Web irão querer manter seus dados proprietários, muitos outros irão optar por tornar suas informações públicas e abertas, usando marcações semânticas nos conteúdos disponibilizados".

Desta forma, continua a autora,

> "a Web Semântica oferece novas capacidades, tornando possível a adição de documentos que representam o 'conhecimento' sobre uma página, foto ou base de dados, público e de acesso livre, em um formato capaz de ser interpretado também pelas máquinas".

Segundo BOZSAK (2002),

> "a Web Semântica será capaz de apoiar serviços eletrônicos automatizados que usam as descrições semânticas armazenadas. Estas descrições são vistas como um fator chave para se achar o caminho para os crescentes problemas ocasionados pela expansão da Web".

O autor destaca que

> "a Web Semântica engloba esforços de se construir uma nova arquitetura Web que realce o conteúdo com uma semântica formal. Isto irá permitir que agentes automatizados 'raciocinem' sobre o conteúdo das páginas e possam realizar tarefas 'mais inteligentes' em nome dos usuários. 'Expressar significados' é a principal tarefa da Web Semântica no momento".

BERNERS-LEE (2001) afirma que, "para que a Web Semântica funcione, os computadores devem acessar coleções estruturadas de informação e conjuntos de regras de inferência que poderão conduzi-los a um 'raciocínio automatizado'". Portanto, nos dias atuais, o desafio da Web Semântica é "prover uma linguagem que expresse tanto dados quanto regras de raciocínio sobre os dados e que permitam regras de qualquer sistema de representação do conhecimento existente para ser exportado para a Web". Esta linguagem é a ontologia. Ainda segundo o autor, "as ontologias típicas para a Web têm uma taxonomia e um conjunto de regras de inferência". A taxonomia define classes de objetos e relações entre eles. Pode-se expressar um grande número de relações entre entidades assinalando propriedades para as classes e permitindo que subclasses herdem suas propriedades. Regras de inferência fornecem poder adicional às ontologias.

Segundo o autor,

> "o computador não 'entende' uma informação, mas com o uso de regras de inferência, ele pode manipular os termos com muito mais eficiência de forma a tornar mais usável e inteligível para o usuário humano".

Os portais podem se favorecer do uso de ferramentas baseadas em ontologias e da marcação semântica das páginas da Web de várias maneiras. BERNERS-LEE (2001) afirma que as ontologias podem aumentar a funcionalidade da Web das seguintes formas:

- Podem ser usadas de uma maneira simples para melhorar a eficácia de pesquisas na Web. O programa pesquisador olha hoje apenas para as páginas que contêm o termo preciso, ao invés de olhar para todas aquelas que usam termos equivalentes;
- Páginas podem ser direcionadas para outras páginas, cuja ontologia define como informações relacionadas. As informações são processadas e podem ser usadas para responder questões que normalmente exigiriam intervenção de um humano;

- Adicionalmente, as marcações com significados específicos tornam mais fácil o desenvolvimento de programas de software que podem responder a questões complexas cujas soluções não residem em uma página Web simples.

BERNERS-LEE (2001) apresenta uma arquitetura de cinco camadas para a Web Semântica :

Camada 1 : A camada de sintaxe (*syntax layer*)

Permite a marcação arbitrária de conteúdo através de significados de elementos atribuídos e aninhados. O nome desses elementos nada diz a respeito do que sua estrutura significa, portanto, algum significado a mais é necessário, ficando o papel do XML (*eXtensible Markup Language*) reduzido a uma posição puramente sintática. Por este motivo, é correto afirmar que a XML é usada para adicionar metadados às aplicações, mas não oferece meios para conectar ou inferir informações das sentenças. A linguagem XML será vista com mais detalhes na Seção 2.6.

Camada 2: A camada de dados (*data layer*)

RDF (*Resource Description Framework*) permite a codificação, troca e reutilização de metadados estruturados. Principalmente, informação é representada por um significado genérico, os pseudográficos (também chamados triplas), escritos em XML. Contrariamente à XML, RDF permite assinalar identificadores globais aos recursos e permite referenciar e estender sentenças feitas em outros documentos. A RDF será visto com mais detalhes na Seção 2.7.

Camada 3: A camada de ontologias (*ontology layer*)

O terceiro componente básico da Web Semântica é a ontologia. Por definir vocabulários e teorias de domínio comuns e compartilhadas, as ontologias ajudam tanto pessoas quanto máquinas a se comunicarem de forma mais concisa, apoiando-as de maneira semântica e não apenas em relação à sintaxe. A linguagem de criação de ontologias OWL (*Ontology Web Language*) será vista com mais detalhes na Seção 2.10.

Camada 4: A camada lógica (*logic layer*)

As ontologias são expressas em linguagens lógicas. Para codificar relacionamentos em uma linguagem lógica, várias regras de inferência devem ser elaboradas, para que sejam permitidos escolher cursos de ação e responder questões. Esta camada consiste exatamente nestas regras que permitem inferências. Mecanismos de inferência[11] são ferramentas que usam sentenças lógicas para inferir sobre classes, propriedades, instâncias e seus relacionamentos.

Camada 5: A camada de inferência (*proof layer*)

Permite o acompanhamento de respostas geradas por agentes automatizados. Isto irá requerer a tradução de seus mecanismos de raciocínio internos em alguma linguagem unificada de realização de inferências.

Na Figura 2.1 pode-se ver a arquitetura de cinco camadas para a Web Semântica. Os padrões e linguagens apresentados anteriormente serão detalhados nas próximas seções.

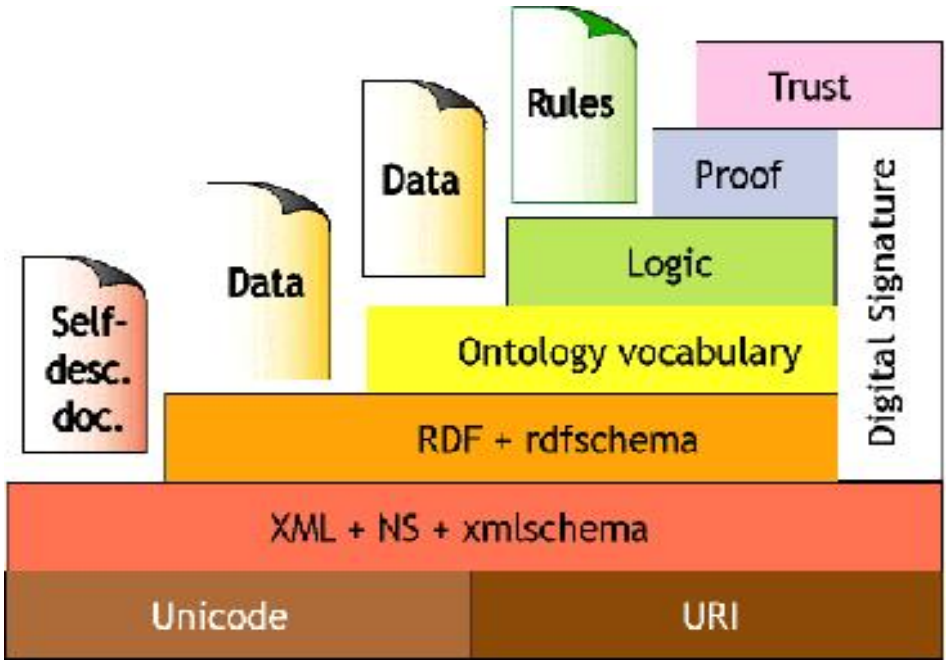


Figura 2.1 - Arquitetura de cinco camadas proposta para a Web Semântica por BERNERS-LEE (2001)

[11] do termo inglês *Reasoners*

2.6) XML – eXtensible Markup Language

Segundo CHAVES et al (2001),

> "a maior parte dos documentos disponíveis na Web hoje está identificada apenas com informações de exibição. Diante da necessidade de se prover meios mais adequados para possibilitar acesso – mediado por máquina – ao repositório de informações disponível na Web, é necessária a utilização da tecnologia XML[12], que possibilita a exploração da semântica dos conteúdos desses documentos".

Torna-se necessário, neste momento, uma distinção entre as linguagens XML e HTML (*HiperText Markup Language*). A XML é utilizada para estruturar dados, enquanto a HTML é usado basicamente para a apresentação de conteúdo. Neste contexto, é importante ficar explícito que a XML não é substituta à HTML. A XML é uma maneira de solucionar limitações da HTML.

De acordo com o W3C,

> "XML é uma linguagem simples, de formato texto muito flexível, derivada do SGML. Originalmente projetada para enfrentar os desafios de publicações eletrônicas de larga escala, está desempenhando um papel cada vez mais importante nas trocas de uma grande variedade de dados na Web".

Segundo PITTS-MOULTIS e KIRK (2000), XML "é a base para o processamento de metadados e permite a exploração de consultas através destes metadados". Um catálogo de biblioteca é um bom exemplo de uso de metadados, porque nos permite obter dados sobre os livros contidos nela. Eles nos dão informações sobre quem escreveu, quando foi publicado, que assunto é discutido, etc.

A linguagem XML é uma das tecnologias mais amplamente adotadas para intercâmbio e representação de informações na Web e permite, a cada um, criar suas próprias tags, rótulos ocultos que anotam as páginas Web, ou seções de texto em uma página. XML permite aos usuários adicionar arbitrariamente estrutura a seus documentos, entretanto, não diz nada sobre o que a estrutura significa. Significado é expresso pela RDF, que trabalha com conjuntos de triplas, que podem ser escritas usando tags XML. A RDF é detalhada na seção a seguir.

[12] http://www.w3c.org/XML

2.7) RDF - Resource Description Framework

A RDF[13] é um padrão recomendado pelo W3C que usa notação XML como sintaxe de codificação para descrição de metadados. PITTS-MOULTIS e KIRK (2000) afirmam que "um dos alicerces importantes da RDF é sua habilidade de automatizar o processamento de recursos Web". Segundo os autores,

> "RDF é uma aplicação XML que serve como uma base para o processamento de metadados. Seu principal objetivo é o de facilitar o intercâmbio de informações, que podem ser interpretadas por máquinas, entre aplicativos via Web".

LASSILA e SWICK (1999) afirmam que "a RDF é o padrão para representação de metadados na Web. Essa linguagem permite aos computadores representar e compartilhar dados semânticos na Web".

Na Web que conhecemos hoje, os mecanismos de pesquisa indexam páginas por palavras-chave existentes nos textos dos documentos e pelos *links* existentes na página, não havendo formas de uma máquina compreender algo sobre a página. Na Web Semântica, a marcação é feita tomando como base os recursos existentes na página. Usando RDF e OWL (a ser detalhada na Seção 2.10), os usuários podem codificar recursos existentes em uma página Web e apontar para outros recursos existentes em outros sítios. Desta forma, as triplas RDF, como são chamadas as unidades básicas de informação, formam teias de informação sobre coisas relacionadas.

Um documento ou elemento RDF faz declarações sobre recursos. Uma declaração diz que um certo recurso tem uma ou mais propriedades. Cada propriedade tem um tipo, que é um nome, e um valor. O valor de uma propriedade pode ser um literal, como um texto, um número, ou pode ser outro recurso. Uma declaração é composta de três partes:

- *Recurso* (ou Sujeito) – Qualquer objeto que é identificável unicamente. Podem ser livros, elementos em uma página e pessoas individuais, por exemplo;
- *Propriedade* (ou Verbo) – É uma característica, um atributo ou um relacionamento específico de um recurso;

[13] http://www.w3c.org/RDF

- *Valor* (ou Objeto) - Deve ser atômico na natureza (literal de texto ou números, por exemplo) ou outros recursos, os quais podem ter suas próprias propriedades.

É apresentado, na Tabela 2.2, um exemplo de uma tripla RDF, dentro do contexto do ensino superior, utilizando-se da ontologia ONTO-POINT e gravada no formato OWL.

Recurso	Http://www.unipac.br/onto-point.owl#PauloLima
Propriedade	Http://www.unipac.br/ onto-point.owl#coordena curso
Valor	"Sistemas de Informação"

Tabela 2.2 – Exemplo de tripla RDF a partir da ontologia ONTO-POINT

Na Web Semântica, a todo recurso é dado um nome único, chamado Identificador Uniforme de Recurso, ou URI [14], que é o endereço Web do documento que contém a marcação, seguido de "#", e depois do nome do objeto. Uma vez que cada recurso tem um nome único, ele poderá ser referenciado em qualquer lugar, por qualquer pessoa, que tenha o tema como interesse. Também torna-se possível fazer referência a instâncias definidas em outras ontologias.

Na Tabela 2.3, é apresentado outro exemplo, onde o valor não é um literal fixo, como no exemplo acima, mas um objeto, que se utiliza de recursos de outra ontologia :

Recurso	http://www.unipac.br/onto-point.owl#UNIPAC
Propriedade	http://www.unipac.br/ onto-point.owl#pertence
Valor	http://www.ConselhosEstaduaisDeEducacao.owl#CEE-MG

Tabela 2.3 – Exemplo de tripla RDF utilizando recursos de outra ontologia

Sujeito e objeto são identificados por URIs, exatamente como uma página Web, que usa o tradicional Localizador Uniforme de Recurso ou URL [15], o tipo mais comum de URI. Os verbos também são identificados por URIs, o que capacita a qualquer pessoa definir um novo conceito ou um novo verbo, apenas definindo a sua URI em algum lugar na Web.

[14] do termo inglês *Uniform Resource Identifier*

O benefício desta marcação é similar aos *links* usados nos documentos HTML, que remetem o usuário a outras páginas para que ele possa obter mais informações. A diferença é que a Web Semântica oferece mais do que isso, pois, fazendo referência a um ou mais recursos da página, temos mais informações e informações mais detalhadas a respeito do assunto de nosso interesse, que podem, indubitavelmente, ser processadas por agentes não humanos, se necessário. Além disso, a linguagem humana se modifica quando usamos o mesmo termo para significar coisas diferentes, mas uma linguagem automática não tem esta capacidade. Deve-se usar URIs diferentes para solucionar esse problema.

A Web Semântica, nomeando todos os conceitos simplesmente através de uma URI, permite que alguém expresse novos conceitos por ele criados. A unificação de uma linguagem lógica permitirá que esses conceitos sejam progressivamente ligados em uma teia universal. BERNERS-LEE (2001) afirma que

> "essa estrutura vai abrir o conhecimento e os trabalhos da humanidade para uma análise semântica por agentes de software, provendo uma nova classe de ferramentas através das quais todos poderão viver, trabalhar e aprender juntos".

2.8) RDFS - Resource Description Framework Schema

A RDF provê uma maneira de se criar instâncias e associar propriedades descritivas a cada uma delas. Entretanto, a RDF não provê sintaxe para se definir classes, propriedades e descrever como elas se relacionam umas com as outras. Para realizar esta tarefa, utiliza-se a RDF Schema (RDFS). RDFS [16] usa o RDF como base para especificar um conjunto de recursos RDF pré-definidos e propriedades que permitem então ao usuário definir classes, suas sub-classes, propriedades, domínios e restrições.

BRICKLEY e GUHA (2000) relatam que "RDF em si é uma linguagem simples que é capaz de fazer relacionamentos entre informações, mas, além disso, é necessário um meio para definição de dados. (...) RDF Schema foi criada com essa finalidade".

---

[15] do termo inglês *Uniform Resource Locator*
[16] http://www.w3c.org/RDF/#schemas

O sítio do W3C define RDF como "uma linguagem de propósito geral para representar informações na Web" e RDFS como "um padrão que descreve como utilizar RDF para descrever vocabulários RDF na Web".

Grande parte das ferramentas utilizadas para o desenvolvimento de ontologias possibilita a exportação dos dados para o formato RDFS. BRAGANHOLO e HEUSER (2001) afirmam que "RDF Schema é um sistema de classes extensível e genérico que pode ser utilizado como base para esquemas de um domínio específico". Esses esquemas podem ser compartilhados e estendidos através de refinamento de subclasses. Além disso, segundo os autores, "definições de metadados podem ser reutilizadas através do compartilhamento de esquemas".

2.9) RDQL - Resource Description Framework Data Query Language

RQL (*RDF Query Language*) foi a primeira linguagem de consulta declarativa para metadados RDF. A linguagem conta com um modelo de grafos formal, que captura a modelagem em RDF e RDF Schema e permite a interpretação de descrições sobrepostas por meio de um ou mais esquemas. A linguagem RDQL (*RDF Data Query Language*) vem se destacando como uma evolução de RQL e tem sido implementada em um grande número de sistemas RDF com o objetivo de extrair informação de modelos em RDF. A linguagem é encontrada hoje nos principais sistemas de armazenamento de triplas RDF.

Segundo SEABORNE (2003),

> "RDQL consiste em um padrão gráfico, representado como uma lista de padrões de triplas. Cada padrão de triplas é constituído de variáveis com nomes e valores RDF (URIs e literais). (...) Uma consulta RDQL pode adicionalmente ter um conjunto de restrições de valores destas variáveis e uma lista de variáveis requeridas no conjunto de respostas".

A Figura 2.2 mostra a sintaxe padrão de consulta em RDQL, ressaltando as similaridades com a linguagem SQL [17] encontrada nos SGBDs [18] relacionais.

[17] *Structured Query Language*
[18] Sistemas Gerenciadores de Bancos de Dados

```
SELECT variables listing
FROM rdf documents
WHERE patterns
AND filter expressions
USING prefix declaration
```

Figura 2.2 – Exemplo de consulta RDQL, com sintaxe semelhante ao SQL

2.10) OWL – Ontology Web Language

Em março de 2001, uma linguagem chamada DAML+OIL foi elaborada como resultado do trabalho conjunto de um comitê de pesquisadores dos EUA e Europa. Em novembro de 2001, o W3C criou o Web Ontology Working Group para desenvolver uma linguagem para representar ontologias. A linguagem resultante foi chamada de OWL [19] (*Ontology Web Language*) e se baseia em RDF e RDFS, que, por sua vez, utilizam a sintaxe de XML.

OWL é um vocabulário que estende a RDFS, definindo classes e relacionamentos de maneira mais detalhada. Desde que OWL é construído em RDF, um grafo qualquer RDF forma uma ontologia OWL válida. Entretanto, OWL adiciona semântica e vocabulário a RDF e RDFS, dando-lhe mais poder para expressar relacionamentos complexos.

De acordo com o W3C, "a linguagem OWL é projetada para ser usada por aplicações que necessitam de processar o conteúdo das informações, ao invés de apenas apresentar informações aos seres humanos". OWL facilita um maior interpretação do conteúdo Web por máquinas do que XML, RDF e RDFS, por prover vocabulário adicional juntamente com a semântica formal. OWL possui três sub-linguagens que aumentam gradativamente o nível de expressividade em ordem crescente: OWL Lite, OWL DL e OWL Full. Ainda segundo o W3C, OWL pode ser usada para:

1. Formalizar um domínio pela definição de classes e propriedades destas classes;
2. Definir instâncias e propriedades sobre elas;
3. Raciocinar a respeito destas classes e instâncias até o grau permitido pela semântica formal da linguagem.

[19] http://www.w3c.org/TR/owl-features

Uma ontologia OWL consiste em classes, propriedades e instâncias. Segundo HORRIDGE et al (2004), instâncias representam os objetos do domínio nos quais temos interesse, podendo também serem chamadas de instâncias de classes, enquanto classes são interpretadas como conjuntos que contém instâncias. Segundo o autor, "a palavra conceito é ,algumas vezes, usada no lugar de classe. (...) Classes são, na verdade, uma representação concreta de conceitos". Ele define propriedades como relações binárias entre indivíduos, ou seja, propriedades ligam dois indivíduos, um ao outro. Também são chamadas de papéis na lógica descritiva e relações na UML[20] e em outros ambientes orientados a objetos. Usualmente também podem ser chamadas de atributos.

Segundo HORRIDGE et al (2004), em uma primeira classificação, as propriedades OWL podem ser de três tipos, a saber:

- Propriedades de Objetos [21]: ligam uma instância a outra instância;
- Propriedades de Tipos de Dados [22]: ligam uma instância a um valor ou a um literal;
- Propriedades de Anotação [23]: podem ser usadas para adicionar informações, ou metadados, a classes, instâncias ou propriedades de objetos ou de tipos de dados.

Em uma outra classificação, o autor diferencia outros tipos de propriedades:

- Propriedades Inversas: cada Propriedade de Objetos possui uma Propriedade Inversa correspondente. Por exemplo, a Figura 2.3 mostra que, se João é pai de José, então, tem-se, por inferência, que José é filho de João.

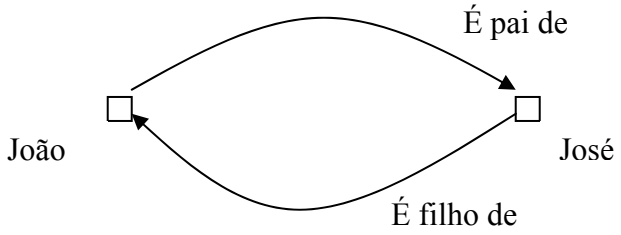


Figura 2.3 – Exemplo de Propriedade Inversa

[20] do termo inglês *Unified Modelling Language* ou Linguagem de Modelagem Unificada
[21] do termo inglês *Object Properties*
[22] do termo inglês *Datatype Properties*
[23] do termo inglês *Annotation Properties*

- Propriedades Funcionais: se uma propriedade é funcional para um indivíduo, deve existir no máximo um outro indivíduo relacionado a ele através desta propriedade. Por exemplo, a Figura 2.4 mostra que, se Pedro tem como avó materna Maria Helena e a propriedade é funcional, e se Pedro tem como avó materna "Vó Maria", então Maria Helena e "Vó Maria" são, necessariamente, a mesma pessoa

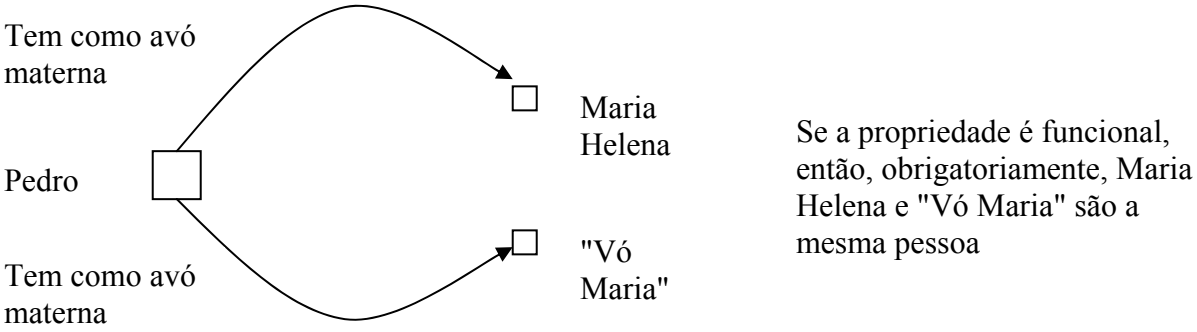


Figura 2.4 – Exemplo de Propriedade Funcional

- Propriedade Funcional Inversa: uma propriedade é funcional inversa quando sua propriedade inversa é funcional, ou seja, para um certo indivíduo, deve existir no máximo um outro indivíduo relacionado a ele através desta propriedade. Por exemplo, a Figura 2.5 mostra que, se Maria Helena é avó materna de Pedro e a propriedade é funcional inversa, e se "Vó Maria" é avó materna de Pedro, então, novamente, pode-se inferir que Maria Helena e "Vó Maria" são a mesma pessoa.

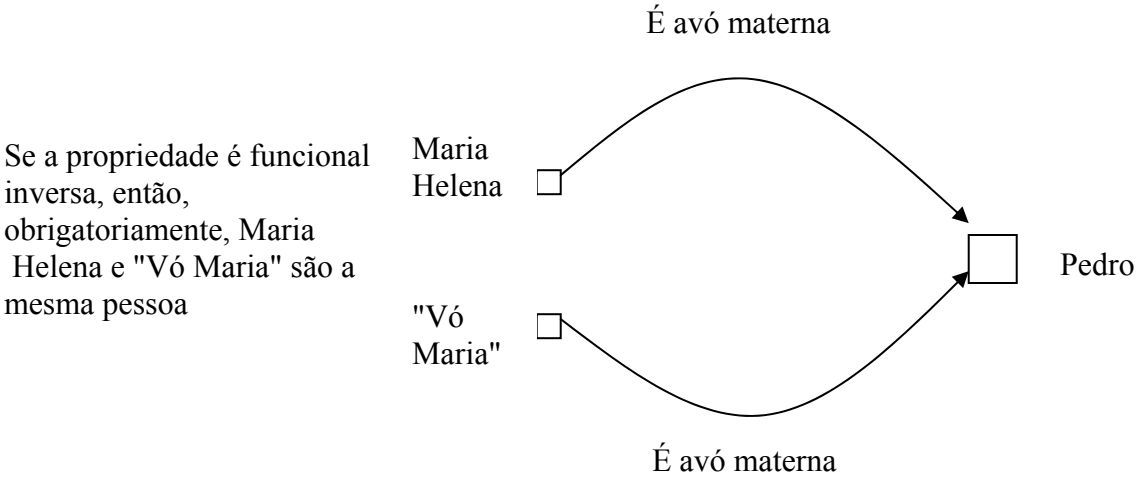


Figura 2.5 – Exemplo de Propriedade Funcional Inversa

- Propriedades Transitivas: se a propriedade transitiva relaciona o indivíduo A e o indivíduo B e também o indivíduo B ao indivíduo C, então pode-se inferir que o indivíduo A se relaciona com o indivíduo C via propriedade transitiva. Por exemplo, a Figura 2.6 mostra que, se Mateus é descendente de Tiago, que é descendente de William, então Mateus é descendente de William.

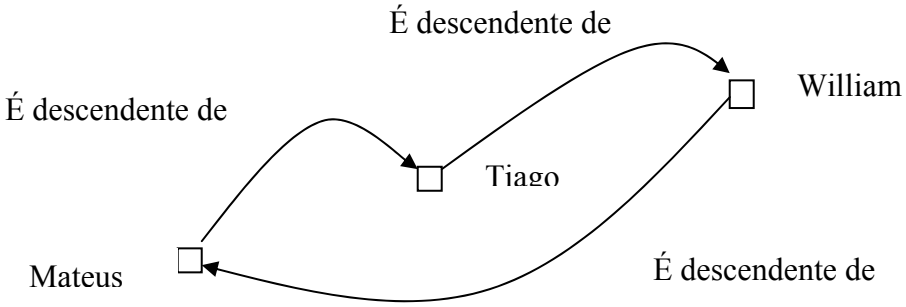


Figura 2.6 – Exemplo de Propriedade Transitiva

- Propriedades Simétricas: se a propriedade simétrica relaciona o indivíduo A ao indivíduo B, então o indivíduo B também se relaciona com o indivíduo A via propriedade simétrica. Por exemplo, a Figura 2.7 mostra que, se Carla é prima de Adriana, então Adriana é prima de Carla.

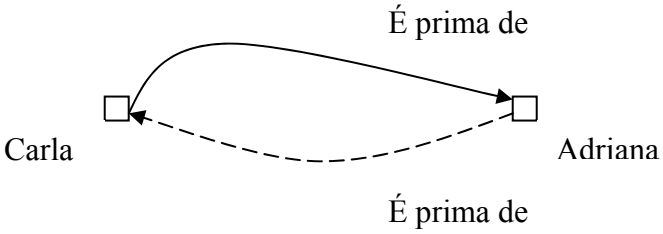


Figura 2.7 – Exemplo de Propriedade Simétrica

CAPÍTULO 3 - Ferramentas semânticas

O objetivo deste capítulo é fazer uma investigação de ferramentas semânticas existentes na literatura científica contemporânea da área que utilizam as tecnologias mencionadas no capítulo anterior e que agregam valor aos portais Web, oferecendo funcionalidades de interesse desta pesquisa.

A meta não é analisar o maior número possível de ferramentas, mas dar destaque àquelas que se mostraram mais adequadas à proposta desse estudo e que podem, por serem de código aberto, integrar o portal em sua futura implementação.

As ferramentas foram divididas em seis grupos, de acordo com seu foco de atuação funcional. Algumas das ferramentas apresentadas nesse capítulo serão utilizadas, na prática, para ilustrar, através de exemplos no Capítulo 6, as funções do portal semântico proposto.

O primeiro grupo diz respeito a ferramentas que auxiliam na criação e manutenção de ontologias, sendo apresentada apenas a ferramenta Protégé 2000, que se destaca bastante no cenário atual, sendo escolhida para a criação da ontologia utilizada no trabalho. O segundo grupo abrange ferramentas de autoria e marcação semântica, sendo descritos os aplicativos SMORE e Semantic Word. O terceiro grupo apresenta ferramentas de navegação, representadas por MAGPIE e Espaço Semântico (este último, utilizado para navegação guiada por taxonomias). O quarto grupo engloba ferramentas de busca semântica, sendo apresentadas as ferramentas KIM e TOSS. Um quinto grupo foi criado para englobar ferramentas de gestão de conteúdo [24]. Apesar de não serem propriamente ferramentas semânticas, serão apresentadas aqui por serem utilizadas como arcabouço de código aberto para a construção de portais, devendo servir de base para a integração das outras ferramentas na implementação do portal proposto nesta pesquisa. Portanto, são apresentados nesse grupo o Plone e o Zemantic. Finalmente, no sexto grupo, é apresentado um estudo sobre a identificação de assuntos em mensagens de *chat*, considerando-se sua utilidade como ferramenta para colaboração entre usuários do portal que será proposto no Capítulo 5 deste trabalho.

[24] O termo "gestão de conteúdo" ou "sistema de gestão de conteúdo" do inglês *Content Management System* (CMS) surgiu na indústria de software para designar o conjunto de tecnologias de gerenciamento de sítios Web. Não deve ser interpretado como sinônimo de "gestão de informação", conceito mais antigo e genérico.

3.1) Ferramentas para criação e manutenção de ontologias

- **Protégé 2000 com Plug-In OWL**

Segundo KNUBLAUCH e MUSEN (2004), "Protégé é um ambiente para desenvolvimento de ontologias de código aberto com funcionalidades para editar classes, propriedades e instâncias". É uma plataforma para modelagem do conhecimento desenvolvida pela Universidade de Stanford [25] que apóia uma comunidade de usuários há quase duas décadas.

A ferramenta, ao longo de todo esse tempo, passou por inúmeras evoluções, sendo que a mais recente lhe deu a capacidade de trabalhar com ontologias OWL através da extensão OWL [26]. Ela permite carregar e salvar arquivos OWL em vários formatos e realizar classificações inteligentes baseadas em descrição lógica[27] (DL). DL é a base da tecnologia da Web Semântica, assim como a OWL. Máquinas de inferência DL ajudam a construir e manter ontologias compartilháveis, descobrindo inconsistências, dependências escondidas e possíveis redundâncias.

A extensão OWL permite acesso direto a mecanismos de inferência (como o aplicativo *Racer*, por exemplo). Na versão existente hoje, são permitidos dois tipos de inferência em DL: checagem de consistência e classificação (ou subsunção). Outros tipos de inferência, como checagem de instâncias, ainda estão sendo construídas e devem ser disponibilizadas nas próximas versões do produto.

A checagem de consistência, ou seja, o teste se uma classe poderia ter instâncias, pode ser chamada por todas as classes ou para apenas as classes selecionadas. As classes inconsistentes, caso existam, são marcadas com um ícone de cor vermelha.

A classificação, ou seja, inferência de uma nova árvore a partir das definições existentes, pode ser chamada a partir de um comando especial. É feito um teste completo de consistência e as classes inconsistentes não são classificadas corretamente. Uma classificação inferida é mostrada ao lado da classificação original definida. Uma lista completa de mudanças sugeridas também é mostrada ao usuário, facilitando a análise do usuário para que ele consiga criar e manter com sucesso suas ontologias.

[25] http://protege.stanford.edu
[26] do termo inglês OWL *Plug-in*
[27] do termo inglês *Description Logic*

3.2) Ferramentas de autoria e marcação semântica

o SMORE - Semantic Markup, Ontology and RDF Editor

A ferramenta SMORE [28] foi projetada a partir da idéia de que o compartilhamento de conhecimento na Web Semântica deve se parecer mais com a criação de páginas Web do que uma complexa engenharia do conhecimento, reduzindo, desta forma, a linha que separa a criação normal de um conteúdo e a sua anotação semântica.

A ferramenta faz parte do projeto Mindswap [29] e é uma ferramenta que permite ao usuário realizar a marcação semântica de documentos em RDF através do uso de ontologias. A ferramenta também é útil para criação de ontologias, podendo também se referenciar a outras ontologias existentes na Internet.

SMORE integra a criação de conteúdo e sua marcação. De um lado, apresenta um editor texto/HTML que permite aos usuários criar e disponibilizar páginas Web. De outro lado, um editor facilita a geração de marcação semântica, em que o autor pode criar triplas com significado semântico de seus dados. O usuário pode selecionar porções do texto da página Web e arrastá-lo até as classes da ontologia, inserindo-as nas triplas que seguem o padrão do modelo sujeito-predicado-objeto.

Segundo o sítio do projeto,

> "SMORE provê ao usuário um ambiente no qual ele pode criar páginas sem muitos impedimentos advindos da marcação semântica, pois ele marca seus documentos com um mínimo de conhecimento de termos e sintaxe RDF. (...) Entretanto, ele deve ser capaz de classificar semanticamente seus dados para anotação, em um modelo básico de sujeito-predicado-objeto".

O processo de criação de metadados se aplica também a recursos não-textuais, como fotografias, por exemplo. SMORE permite aos usuários adicionar triplas em um documento para descrever uma fotografia. A ferramenta também permite a anotação de informações detalhadas de sub-imagens, ou uma região delimitada da imagem.

---

[28] http://www.mindswap.org/~aditkal/editor.shtml

[29] *Maryland Information and Network Dynamics Lab Semantic Web Agents Project*

o SEMANTIC WORD

Semantic Word foi desenvolvido levando em consideração as premissas de que o processo de produção de documentos consome a maior parte do tempo dos autores e eles não têm tempo suficiente para elaborar anotações semânticas nos mesmos. É uma ferramenta de anotação semântica que se utiliza do ambiente do aplicativo Microsoft Word para disponibilizar ferramentas customizáveis que permitem, simultaneamente, a geração de conteúdo e a realização de marcações semânticas nos documentos.

Desenvolvida pela empresa TEKNOWLEDGE [30], tem como características principais um esquema de anotações que permite o reuso das mesmas quando o conteúdo é reusado, uma biblioteca customizável de moldes contendo textos parcialmente anotados e um sistema de extração automática de informações.

Semantic Word estende o MS Word. Primeiramente, uma nova barra de ferramentas é acrescentada para dar apoio à criação de descrições semânticas. A interface com o usuário também é estendida para mostrar estas anotações feitas junto com o texto e permitir que elas sejam manipuladas. Em segundo lugar, a ferramenta estende o editor de textos através da abertura de um canal com a Web Semântica. O conteúdo da Web Semântica, como definições da ontologia, por exemplo, é trazido para o ambiente de forma a compor as anotações que, mais tarde, são levadas de volta à Web Semântica. Em terceiro lugar, a ferramenta estende os serviços do MS Word de forma a integrar nele um aplicativo responsável pela extração automática de informações denominado AeroDAML. Este aplicativo analisa e anota o texto do documento no momento em que ele é produzido, aparecendo para o autor como um serviço análogo ao '*Word spelling'* ou a checagem gramatical existentes no editor de textos.

Existem dois tipos de anotações: referências de instâncias, que associam uma região do texto com instâncias de classes referenciáveis, e as chamadas '*triple bags*', que descrevem o conteúdo da região do texto como uma coleção de triplas que seguem o modelo RDF sujeito-predicado-objeto, onde o sujeito é uma instância, o predicado é uma propriedade da ontologia e o objeto pode ser tanto uma instância quanto um valor.

[30] http://mr.teknowledge.com/daml/

3.3) Ferramentas de navegação semântica

- **MAGPIE**

MagPie é uma ferramenta desenvolvida pelo *Knowledge Media Institute* (KMI) [31], da *Open Unversity*, Milton Keynes, Inglaterra e tem por objetivo auxiliar o usuário a acessar a Web. Foi criada partindo da consideração de que a atividade de navegação na Web envolve duas etapas: encontrar a página certa e fazer sentido do seu conteúdo.

DZBOR et al (2003) afirmam que

> "grande parte dos esforços atuais focam a busca de páginas Web, sendo dada pouca atenção ao outro aspecto da interpretação das páginas, como por exemplo, achar páginas visitadas anteriormente [32] e muito menos atenção é dada ao processo de construção de sentido [33]".

Segundo os autores,

> "a anotação semântica permite ao usuário associar metadados aos recursos Web, os quais podem facilitar a interpretação. (...) A ferramenta é então acrescentada a um navegador padrão e apoia a interpretação de páginas e gerenciamento da história de navegação semanticamente enriquecida, através da adição de uma camada baseada em ontologias, permitindo que serviços relevantes possam ser chamados a partir do navegador".

Desta forma, o usuário Web passa a ter acesso não só às informações registradas pelos autores da página, mas a um montante de informações de fundo[34] e conhecimento armazenado. Segundo DOMINGUE et al (2004) "é justamente esta informação de fundo que, na maioria das vezes, o usuário está procurando e que leva um longo tempo para ser encontrada e processada".

Os autores também afirmam que

> "diferentes usuários têm diferentes níveis de familiaridade com uma informação em uma página Web e com o domínio do qual a informação faz parte. (...) Desta forma, eles necessitam de níveis diferentes de apoio à construção de sentido e a camada semântica existente em MagPie foi projetada tendo isto em mente".

[31] http://kmi.open.ac.uk
[32] do termo em inglês *browsing history*
[33] do termo em inglês *sense making*
[34] do termo inglês *background information*

Baseado em estatísticas que afirmam que mais da metade das páginas Web visitadas pelas pessoas são revisitadas, o autor ainda conclui que os usuários necessitam de apoio para capturar aquilo que já foi visto anteriormente. MagPie permite marcar itens de informação encontrados em uma seção de navegação em uma lista semântica, para serem posteriormente recuperadas.

Para se usar Magpie é necessário instalá-lo como uma extensão de um navegador Web. Uma barra de ferramentas especial é acrescentada às demais existentes no navegador. A partir daí, deve-se carregar a ontologia de domínio que dará suporte ao processo de navegação semântica, através de opção existente em um menu na barra de ferramentas. Automaticamente, as classes da ontologia são expostas nesta barra em forma de botões de comando com cores diferentes. Ao clicar cada um desses botões, os conceitos existentes na página visualizada são coloridos exatamente com a cor que representa a classe a que ela pertence na ontologia de domínio carregada anteriormente.

Para ter flexibilidade, o aplicativo foi projetado em forma de módulos. Apenas uma pequena parte reside localmente no navegador do usuário. A barra de ferramentas é carregada no navegador e a ontologia também pode residir no computador local, mas grande parte das funcionalidades é provida remotamente, através do acesso a serviços.

Assim, depois que as palavras ou frases estão marcadas semanticamente e coloridas de acordo com os conceitos que representam dentro da ontologia, o botão direito do mouse pode ser acionado para requisitar um menu dinâmico usado para invocar serviços Web, que podem processar cada conceito particularmente. Um gerenciador de serviços [35] é necessário para traduzir estas requisições em uma linguagem interna.

- **Espaço Semântico**

O aplicativo "Espaço Semântico" foi desenvolvido durante o trabalho de Iniciação Científica do aluno Leonardo Souza (SOUZA, 2005), do curso de graduação em Ciência da Computação da UFMG e tem como proposta ser um sistema que relaciona semanticamente o conteúdo de um portal de comunidades de práticas, permitindo a busca por conteúdos a

[35] do termo inglês *service handler*

partir da navegação em uma taxonomia. Para tal, essa aplicação deve permitir associar os termos da taxonomia aos conteúdos do portal.

O objetivo do aplicativo é permitir o relacionamento entre os documentos do portal, que deverão ser previamente classificados pelos usuários em uma taxonomia ou uma ontologia simples. No momento de se visualizar um documento, o sistema exibe os documentos correlatos.

A ferramenta foi desenvolvida como um produto do sistema Plone e suas principais funcionalidades são associar termos da taxonomia a documentos, navegar nos documentos a partir da taxonomia e atualizar a taxonomia.

Assim é possível associar um conjunto de palavras chaves (termos da taxonomia) a um documento. Para tal, basta visualizar a página de propriedades do documento e clicar no termo desejado da taxonomia exibida.

Uma pequena janela que aparece na esquerda da tela exibe a taxonomia em um formato hierárquico. É possível então visualizar sub-termos de um termo clicando no sinal de "+" ao lado de cada um deles. Para fazer a busca por um conteúdo associado a um termo, basta clicar sobre o termo escolhido e uma página com os documentos associados é exibida.

3.4) Ferramentas de busca semântica

o TOSS – motor de pesquisa semântica

KAON [36] é um projeto da Universidade de Karlshure, Alemanha, e consiste em uma infra-estrutura de código aberto, baseado em ontologias, para gerenciamento de aplicações de negócios. Faz parte do KAON um conjunto de ferramentas que facilita a criação e manutenção de ontologias, além de uma arquitetura para a construção de aplicações baseadas em ontologias e nos padrões da Web Semântica.

O portal PASS, que faz parte do arcabouço KAON, é um exemplo de portal semântico que faz uso de uma ferramenta semântica de busca. O portal semântico PASS é composto de dois módulos: ONTOPASS (*Ontologies for the PASS*), um editor de

[36] *The **Ka**rlshure **On**tology and Semantic Web Tool Suíte.*

ontologias, e TOSS (*Tools for Semantic Search*), um motor de pesquisa semântica, que possui uma estratégia de aumentar o poder de busca no contexto do portal.

O primeiro permite ao usuário a criação e a manutenção de suas ontologias através de um navegador Web, enquanto o segundo utiliza os conceitos das ontologias armazenadas no portal para orientar e favorecer as pesquisas dos usuários.

O principal objetivo de TOSS é ajudar o usuário a contextualizar a sua busca. Ele utiliza ontologias de domínio para expandir os termos de pesquisa com o objetivo de alcançar resultados mais relevantes. Esta expansão de termos é realizada transformando a pesquisa original e usando conceitos da ontologia de domínio. Relacionamentos hierárquicos como generalização-especialização, todo-parte, além de sinônimos e propriedades do tipo simétricas, transitivas e inversas, são utilizados para este fim. A ferramenta utiliza o mecanismo Google[37], através de sua *Application Programming Interface* (API) [38] para realizar efetivamente a pesquisa.

Avaliações provaram que uma maior relevância dos resultados das pesquisas foi alcançado, utilizando-se os conceitos de revocação e precisão. Os resultados da pesquisa através do TOSS foram comparadas às pesquisas realizadas através de motores tradicionais de busca, como o próprio Google, por exemplo, demonstrando resultados significativos.

o KIM – Knowledge & Information Management

A ferramenta KIM [39], criada pelo *Knowledge and Language Engineering Lab*, em Sirma, Bulgária, é uma plataforma de software que fornece serviços para anotação semântica semi-automática de texto, povoamento automático de ontologias, indexação semântica e recuperação de informações semi e não estruturadas, além de prover possibilidades de consultas e navegação em bases de conhecimento formal. No presente estudo, preferência foi dada à característica de recuperação de informações marcadas semanticamente e armazenadas em um repositório de metadados.

[37] http://www.google.com
[38] Conjunto de rotinas e padrões estabelecidos por um software para o uso de suas funcionalidades. Composta por funções acessíveis somente por programação, e que permitem utilizar características do software menos evidentes ao usuário tradicional.
[39] http://www.ontotext.com/kim/index.html

O pacote KIM inclui a chamada ontologia KIM (KIMO), a base de conhecimento KIM World KB, o servidor KIM Server com API para acesso remoto e integração, a tela KIM Web User Interface e uma extensão para o navegador Internet Explorer.

É através do KIM Web User Interface que a ferramenta oferece as melhores funcionalidades para recuperação de informações. São disponibilizados três níveis de complexidade para consultas semânticas, detalhados a seguir.

O primeiro nível, denominado Pesquisa de Entidades[40], permite definições de restrições de consulta pelo tipo ou *alias* da entidade. Por exemplo, 'mostre-me todas as organizações que terminam com Ltda'. O segundo nível, chamado Pesquisa de Padrões pré-definidos[41], provê um conjunto de consultas usadas para auxiliar o usuário. São usadas entidades com tipos específicos, 'Pessoa', por exemplo, sendo permitido ao usuário restringi-las pelo seu nome. 'Mostre-me as pessoas que tem CEO como cargo' seria um exemplo. O terceiro nível, chamado Pesquisa de Padrões de Entidade[42], provê a flexibilidade de especificar entidades e relações entre entidades, e assim, criar o padrão. Também neste modelo de consulta podem ser criadas restrições a partir dos atributos das entidades.

A partir da definição de cada consulta, a recuperação de entidades ou documentos é realizada e os resultados podem ser refinados através de alterações nas consultas. Os documentos recuperados podem ser analisados e aqueles provenientes de entidades mais relevantes são destacados e ligados à base de conhecimento e apresentados pelo aplicativo KB Explorer que é capaz de mostrar também as outras marcações semânticas existentes.

[40] do termo inglês *Entity Lookup*
[41] do termo inglês *Predefined Pattern Search*
[42] do termo inglês *Entity Pattern Search*

3.5) Ferramentas de gestão de conteúdo e repositório de metadados

- **PLONE / ZEMANTIC**

Plone[43] é um Sistema de Gestão de Conteúdo (CMS) [44] de código aberto, distribuído sobre a licença pública Geral GNU GPL[45] e construído em Zope[46], um servidor de aplicações. Ambos utilizam a linguagem Python[47].

Plone dá poder aos usuários finais para publicar, sozinhos, no sítio da Web da organização em que trabalham, o conteúdo por eles criados, sem a necessidade de intervenção de profissionais especializados em programação. Plone oferece uma conjunto de funcionalidades como gerenciamento de membros da comunidade, *workflow*, padrões de publicação, além de uma série de funções de segurança para restringir os privilégios de publicação dos usuários, de forma a permitir a eles executarem o seu trabalho sem interferirem de forma inadvertida no trabalho de outros colegas.

De acordo com PEREIRA e BAX (2002), "a idéia que está por detrás de um Sistema de Gestão de Conteúdo é a de separar o conteúdo, estrutura e projeto gráfico". A estrutura é definida em arquivos molde, o projeto é codificado em padrões de estilo, geralmente em arquivos do tipo CSS[48], enquanto o conteúdo é armazenado em bancos de dados ou arquivos separados. Quando um usuário solicita uma página, as partes são combinadas para produzirem uma página XHTML padrão.

Segundo o sítio do Plone,

> "o sistema é ideal como um servidor de intranet e extranet, como um sistema de publicação de documentos, servidor de portal e como uma ferramenta de trabalho em grupo para colaboração entre entidades separadas geograficamente".

A comunidade Plone conta hoje com uma larga equipe de desenvolvedores ao redor do mundo e muitas companhias especializadas em desenvolvimento e suporte. Existe interoperabilidade entre Plone e a maioria dos sistemas de bancos de dados relacionais de

---

[43] http://www.plone.org
[44] do termo inglês *Content Management System*
[45] do termo inglês *General Public License*
[46] http://www.zope.org
[47] http://www.python.org
[48] *Cascading Style Sheets* (padrão definido pelo W3C).

código livre e também comerciais, além de rodar em um grande número de sistemas operacionais e plataformas, incluindo Linux, Windows, Mac OS X, Solaris, entre outros.

Por sua vez, o Zope é um servidor de aplicações[49], infra-estrutura para o Plone, construído por um conjunto de componentes que se integram em torno de um servidor Web, um banco de dados e um interpretador da linguagem Python.

Zemantic é um catálogo baseado em Web Semântica e um motor de pesquisa para Zope versões 2 e 3. O sistema funciona catalogando dados RDF que podem ser provenientes de objetos internos ou recursos externos ao Zope. Para seu funcionamento, Zemantic requer uma biblioteca especial construída na linguagem Python.

3.6) Utilização de ontologias em ferramentas de colaboração

- **Identificação de Assuntos em Mensagens de Chat**

Um estudo sobre a identificação de assuntos em mensagens de *chat* [50] foi desenvolvido por um grupo de pesquisadores da Universidade Católica de Pelotas (UCPEL) – Grupo de Pesquisa em Sistemas de Informação, e tem como objetivo investigar o processo de classificação de textos sobre mensagens de um *chat* na Web. O sistema utiliza uma ontologia de domínio para classificar documentos, para identificar temas nas mensagens e para traçar o perfil dos usuários. Acredita-se que, a partir da identificação do assunto de uma conversa informal de *chat* por parte de uma ferramenta baseada em ontologias, torna-se possível que esta ferramenta possa sugerir aos usuários informações e conteúdo armazenado em um repositório de um portal semântico, que tenham relação com o assunto sobre o qual as pessoas estão conversando. Esta seria uma ferramenta de colaboração baseada em ontologia que poderia ser utilizada pelo portal semântico para incrementar a interação entre os membros da comunidade universitária.

As mensagens de *chat* possuem algumas particularidades como concisão, pouco cuidado com correções ortográficas e revisões, devido a serem escritas às pressas e por pessoas de diferentes níveis culturais.

[49] Plataforma de software para construção e gerenciamento de aplicações Web dinâmicas.
[50] *chat*, que em português significa conversação, é um neologismo para designar aplicações de conversação em tempo real. Esta definição inclui programas de IRC, ou mensageiros instantâneos.

Segundo LOHL (2004),

> "o método de identificação de assuntos funciona examinando cada uma das mensagens enviadas pelos usuários participantes do *chat*. Os assuntos ou temas são identificados pela comparação dos termos que aparecem nas mensagens com termos que estão relacionados a conceitos presentes numa ontologia. O método faz uma classificação de textos, isto é, identifica assuntos nos textos presentes nas mensagens".

O método está baseado em técnicas probabilísticas e, portanto, não utiliza técnicas de Processamento de Linguagem Natural para analisar a sintaxe de cada mensagem. O algoritmo usado é baseado em vetores de texto que são usados para representar textos e assuntos. O método avalia a similaridade entre o texto e um assunto usando uma função de similaridade que calcula a distância entre os dois vetores, um representando a mensagem do *chat* e outro representando o assunto, presente na ontologia.

O método é baseado em um índice de relevância, que é um conjunto de características que, juntas, podem predizer, com confiança, a descrição ou existência de um evento. Partindo desta premissa, o método considera que alguns termos presentes nas mensagens do *chat* possam, portanto, indicar a presença de um assunto com um certo grau de certeza. Conseqüentemente, o processo de raciocínio deverá avaliar a probabilidade de um assunto estar presente em um texto, analisando a intensidade destas indicações, ou seja a presença de termos. Isto significa que, se as palavras que descrevem um assunto qualquer aparecem em um texto, existe uma probabilidade alta desse assunto estar presente no texto. A soma das indicações deve resolver problemas de ambigüidade. Por exemplo, a presença do termo ‘inteligência’ gera uma ambigüidade por poder se referir tanto ao tema ‘inteligência artificial’ quanto ao tema ‘inteligência competitiva’, mas a presença de outros termos ajuda a resolver tal conflito.

No contexto do portal proposto neste estudo, acredita-se que a aplicação dos resultados do trabalho descrito poderá ajudar no processo de identificação de temas em mensagens de *chat*, de forma a servir de auxílio para que pessoas especialistas nos assuntos discutidos possam ser encontradas ou de forma a indicar itens de forma sensível ao contexto, oferecendo conteúdos personalizados conforme o interesse de cada pessoa que faz uso da ferramenta de colaboração.

CAPÍTULO 4 - Desenvolvimento da ontologia do portal semântico

Tendo em vista o objetivo desta pesquisa, que é buscar compreender como o uso de ontologias apóia a organização do conhecimento, este capítulo mostra o processo de construção de uma ontologia denominada ONTO-POINT, que servirá de ponto de apoio para o portal semântico a ser proposto e servirá de base para todas as funções semânticas que também serão propostas na seqüência deste trabalho.

4.1) A ontologia ONTO-POINT

A ontologia ONTO-POINT deve prover uma representação conceitual do ambiente da Universidade Presidente Antônio Carlos (UNIPAC), suas unidades, cursos oferecidos, o projeto pedagógico de cada um desses cursos, além de classificar os atores envolvidos no dia-a-dia da universidade. Os assuntos que são de interesse de cada um desses atores e os assuntos em que cada um se considera especialista também devem fazer parte da modelagem. A ontologia também deve prover uma representação de todo o conteúdo que será publicado e recuperado por estes atores que freqüentarão o portal semântico.

De acordo com a classificação quanto ao nível de generalidade proposta por GUARINO (1998), a ontologia a ser construída neste trabalho pode ser considerada como uma ontologia de aplicação, pelo fato de descrever conceitos que são dependentes de um domínio particular ou de tarefa. Este tipo de ontologia pode ser entendido como uma especialização de outras ontologias de um nível superior, chamadas ontologias de domínio e de tarefas, que, por sua vez, descrevem, respectivamente, o vocabulário relacionado a domínios genéricos ou a tarefas e atividades genéricas, através da especialização de termos introduzidos nas ontologias de alto-nível.

Desta forma, para a construção da ontologia ONTO-POINT, foram necessárias pesquisas em duas frentes distintas: a primeira procurou descobrir a melhor forma de representar a estrutura da universidade e seus diversos elementos constituintes, enquanto a segunda focou a representação do conteúdo e dos diversos tipos de documentos a serem armazenados e posteriormente disponibilizados no portal.

4.2) A construção da ontologia ONTO-POINT

Para o processo de elaboração da ontologia deste estudo, optou-se por utilizar a versão 2000 da ferramenta Protegé, desenvolvida pela Universidade de Stanford, descrita na Seção 3.1. Foram utilizadas as funcionalidades do editor de ontologias, através da extensão OWL, que apóia a construção de ontologias para a Web Semântica.

A extensão OWL pode ser utilizada para editar ontologias OWL, acessar classificadores de descrição lógica e construir instâncias para marcação semântica. A extensão OWL tem o potencial de tornar-se uma infra-estrutura padrão para a construção de aplicações para a Web Semântica baseada em ontologias (KNUBLAUCH e MUSEN, 2004).

Diversas metodologias têm sido desenvolvidas no intuito de sistematizar a construção e a manipulação de ontologias (LOPEZ, 1999). Encontra-se na literatura propostas diversas para a construção de ontologias com abordagens e características diversas, muitas vezes desenvolvidas para aplicações distintas ou diferentes objetivos. Neste estudo, optou-se por adotar o roteiro proposto por NOY e McGUINESS (2001), que consiste em sete passos e será detalhado a seguir.

4.2.1) Passo 1 - Determinação do domínio e escopo da ontologia

Conforme declarado anteriormente, a ontologia de aplicação a ser construída deve representar o ambiente da universidade, além de prover uma classificação hierárquica e relacionamentos que contemplem os diversos assuntos e documentos que serão objeto de trabalho do portal semântico a ser proposto. Para melhor entendimento e clareza, a partir dos próximos passos, será chamada de Tarefa 1, o trabalho de representação da estrutura interna da universidade, e Tarefa 2, o trabalho de representação dos assuntos de interesse e documentos a serem publicados e recuperados no portal.

4.2.2) Passo 2 - Considerar a reutilização de ontologias existentes

Para a Tarefa 1, o primeiro desafio encontrado foi a inexistência de padrões que pudessem representar a estrutura interna de uma IES (Instituição de Ensino Superior), devido às óbvias diferenças existentes entre elas. Procurou-se então, construir a ontologia baseando-se em uma possível padronização, por parte do Ministério da Educação (MEC), no que diz respeito a diretrizes curriculares nacionais para as instituições de ensino e os cursos de graduação, de uma forma geral. Pesquisando o sítio do MEC [51], constatou-se não existir uma regra geral de nomenclatura para formação dos currículos dos cursos superiores no país e que uma comissão de especialistas de cada área formula as diretrizes de seus cursos sem se preocupar com as outras existentes. Acredita-se que as classes propostas na ontologia deste trabalho possam representar, de forma genérica, as diferentes sugestões das diferentes comissões de especialistas que definem para cada curso, ou área de cursos, diferentes conteúdos curriculares.

Para a Tarefa 2, foi feita uma pesquisa tentando encontrar alguma ontologia existente que pudesse representar documentos e assuntos a serem tratados pelos atores do portal proposto. O KSL - *Knowledge Systems, AI Laboratory* [52] - conduz pesquisas na área de representação do conhecimento e processos automatizados no Laboratório de Inteligência Artificial do Departamento de Ciência da Computação da Universidade de Stanford. Os trabalhos atuais enfatizam a disponibilização e o uso de tecnologias para a Web Semântica. Dentro do portal Stanford KSL Network Services são encontrados vários serviços disponíveis, entre os quais um editor de ontologias. Dentro da página deste editor, é oferecida uma biblioteca de ontologias, que se propõe a ser um padrão de utilização mundial. Nesta biblioteca foi encontrada uma ontologia de documentos, denominada *Documents*, que atende aos requisitos necessários a este estudo e por isso, passa a ser utilizada neste trabalho, fazendo parte da ontologia ONTO-POINT.

[51] http://www.mec.gov.br
[52] http://www.ksl.stanford.edu

4.2.3) Passo 3 - Enumerar termos importantes da ontologia

Durante a execução deste passo, os conceitos mais importantes do domínio foram listados, sem a preocupação de classificá-los ou relacioná-los uns aos outros. Em grande parte dos casos, cada conceito listado se transformou em uma classe da modelagem da ontologia (que será apresentada detalhadamente no Passo 4), momento em que houve a preocupação de agrupamento e classificação hierárquica, para que pudessem fazer parte de uma taxonomia que passaria posteriormente a dar origem à ontologia ONTO-POINT.

4.2.4) Passo 4 - Definir as classes e a hierarquia de classes

Segundo USCHOLD e GRUNINGER (1996), existem algumas abordagens possíveis no desenvolvimento de uma hierarquia de classes: *top-down*, *bottom-up* e a combinação de ambos os processos. A abordagem *top-down* começa com a definição dos conceitos mais gerais do domínio e a subseqüente especialização dos mesmos. Já a abordagem *bottom-up* se inicia com a definição dos conceitos mais específicos da hierarquia para depois serem agrupados em conceitos mais genéricos. Para a construção da ontologia ONTO-POINT, foi realizada a combinação dessas duas abordagens, sendo os conceitos mais importantes descobertos e posteriormente realizadas atividades de generalização ou especialização dos mesmos.

Desta forma, durante a Tarefa 1, foram criadas as classes UNIVERSIDADE, UNIDADE e CURSO. Adotando as diretrizes dos cursos da área da computação como padrão dentre as muitas encontradas na pesquisa ao sítio do MEC, por considerar que elas podem representar satisfatoriamente a maioria dos demais cursos existentes, sugere-se uma distribuição dos currículos dos cursos em áreas de formação, que por sua vez são divididas em áreas e sub-áreas do conhecimento, sendo estas últimas similares a matérias ou disciplinas. Esta divisão proposta pela área da computação pode ser vista na Figura 4.1.

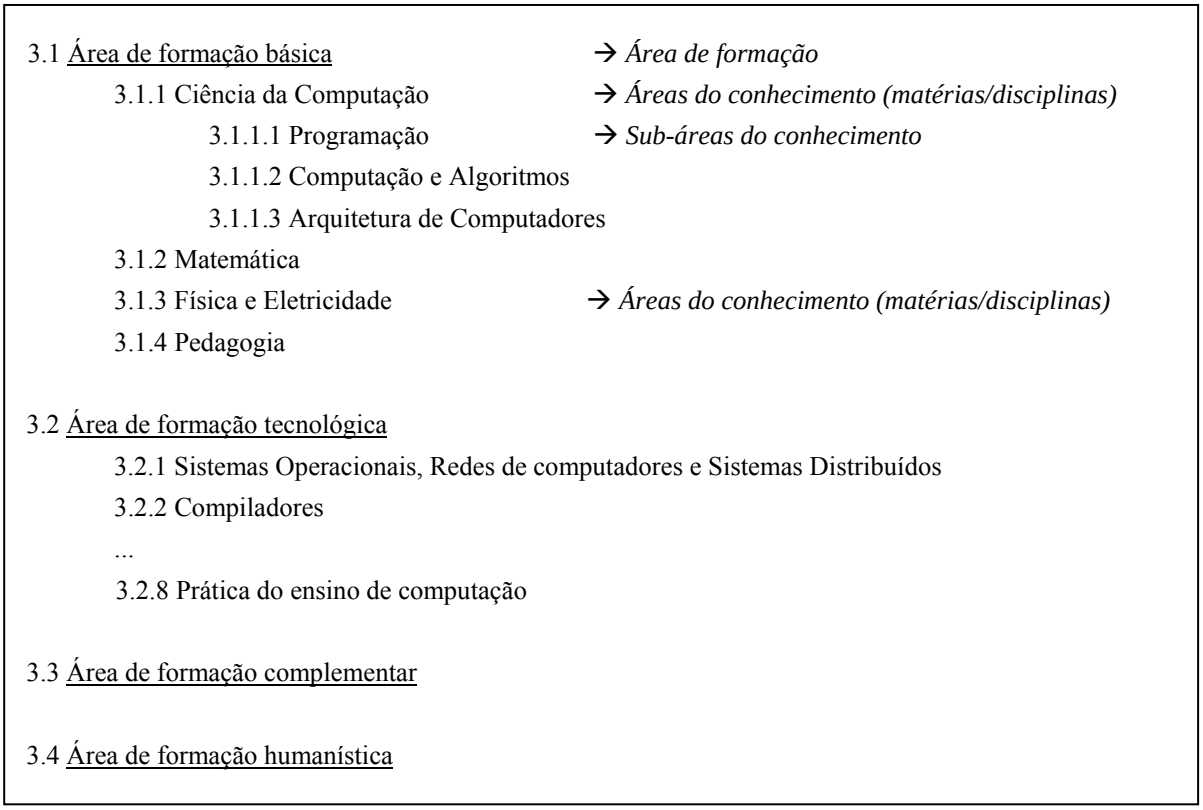


Figura 4.1 - Sugestão de áreas e sub-áreas do conhecimento referente a cursos da Área da Computação.

Foram criadas então as classes CURRÍCULO, ÁREA DE FORMAÇÃO, DISCIPLINA, EMENTA, CONTEÚDO PROGRAMÁTICO e BIBLIOGRAFIA. A classe DISCIPLINA foi declarada como sinônimo de MATÉRIA e ÁREA DO CONHECIMENTO. As classes ALUNO**,** PROFESSOR**,** COORDENADOR DE CURSO e PESQUISADOR também foram criadas, herdando as propriedades da classe ATOR, a que elas estão ligadas em uma estrutura do tipo Genaralização-Especialização. A classe ATOR, por sua vez, é sub-classe da classe PESSOA, assim como a classe VISITANTE**.**

Todas estas classes referentes à Tarefa 1 podem ser visualizadas na Figura 4.2, que mostra um fragmento da ontologia elaborada através da ferramenta Protégé 2000 com extensão OWL, com destaque para as classes que compõem a representação interna da Universidade.

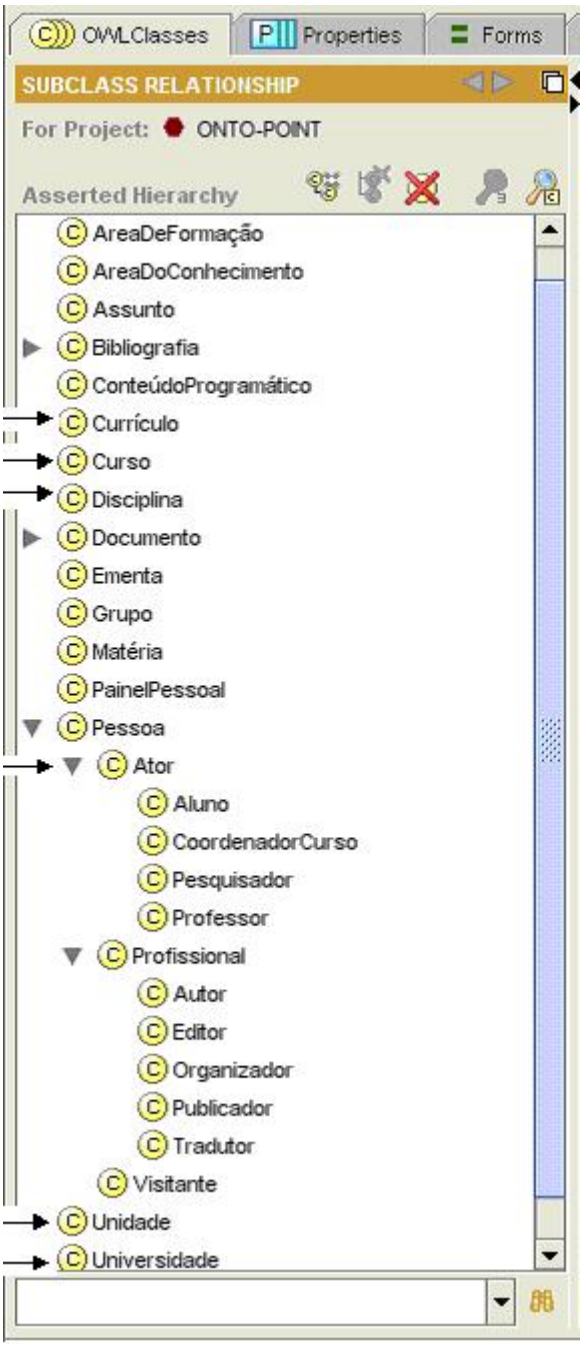


Figura 4.2 – Primeiro fragmento da ontologia ONTO-POINT, montada com a ferramenta Protégé 2000 com extensão OWL

Durante a Tarefa 2, foi tomada como base a ontologia *Documents,* mencionada anteriormente e construída na língua inglesa. Optou-se por fazer a tradução dos termos para o português, para não haver conflitos de línguas dentro da ontologia e também por causa do caráter de regionalismo que deve ser preservado dentro do estudo que está sendo proposto.

Assim, foram aproveitadas as classes DOCUMENTS (DOCUMENTOS) e todas as suas sub-classes: BOOK (LIVRO), MISCELANNEOUS PUBLICATION (PUBLICAÇÃO DIVERSA), PERIODICAL PUBLICATION (PUBLICAÇÃO PERIÓDICA), PROCEEDINGS (ANAIS), TECHNICAL REPORT (RELATÓRIOS TÉCNICOS) E THESIS (TESE).

Algumas dessas sub-classes também têm classes filhas, como por exemplo, BOOK que tem a sub-classe EDITED BOOK (LIVRO EDITADO), a classe MISCELLANEOUS PUBLICATION que tem as sub-classes ARTWORK (TRABALHO ARTÍSTICO), CARTOGRAPHIC MAP (MAPA CARTOGRÁFICO), COMPUTER PROGRAM (PROGRAMA DE COMPUTADOR), MULTIMEDIA DOCUMENT (DOCUMENTO MULTIMÍDIA) e TECHNICAL MANUAL (MANUAL TÉCNICO).

Também a classe PERIODICAL PUBLICATION tem filhas que são JOURNAL (não traduzido e mantido como JOURNAL), MAGAZINE (REVISTA) E NEWSPAPER (JORNAL) e a classe THESIS tem filhas que são MASTER THESIS (DISSERTAÇÃO DE MESTRADO) e DOCTORAL THESIS (TESE DE DOUTORADO).

Ainda dentro desta segunda tarefa, estão presentes na ontologia as classes PROFISSIONAL**,** filha da classe PESSOA e pai das classes AUTOR, EDITOR, TRADUTOR, PUBLICADOR e ORGANIZADOR, necessárias à elaboração de muitos dos relacionamentos entre as classes aqui propostas. Todas estas classes referentes à Tarefa 2 podem ser visualizadas na Figura 4.3, que mostra um fragmento da ontologia elaborada através da ferramenta Protégé 2000 com extensão OWL, com destaque para as classes que representam os assuntos e documentos publicados e recuperados no portal

Figura 4.3 – Segundo fragmento da ontologia ONTO-POINT, montada com a ferramenta Protégé 2000 com extensão OWL

4.2.5) Passo 5 - Definir as propriedades das classes

As classes sozinhas não provêm informações suficientes para responder as questões advindas do uso da ontologia. Por isso, é necessária a definição de propriedades para cada uma delas, ou seja, estabelecer a estrutura interna dos conceitos. As propriedades[53] podem ser de dados ou de objetos, que representam relacionamentos com outras classes.

Para a Tarefa 1, foram criados os seguintes relacionamentos: *"tem_unidade"* e o relacionamento inverso *"é_da_universidade"* entre as classes UNIVERSIDADE e UNIDADE; *"tem_curso"* e o relacionamento inverso *"é_da_unidade"* entre as classes UNIDADE e CURSO; *"tem_currículo"* e o relacionamento inverso *"pertence_ao_curso"* entre as classes CURSO e CURRÍCULO; *"tem_área_formação"* e o relacionamento inverso *"é_do_currículo"* entre as classes CURRÍCULO e ÁREADEFORMAÇÃO; *"contém_disciplinas"* e o relacionamento inverso *"é_da_área_formação"* entre as classes ÁREADEFORMAÇÃO e DISCIPLINA; *"é_dividida_em_disciplina"* e o relacionamento inverso *"faz_parte_da_disciplina"* entre as classes DISCIPLINA e DISCIPLINA (auto-relacionamento); *"tem_ementa"* e o relacionamento inverso *"descreve_a_disciplina"* entre as classes DISCIPLINA e EMENTA; *"tem_conteúdo_programático"* e o relacionamento inverso *"compõe_a_disciplina"* entre as classes DISCIPLINA e CONTEÚDOPROGRAMÁTICO, além do relacionamento *"tem_bibliografia"* e o relacionamento inverso *"é_da_disciplina"* entre as classes DISCIPINA e BIBLIOGRAFIA.

O relacionamento *"tem_ator"* e o relacionamento inverso *"pertence_a_unidade"* foram criados entre as classes UNIDADE e ATOR e outros relacionamentos também foram criados entre os diversos atores e as demais classes, a saber : *"está_matriculado"* e o inverso *"tem_matrículas"* entre as classes ALUNO e CURSO; *"estuda"* e o inverso *"é_estudada"* entre as classes ALUNO e DISCIPLINA; *"ministra"* e o inverso *"é_ministada"* entre as classes PROFESSOR e DISCIPLINA; *"coordena"* e *"é_coordenado"* entre as classes COORDENADOR DE CURSO e CURSO.

Uma classe ASSUNTO também foi criada para efeitos de representar os relacionamentos *"é-especialista_em"* e o seu inverso *"tem_especialista"* e o relacionamento

[53] do termo inglês *slots*

"se_interessa_por" e o seu inverso *"é_interessante_para"* que existem entre ela e a classe ATOR.

A Figura 4.4 mostra no detalhe a classe ALUNO (A) com seus relacionamentos (B) e propriedades (C), que fazem parte da ontologia ONTO-POINT.

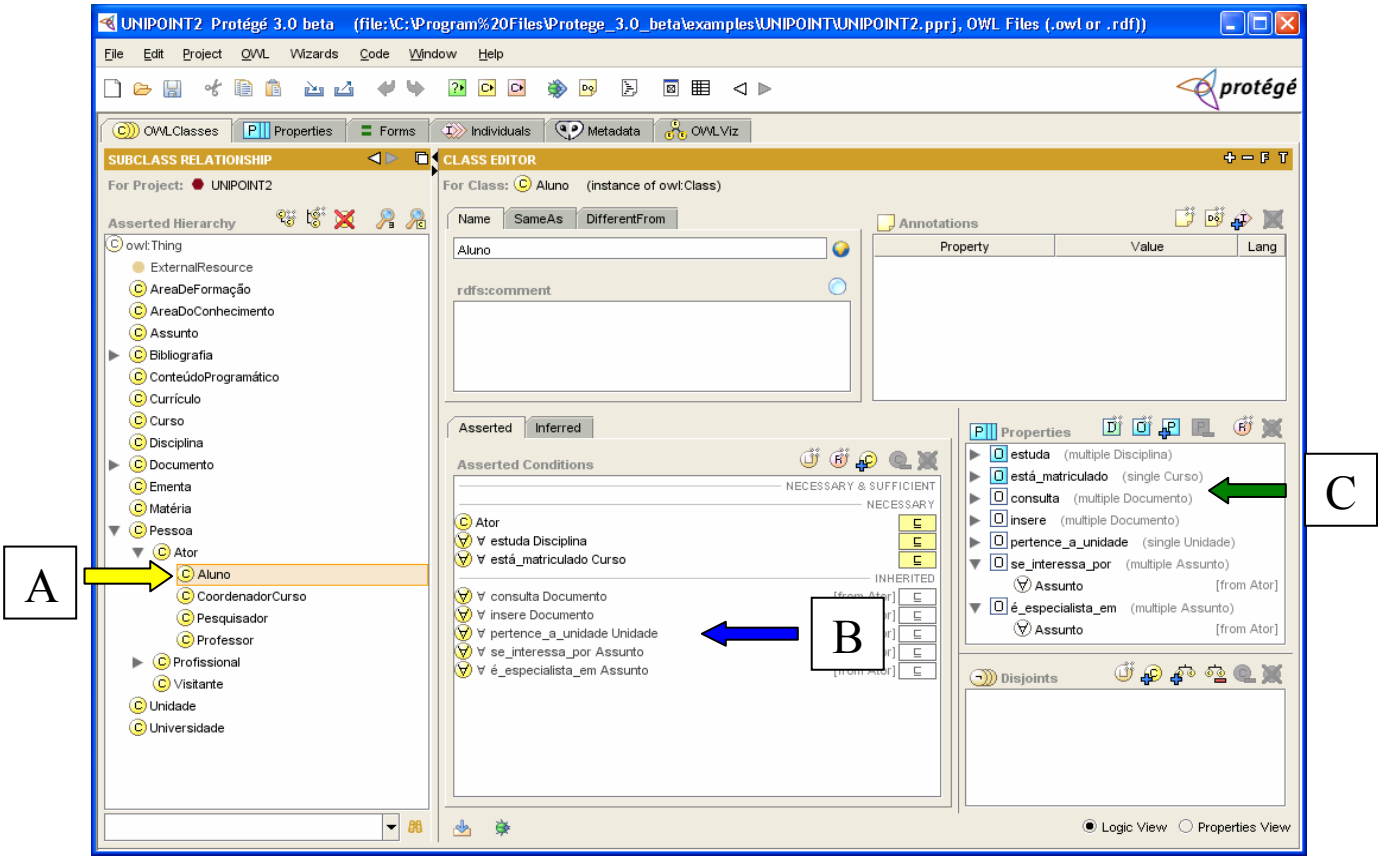


Figura 4.4 - Ontologia ONTO-POINT com detalhe para a classe Aluno (A), seus relacionamentos (B) e suas propriedades (C).

Para a Tarefa 2, foram aproveitados também os relacionamentos: *"tem_tradutor"* e o relacionamento inverso *"é_tradutor"* entre as classes DOCUMENTOS e TRADUTOR; *"tem_autor"* e o relacionamento inverso *"é_autor"* entre as classes DOCUMENTOS e AUTOR; *"tem_publicador"* e o relacionamento inverso *"é_publicador"* entre as classes DOCUMENTOS e PUBLICADOR; *"tem_organizador"* e o relacionamento inverso *"é_organizador"* entre as classes DOCUMENTOS e ORGANIZADOR; *"é_editor"* e o relacionamento inverso *"tem_editor"* entre as classes LIVRO EDITADO e EDITOR.

A classe ASSUNTO se liga à classe DOCUMENTOS representando o relacionamento *"diz_respeito"* e o seu inverso *"é_encontrado"* e a classe ATOR também se liga à classe DOCUMENTOS representando o relacionamento *"insere"* e o seu inverso *"inserido_por"* e o relacionamento *"consulta"* e o inverso *"consultado_por"*.

Na Figura 4.5 é apresentada a ontologia ONTO-POINT com detalhe para a classe DOCUMENTOS e todas as suas sub-classes(A), seus relacionamentos (B) e suas propriedades (C).

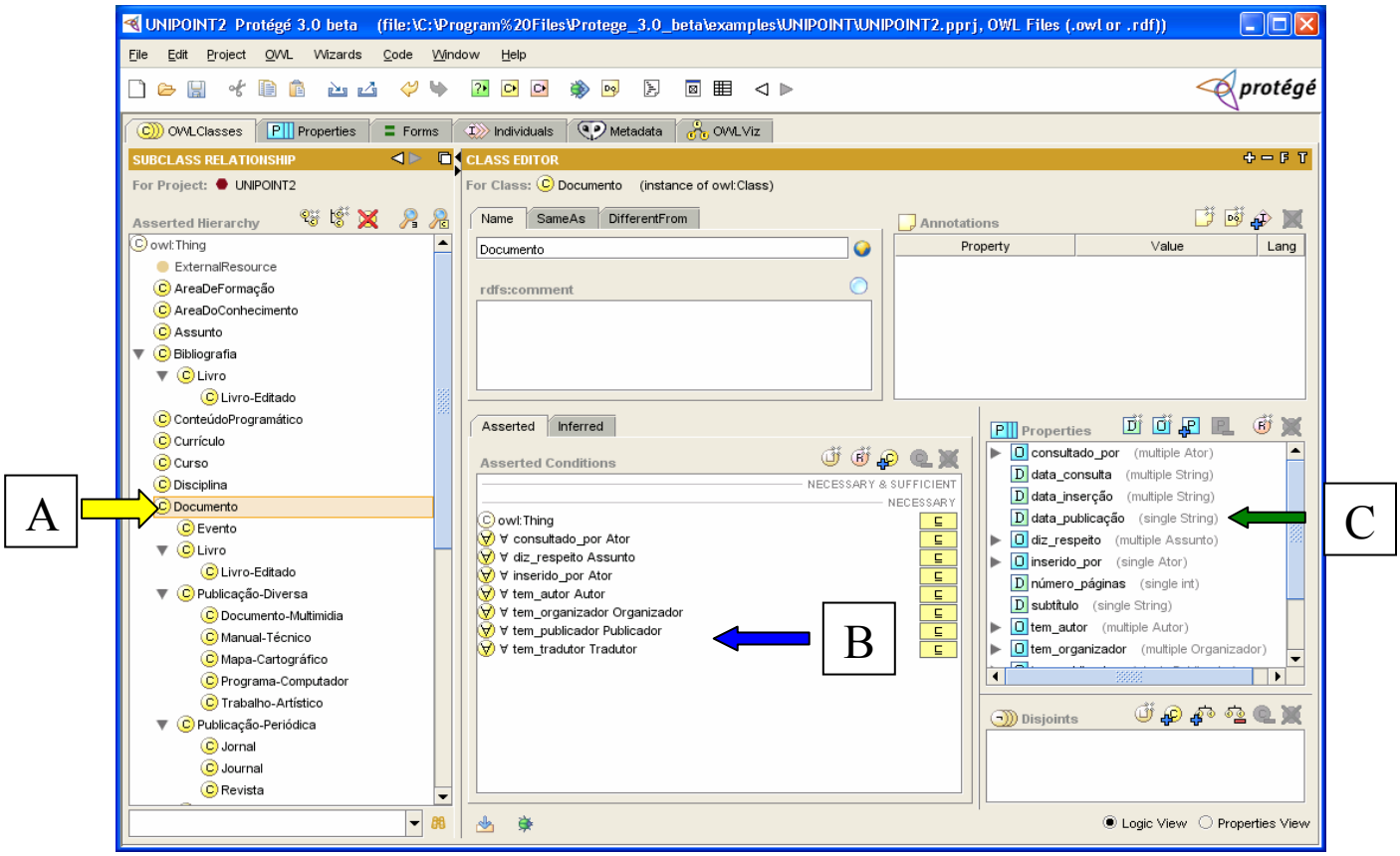


Figura 4.5 - Ontologia ONTO-POINT, com detalhe para a classe DOCUMENTO (A), seus relacionamentos (B) e suas propriedades (C).

O Anexo I mostra um resumo de todas as classes e relacionamentos criados durante a elaboração da ontologia ONTO-POINT.

4.2.6) Passo 6 - Definir as restrições das propriedades

As classes possuem características limitadoras que são chamadas restrições[54]. Algumas restrições possíveis são a cardinalidade do relacionamento[55], o tipo de valor do relacionamento[56], a faixa de valores possível[57] para o relacionamento e o domínio de um relacionamento[58]. Neste passo, foram incluídos estes tipos de características para algumas classes e relacionamentos, observando que nem todas as classes ou relacionamentos possuem características limitadoras a serem definidas.

Como exemplo de cardinalidade, uma disciplina pode ter até seis livros em sua bibliografia básica. Quanto ao tipo de valor do relacionamento, os mais comuns são boleanos, texto e números, mas também existe o 'tipo-instância' que é um tipo de valor que define relacionamentos entre os indivíduos. Como exemplo pode-se citar o relacionamento "*tem_ementa*" entre as classes Disciplina e Ementa. Os relacionamentos ligam instâncias de um domínio - como por exemplo a classe Disciplina retratada anteriormente - a instâncias de uma faixa de valores possível, como a classe Ementa, também retratada no exemplo anterior.

4.2.7) Passo 7 - Criar as instâncias de classes

A última etapa no processo de criação da ontologia é o povoamento, ou preenchimento da mesma. Neste momento escolhe-se uma classe e cria-se uma instância individual para ela, preenchendo os valores para as propriedades e relacionamentos, além de possíveis restrições existentes. Durante este passo, várias classes podem ser instanciadas, mas outras permanecem vazias, pois só serão instanciadas quando da efetiva utilização da ontologia, através dos serviços disponibilizados no portal semântico.

[54] do termo inglês *facets*
[55] do termo inglês *slot cardinality*
[56] do termo inglês *slot-value type*
[57] do termo inglês *range*
[58] do termo inglês *domain*

4.3) A ontologia ONTO-POINT em RDF e RDFS

A Figura 4.6 mostra parte do código fonte da ontologia ONTO-POINT construída em OWL / RDF / RDFS, com detalhe para a tripla formada pela classe "*CoordenadorCurso"*, a propriedade "*coordena"* e o recurso "*Curso"*.

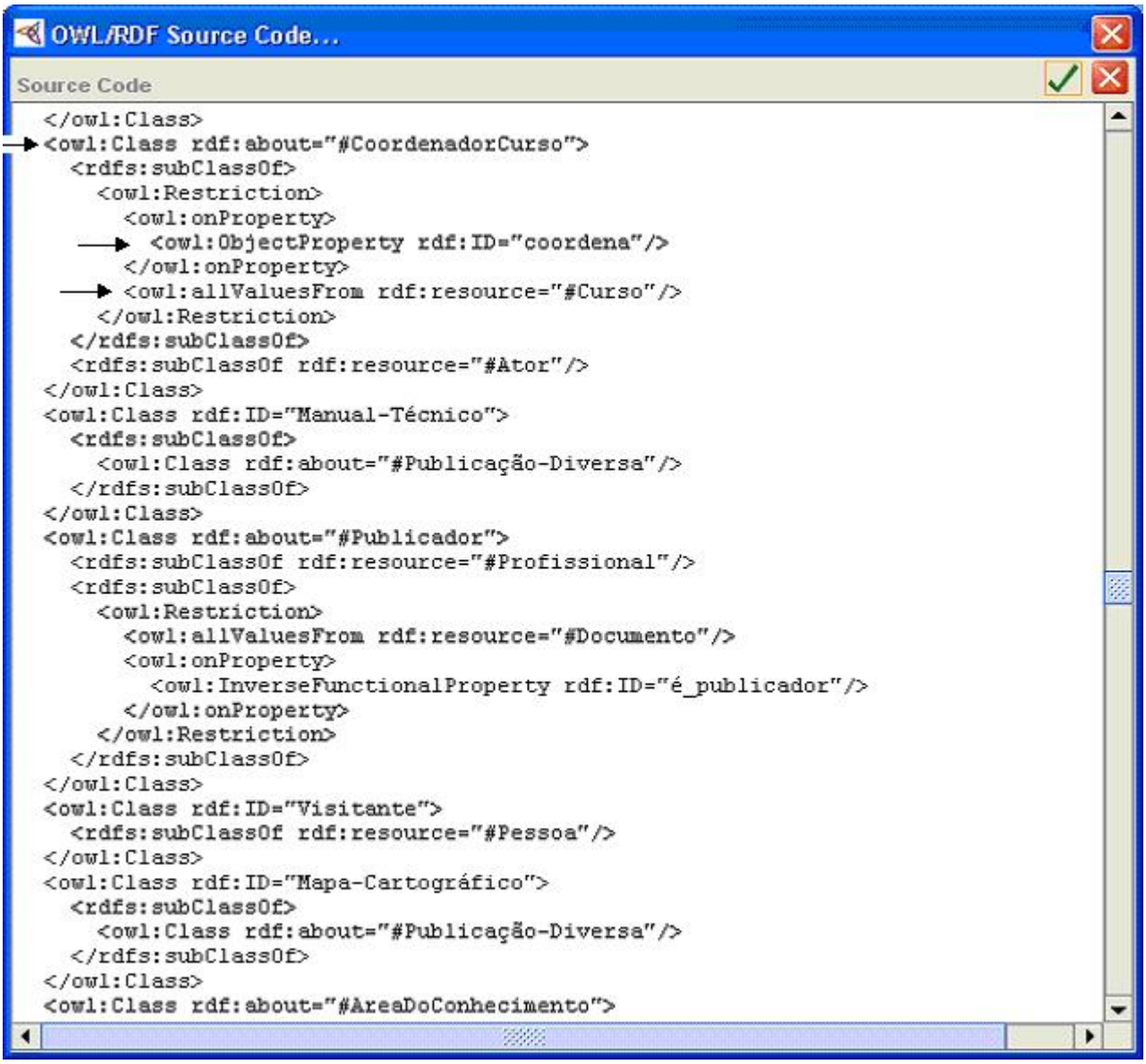


Figura 4.6 – Código fonte da ontologia ONTO-POINT elaborado pelo Protégé 2000 em OWL / RDF

Conforme já visto no Capítulo 2, RDF(S) é um modelo abstrato de dados que define relacionamentos entre recursos. Antes de ser introduzido um modelo de dados RDFS, é exemplificado na Figura 4.7 um modelo de dados em RDF.

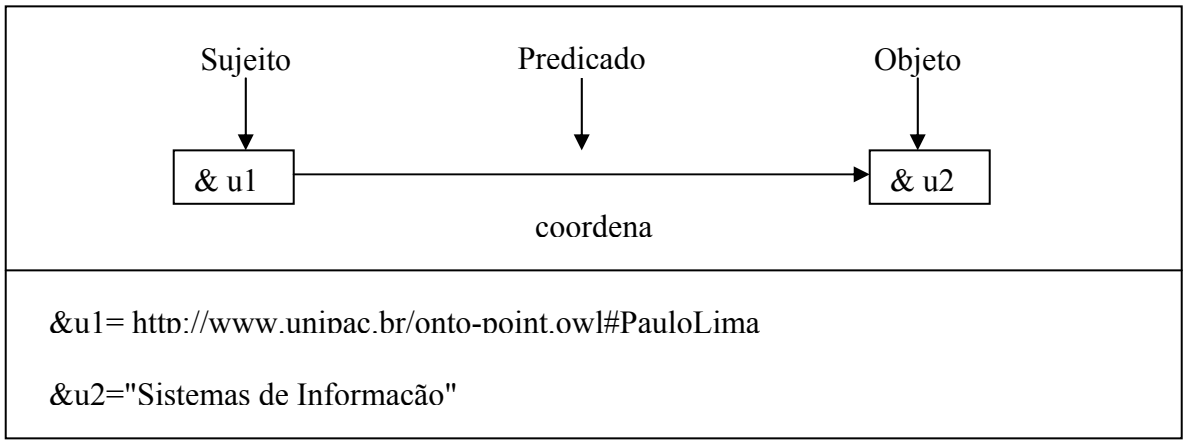
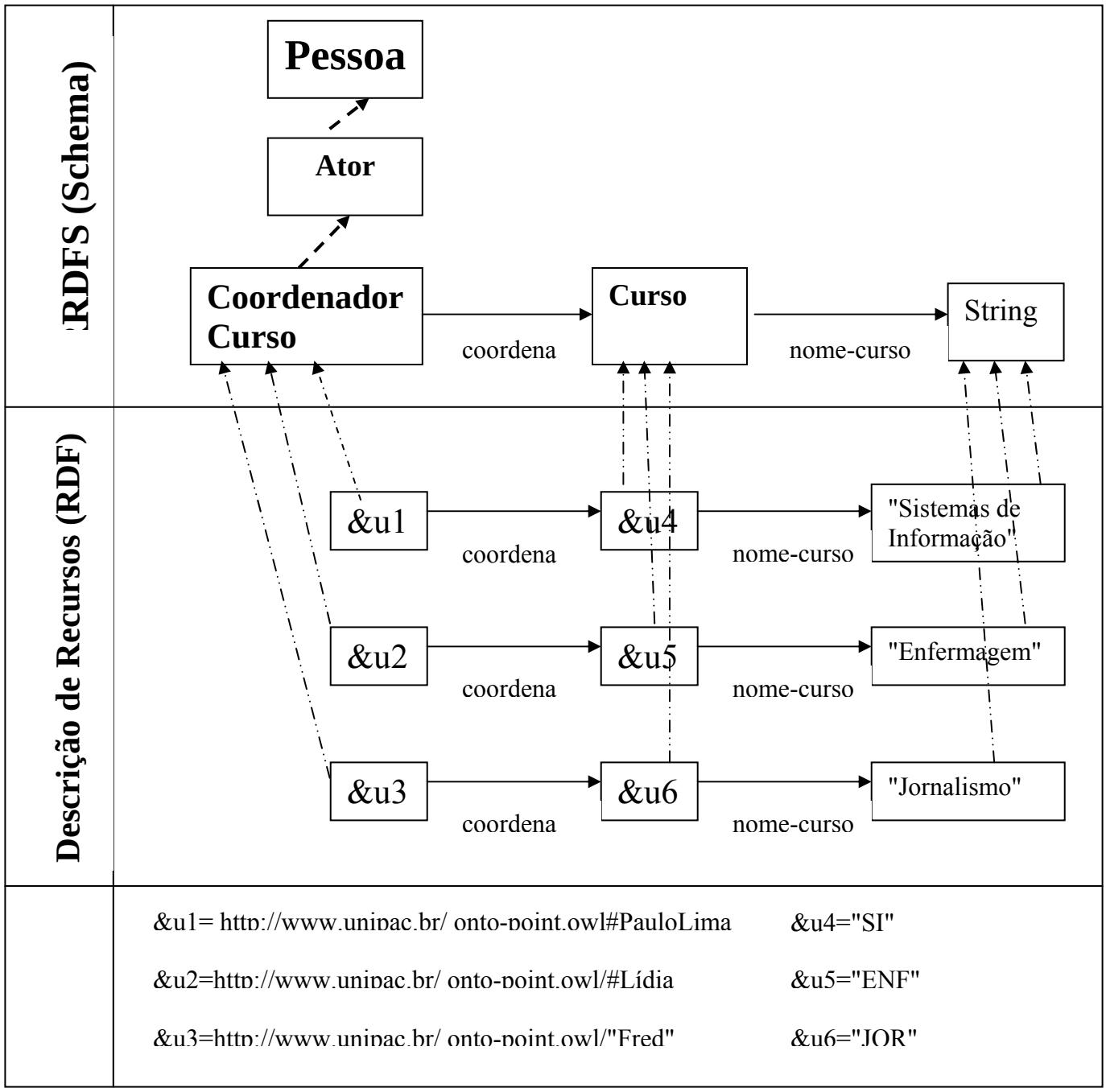


Figura 4.7 - Modelo de dados RDF, com detalhe para o sujeito CORDENADORCURSO, o predicado *coordena* e o objeto CURSO

A figura acima retoma os conceitos de recurso, propriedade e valor, ou, sujeito, predicado e objeto, sendo o sujeito representado pelo recurso &u1(URL de um coordenador de curso), o predicado pela propriedade "*coordena"* e o objeto pelo recurso &u2 (URL de um curso).

Para melhor entendimento das propriedades da orientação a objetos, tais como herança, é apresentado o mesmo modelo da figura anterior ampliado e incorporado a um RDF(S) na Figura 4.8. Na camada superior é apresentado o esquema, que define algumas classes e relações relevantes para o contexto do documento de estudo de caso. Na camada inferior, tem-se os dados atuais que fazem parte do documento de estudo de caso modelados no formato RDF.



Legenda : ——→ **Relações**

- - - - - - - → **Especialização / Generalização**

– · – · – → **Instanciação**

Figura 4.8 - Modelo de dados RDF incorporado de um Schema

4.4) Consultas nas triplas RDF utilizando RDQL

A partir do momento em que são gravadas as triplas RDF em um sistema de armazenamento de recursos do tipo RDF, torna-se possível a recuperação das informações a partir de linguagens de busca especializadas, como por exemplo, RDQL.

Se o objetivo de uma consulta fosse, por exemplo, encontrar todos os coordenadores de curso e seus respectivos endereços eletrônicos em toda a universidade, ela seria representada em RDQL conforme mostra a Figura 4.9.

```
SELECT ?coordenadorcurso, ?email
WHERE (?coordenadorcurso, ontopoint:EMAIL, ?email)
 USING ontopoint FOR <http://www.unipac.br/ontopoint/ontopoint-rdf#>
```

Figura 4.9 - Exemplo de consulta em RDQL

Alguns dos resultados possíveis para a consulta acima são mostrados na Figura 4.10:

```
          coordenadorcurso              |            email
========================================================================
<http://somewhere/Lidia Santos Soares/>  | lídia@lafaiete.unipac.br
<http://somewhere/Edson de Souza Pinto/> | edson@barbacena.unipac.br
<http://somewhere/Camila Maria da Silva/>| camila@juizdefora.unipac.br
```

Figura 4.10 - Possível resultado da consulta em RDQL

CAPÍTULO 5 - O portal UNIPAC

Este capítulo tem como objetivo apresentar o portal tradicional da UNIPAC [59] (Universidade Presidente Antônio Carlos), instituição de ensino superior mantida pela FUPAC (Fundação Presidente Antônio Carlos), cuja sede está localizada na cidade mineira de Barbacena e possui diversas unidades e campi universitários em várias cidades do interior de Minas Gerais. Será apresentada também a proposta de elaboração do portal semântico UNI-POINT, cujo objetivo é ser uma espécie de "ponto de encontro" da comunidade universitária da UNIPAC.

5.1) O Portal tradicional da UNIPAC

O portal da UNIPAC existente em 2005 na Internet atende aos interesses da instituição enquanto fonte de informações para usuários não alunos e visitantes, interessados em obter mais informações sobre essa instituição de ensino superior e também aos candidatos a vagas nos diversos cursos de graduação, pós-graduação e mestrado existentes, oferecidos pela instituição.

Para os alunos regularmente matriculados, o portal avança na oferta de serviços on-line, oferecendo alguma comodidade ao aluno com relação ao pagamento de suas mensalidades, obtenção de notas e faltas apuradas durante o período letivo e a possibilidade de acesso à biblioteca de sua unidade, onde pode realizar pesquisas no acervo disponível, checar a existência de reservas dos títulos, além de se beneficiar com os serviços de renovação de empréstimo e colocação de reserva para uma obra de seu interesse. Na página dos alunos, os discentes podem consultar a matriz curricular do curso no qual está matriculado e obter o seu histórico escolar, entre outros serviços que são disponibilizados.

Sente-se, entretanto, falta de um ambiente onde os alunos possam se encontrar para trocar informações e material didático, assim como encontrar seus professores para uma possível resolução de dúvidas ou distribuição de material de sala de aula. Além disso, coordenadores de curso, professores e pesquisadores ainda não possuem um espaço próprio de interação disponível dentro do portal atual.

É principalmente por causa desses motivos que o presente trabalho sugere a criação de um portal semântico para a universidade e seus diversos campi, de forma a unir a comunidade acadêmica e permitir o compartilhamento de informações e do conhecimento, matéria prima primordial no trabalho diário de uma instituição de ensino.

5.2) A proposta do portal semântico da UNIPAC

Conforme citado anteriormente, a UNIPAC possui hoje dez campi e outras dez unidades universitárias espalhadas pelo interior de Minas Gerais. Muitas destas unidades e campi têm cursos semelhantes - cinco cursos de Sistemas de Informação, por exemplo - que não necessariamente possuem o mesmo projeto pedagógico e a mesma matriz curricular de disciplinas. Por motivos óbvios, os alunos de mesmo curso não se conhecem, mas a distância geográfica entre as unidades faz com que os professores, coordenadores de cursos similares e pesquisadores também não se conheçam. O portal objetiva diminuir esta distância, colocando os atores de uma determinada unidade em contato com os seus pares, através da Web, onde será possível usufruir de uma série de serviços.

5.2.1) UNI-POINT - O "Ponto de Encontro" da comunidade educacional da UNIPAC

O portal UNI-POINT está sendo sugerido como um portal educacional semântico para a comunidade ligada à UNIPAC, para a troca e compartilhamento de informações e conhecimento. Para que esse portal possa realmente se tornar um "ponto de encontro" onde os atores possam interagir, é necessário que ele ofereça aos usuários alguns serviços essenciais que são apresentados nesta seção. A Figura 5.1 mostra um exemplo do que pode ser a página principal do portal, personalizada por um usuário.

[59] http://www.unipac.br

Figura 5.1 – Um exemplo do que pode ser a página principal do portal semântico UNIPAC

O Diagrama de Contexto da Figura 5.2 é um tipo especial de Diagrama de Caso de Uso, de acordo com o padrão UML (BOOCH et al, 2000), utilizado para a modelagem de aspectos dinâmicos de sistemas. Ele tem como objetivo mostrar em formato gráfico os serviços propostos para o portal e seus relacionamentos com os atores.

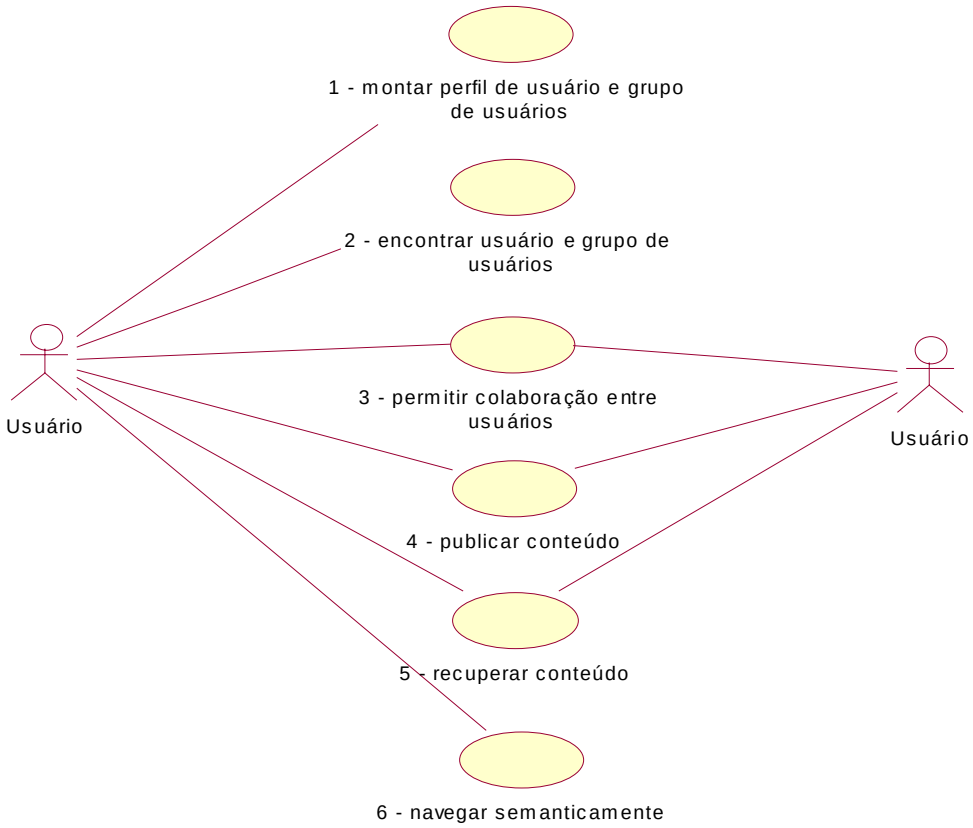


Figura 5.2 – Diagrama de Contexto mostrando os serviços disponíveis no portal

5.2.2) Serviços a serem disponibilizados pelo UNI-POINT

Esta seção mostra os serviços que poderão ser oferecidos ao usuário e tem como objetivo único ajudar a entender o que o portal se propõe a fazer. É entendido que esta organização em serviços, apesar de não trazer ganho no que diz respeito ao objetivo final do trabalho (que é compreender como as ontologias agregam valor aos portais), ajuda a desenhar o pano de fundo do trabalho que é o portal semântico. Posteriormente, no decorrer do trabalho, através da apresentação das funções semânticas e com a utilização da ontologia construída, mostra-se como fazer o que está sendo proposto, sendo, neste momento, possível verificar os benefícios advindos de sua utilização.

A Tabela 5.1 que se segue apresenta cada um dos serviços propostos para o portal semântico, descrevendo suas características principais e os seus benefícios para os usuários.

Serviço do portal	Principal característica
Serviço 1: ***"montar perfil de usuário e grupo de usuários"***	Permite a montagem de perfis de cada usuário, além de montagem de perfis de grupos de usuários e o gerenciamento (entrada e saída) de usuários dentro dos grupos.
Serviço 2: ***"encontrar usuário e grupo de usuários"***	Permite o encontro, através de navegação ou busca, de pessoas ou grupos com os quais os usuários querem fazer contato, trocar ou compartilhar informações e conhecimento.
Serviço 3: ***"permitir colaboração entre os usuários"***	Permite a comunicação (síncrona ou assíncrona) entre os usuários do portal e estende a funcionalidade de outros serviços como publicação e recuperação de conteúdo.
Serviço 4: ***"publicar conteúdo"***	Permite o armazenamento de conteúdo marcado semanticamente por qualquer usuário do portal, de forma que possa ser recuperado pelos outros usuários a qualquer momento.
Serviço 5: ***"recuperar conteúdo"***	Permite, através de navegação ou consulta, descobrir informações ou documentos relevantes no repositório de metadados do portal.
Serviço 6: ***"navegar semanticamente"***	Permite a descoberta de conceitos marcados semanticamente em um documento e a obtenção de informações adicionais relacionadas a eles.

Tabela 5.1 – Serviços propostos para o portal semântico UNI-POINT

Nos anexos deste trabalho podem ser encontradas uma descrição e uma breve especificação de cada um desses serviços, com a apresentação de Diagramas de Atividades e telas de interface.

CAPÍTULO 6 - Funções semânticas do portal

Este capítulo apresenta a arquitetura funcional proposta para o portal, descreve as funções semânticas sugeridas e explora como se daria a aplicação das ferramentas apresentadas no Capítulo 3 no portal semântico, exemplificando tal aplicação com o uso da ontologia construída no Capítulo 4.

6.1) Arquitetura proposta para o portal semântico UNI-POINT

No capítulo anterior foi mostrado que, para que o portal proposto possa se tornar um ambiente de interação entre os usuários, é necessário que ele ofereça alguns serviços essenciais. Esses serviços, por sua vez, utilizam uma ou mais funções específicas, que formariam o núcleo de funcionamento do portal semântico. Estas funções, mostradas na Tabela 6.1 e detalhadas nas próximas seções, utilizam a ontologia ONTO-POINT para realizar suas tarefas.

Serviço do portal	Funções semânticas [60]
Serviço 1: ***"montar perfil de usuário e grupo de usuários"***	*Personalização / Navegação / Colaboração*
Serviço 2: ***"encontrar usuário e grupo de usuários"***	*Personalização / Navegação / Busca*
Serviço 3: ***"permitir colaboração entre os usuários"***	*Personalização / Colaboração / Navegação / Busca*
Serviço 4: ***"publicar conteúdo"***	*Personalização / Publicação / Colaboração*
Serviço 5: ***"recuperar conteúdo"***	*Personalização / Navegação / Busca / Colaboração*
Serviço 6: ***"navegar semanticamente"***	*Navegação Semântica*

Tabela 6.1 – Serviços do portal semântico e funções semânticas utilizadas

[60] A função destacada com sublinhado é a principal para o serviço em questão, enquanto as demais são complementares.

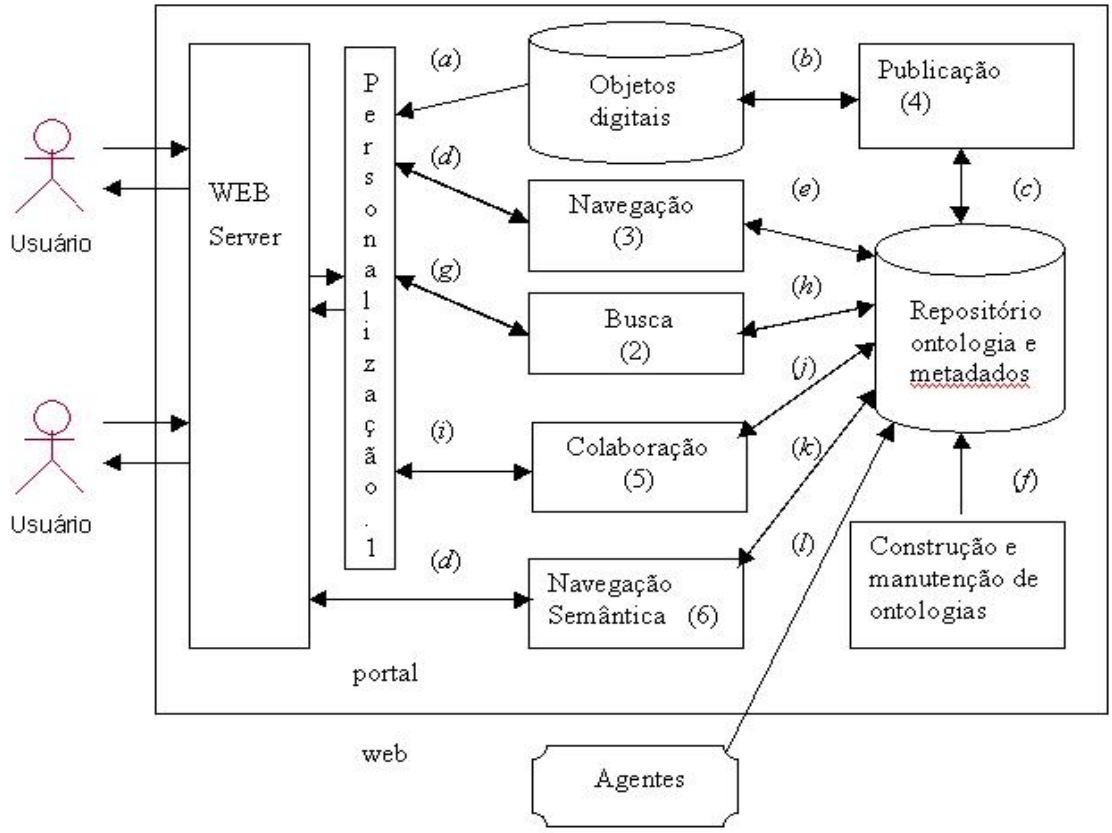


Figura 6.1 – Arquitetura proposta para o portal semântico UNI-POINT

A Figura 6.1 apresenta a arquitetura do portal, onde são encontradas as funções: Personalização (1), Busca (2), Navegação (3), Publicação (4), Colaboração (5) e Navegação Semântica (6).

O conteúdo do portal (*a*), formado por diversos tipos de objetos digitais[61], deverá ser publicado e mantido pelos próprios usuários do portal. Os objetos poderão ser marcados semanticamente por uma ferramenta de marcação semântica (*b*) como por exemplo, SMORE ou Semantic Word, apresentadas no Capítulo 3. O uso de SMORE será exemplificado com a ontologia ONTO-POINT na Seção 6.4 deste capítulo. O armazenamento dos metadados pode ser feito, por exemplo, em um repositório (*c*) do tipo ZEMANTIC, que seja capaz de armazenar triplas RDF, também apresentado no Capítulo 3.

Princípios de gestão de conteúdo (*d*) deverão ser utilizados no gerenciamento dos objetos digitais e na sua recuperação (*e*) através de navegação. Plone, Zope e Espaço Semântico foram apresentados no Capítulo 3 e são utilizadas posteriormente na Seção 6.7.

A ontologia ONTO-POINT deverá ser editada e mantida por uma ferramenta de criação e manutenção de ontologias (*f*). A escolhida foi Protégé 2000 com extensão OWL, utilizada para a construção da ontologia retratada no Capítulo 4. O vocabulário ontológico

deverá ser armazenado no mesmo repositório de metadados, que deverá possuir uma linguagem de consultas semânticas (*h*), do tipo RDQL, que dá suporte a consultas a dados RDF (instâncias) e informações de esquema RDFS (hierarquia de classes, e propriedades). Poderá ser também utilizada uma ferramenta do tipo KIM, que provê uma interface amigável para busca ou do tipo TOSS, também apresentada no Capítulo 3, que permite a expansão dos termos de busca (*g*).

Ferramentas de colaboração (*i*) poderão ser utilizadas, sendo possível que elas estendam as funcionalidades dos aplicativos hoje existentes e passem a interagir com os atores do portal através de consultas à ontologia (*j*) e sugestões de novos atores ou assuntos correlatos, como visto no Capítulo 3, onde foi mostrado um método para identificação de assuntos de uma ferramenta de *chat.*

Uma ferramenta de navegação semântica (*k*) do tipo MAGPIE, deve ser disponibilizada para acesso ao conteúdo anotado do portal. Exemplos são apresentados na Seção 6.6 deste capítulo.

Finalmente, agentes de software (*l*) poderão varrer o repositório de metadados à procura de conteúdo de interesse.

Desta forma, acredita-se que com a arquitetura apresentada na Figura 6.1, o portal facilita a reutilização e o compartilhamento de informações e conhecimento.

6.2) Função de Personalização

O perfil de um usuário contempla as informações básicas de cada ator e também dos assuntos que interessam a ele e os assuntos em que o mesmo se declara especialista, além das informações de quais grupos ele é membro. Também fazem parte do perfil o formato que o mesmo quer dar à sua página principal do portal, com possibilidade de escolher livremente as janelas de informações, serviços e outras funcionalidades disponíveis que lhe interessam.

A Tabela 6.2 mostra parte da ontologia ONTO-POINT projetada para a classe ALUNO. Nota-se os relacionamentos *"estuda_disciplina"* com a classe DISCIPLINA, *"esta_matriculado_curso"*, com a classe CURSO, *"pertence_à_unidade"*, com a classe

[61] Documentos HTML, arquivos em geral, notícias, eventos, imagens etc.

UNIDADE e os relacionamentos "é_especialista_em" e *"se_interessa_por"*, ambos coma classe ASSUNTO. Estas informações, entre outras, compõem o perfil de um ALUNO.

Classe	Relacionamento	Com Classe
ALUNO	*"estuda_disciplina"*	DISCIPLINA
	"esta_matriculado_curso"	CURSO
	"pertence_à_unidade"	UNIDADE
	"é_especialista_em"	ASSUNTO
	"se_interessa_por"	ASSUNTO

Tabela 6.2 – Classe ALUNO e relacionamentos da ontologia, que formam o seu perfil

A criação de instâncias da classe ATOR (superclasse de ALUNO, PROFESSOR, COORDENADOR DE CURSO e PESQUISADOR) representa a criação de novos perfis de usuário no portal. Dentre os atributos da superclasse ATOR que são herdados pela subclasse ALUNO, por exemplo, destacamos "*permite_notificação_publicação"*, *"deseja_ser_notificado"* e *"permite_colaboração"*, visualizados na Tabela 6.3.

Classe	Atributos herdados	Superclasse
ALUNO	*"permite_notificação_publicação"*	ATOR
	"deseja_ser_notificado"	
	"permite_colaboração"	

Tabela 6.3 – Classe ALUNO e atributos herdados da superclasse ATOR, que também formam o seu perfil

A Figura 6.2 retrata a ontologia ONTO-POINT, elaborada com o auxílio da ferramenta Protégé-2000 com extensão OWL e mostra em destaque a classe ALUNO, subclasse da classe ATOR, seus relacionamentos e seus atributos que compõem o perfil de um usuário do tipo aluno.

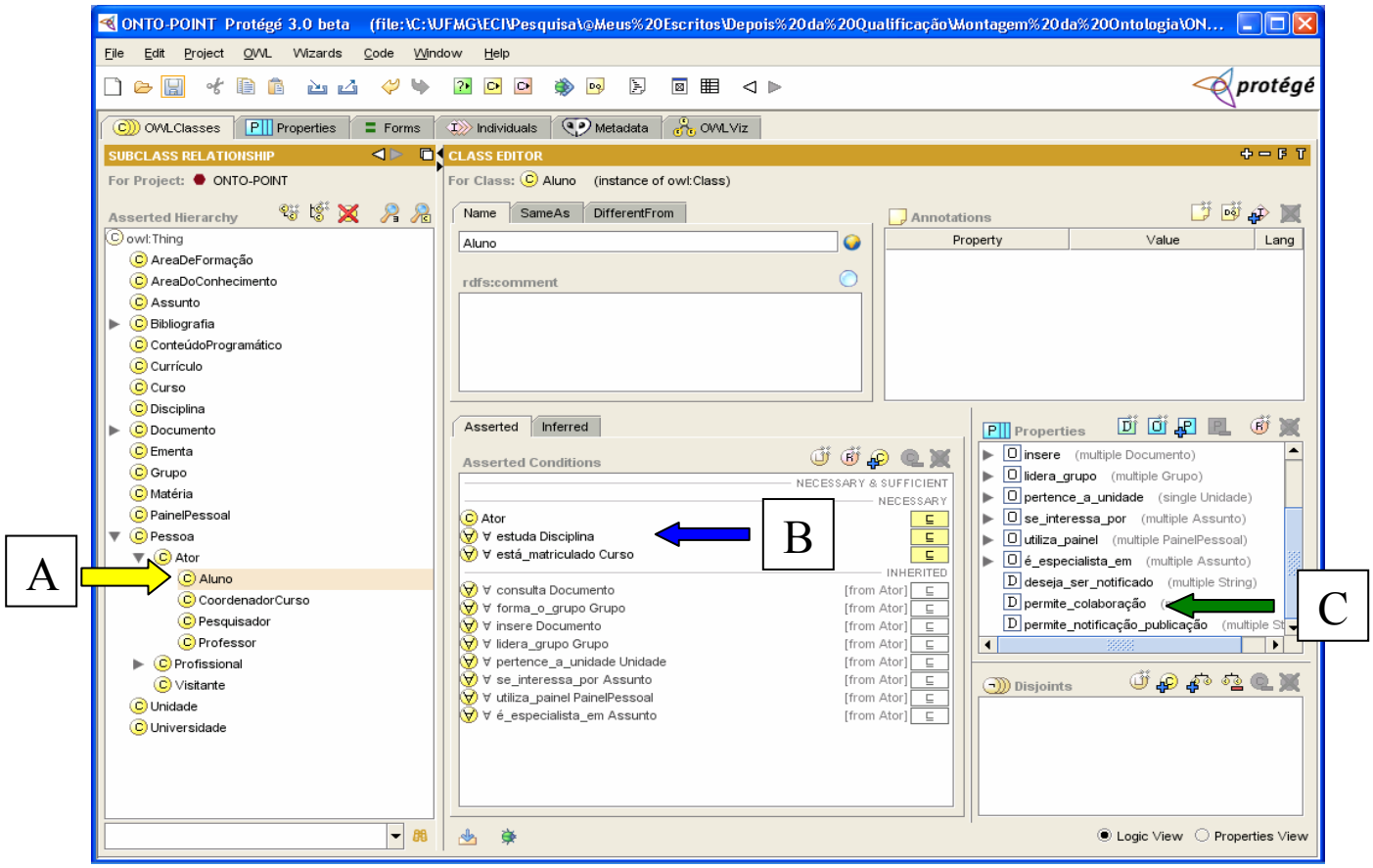


Figura 6.2 – Ontologia ONTO-POINT com detalhe para a classe ALUNO (A), seus relacionamentos (B) e atributos herdados da superclasse ATOR (C)

Também fazem parte do perfil de um usuário os grupos aos quais ele pertence e os assuntos principais que são discutidos em seu âmbito. A Tabela 6.4 mostra parte da ontologia ONTO-POINT projetada para a classe GRUPO. Nota-se o relacionamento *"quer_discutir_assunto"* com a classe ASSUNTO, e os relacionamentos *"é_formado_por"*, e *"tem_como_lider"* com a classe ATOR.

Classe	Relacionamento	Com Classe
GRUPO	*"tem_como_lider"*	ATOR
	"é_formado_por"	ATOR
	"quer_discutir_assunto"	ASSUNTO

Tabela 6.4 – Classe GRUPO e alguns de seus relacionamentos existentes na ontologia ONTO-POINT, que formam o seu perfil

A classe GRUPO possui atributos que compõem o seu perfil, dentre os quais se destaca *"permite_notificação_publicação_dos_membros"*, *"grupo_deseja_ser_notificado"* e *"permite_colaboração_entre_grupos"*, que podem ser visualizadas na Tabela 6.5.

Classe	Atributos
GRUPO	*"permite_notificação_publicação_dos_membros"* *"grupo_deseja_ser_notificado"* *"permite_colaboração_entre_membros"*

Tabela 6.5 – Classe GRUPO e alguns de seus atributos que também formam o perfil

A Figura 6.3 mostra a ontologia ONTO-POINT com destaque para a classe GRUPO, seus relacionamentos e seus atributos, que compõem o perfil de um usuário.

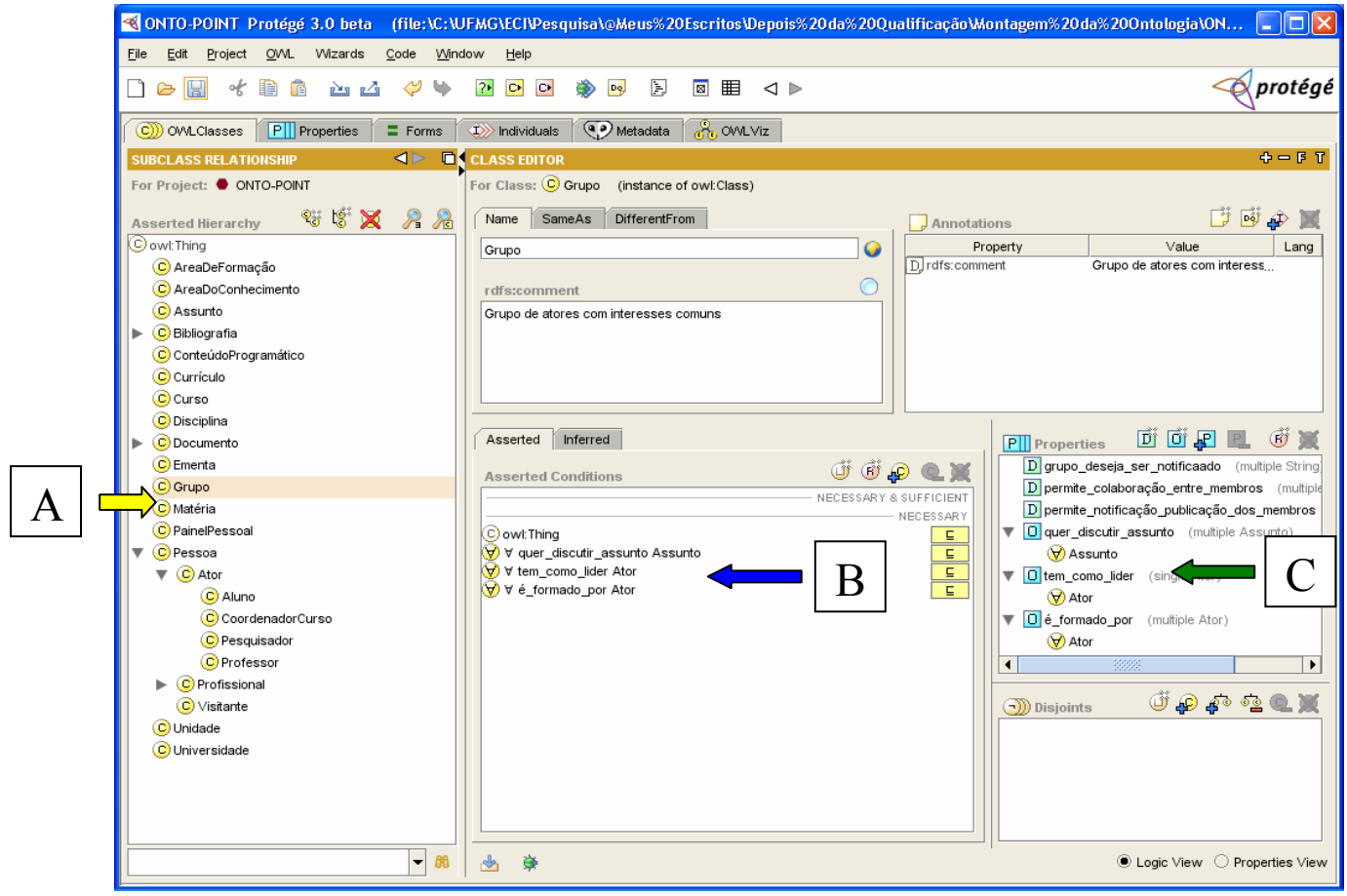


Figura 6.3 – Ontologia ONTO-POINT com detalhe para a classe GRUPO (A), seus relacionamentos (B) e seus atributos (C)

A função de personalização, assim como as demais funções semânticas propostas neste estudo, podem se beneficiar das propriedades das classes OWL da ontologia construída para a realização de inferências entre as instâncias de classe, conforme pode ser visualizado nos exemplos apresentados a seguir.

Exemplo 1: Inferência utilizando uma propriedade inversa: na montagem do perfil de cada aluno, registra-se na propriedade *"estuda_disciplina"* o nome das disciplinas às quais ele assiste às aulas, conforme pode ser visto na Tabela 6.6. Desta forma, pode-se, através de inferência e da propriedade inversa *"é_estudada_por"* da ontologia ONTO-POINT, chegar à conclusão de quais alunos estudam uma certa disciplina, conforme mostrado na Tabela 6.7. Baseado no exemplo apresentado, conclui-se que, se João Carlos e Adriano Sampaio estudam a disciplina "Anatomia Humana", a disciplina "Anatomia Humana" é estudada por João Carlos e Adriano Sampaio.

Instância da Classe ALUNO	**Propriedade**	**Instância da Classe DISCIPLINA**
"João Carlos"	*"estuda_disciplina"*	"Anatomia Humana"
"Silvana Márcia"		"Fisiologia"
"Adriano Sampaio"		"Anatomia Humana"

Tabela 6.6 – Relacionamento existente entre as classes ALUNO e DISCIPLINA, registrando instâncias para a propriedade *"estuda_disciplina"*

Instância da Classe DISCIPLINA	**Propriedade Inversa**	**Instância da Classe ALUNO**
"Anatomia Humana"	*"é_estudada_por"*	"João Carlos"
"Anatomia Humana"		"Adriano Sampaio"
"Fisiologia"		"Silvana Márcia"

Tabela 6.7 – Realização de inferência entre as classes ALUNO E DISCPLINA, através da utilização da propriedade inversa *"é_estudada_por"*

Exemplo 2*:* Inferência utilizando uma propriedade funcional: a Tabela 6.8 mostra um exemplo de relacionamento entre as classes DISCIPLINA e PROFESSOR. A propriedade *"tem_professor_titular"* é funcional, ou seja, cada disciplina deve ter um e somente um professor titular. No exemplo apresentado, pode-se concluir através de inferência que, se a disciplina de "Fisiologia" tem como professor titular "Kênia" e em outro registro, a "Dra. Martins Pedrosa", os dois indivíduos correspondem à mesma pessoa, "Dra. Kênia Martins Pedrosa".

Instância da Classe DISCIPLINA	**Propriedade Funcional**	**Instância da Classe PROFESSOR**
"Anatomia Humana"	*"tem_professor_titular"*	"Jairo Flávio"
"Fisiologia"		"Kênia"
"Fisiologia"		"Dra. Martins Pedrosa"

Tabela 6.8 – Utilização da propriedade funcional *"tem_professor_titular"* que permite concluir por inferência que Kênia e Dra. Martins Pedrosa são a mesma pessoa

Exemplo 3*:* Inferência utilizando uma propriedade funcional inversa: a Tabela 6.9 mostra um relacionamento entre as classes COORDENADOR CURSO e CURSO. A propriedade *"coordena_curso"* é funcional inversa, ou seja, um coordenador até pode coordenar mais de um curso, mas cada curso só pode ser coordenado por um único coordenador. Neste exemplo, pode-se concluir através de inferência que, se o curso de "Administração" tem como coordenador "Dilson" e em outro registro, o "Sr. Campos Júnior", os dois indivíduos correspondem à mesma pessoa, "Sr. Dílson Campos Júnior".

Instância da Classe COORDENADOR CURSO	**Propriedade Funcional Inversa**	**Instância da Classe CURSO**
"Dilson"	*"coordena_curso"*	"Administração"
"Érika"		"Terapia Ocupacional"
"Sr. Campos Júnior"		"Administração"

Tabela 6.9 – Utilização da propriedade funcional inversa *"coordena_curso"* que permite concluir por inferência que Dilson e Sr. Campos Júnior são a mesma pessoa

Exemplo4: Inferência utilizando uma propriedade transitiva: a Tabela 6.10 mostra um auto-relacionamento da classe DISCIPLINA, sendo que a propriedade *"faz_parte"* é declarada como sendo transitiva. Neste exemplo, pode-se concluir através de inferência que, se a disciplina "Bibliotecas Digitais" faz parte da disciplina "Gestão de Conteúdo" e se a disciplina "Gestão de Conteúdo" faz parte da disciplina "Gestão do Conhecimento", então a disciplina "Bibliotecas Digitais" faz parte da disciplina "Gestão do Conhecimento".

Instância da Classe DISCIPLINA	**Propriedade Transitiva**	**Instância da Classe DISCIPLINA**
"Bibliotecas Digitais"		"Gestão de Conteúdo"
"Gestão de Conteúdo"	*"faz_parte"*	"Gestão do Conhecimento"

Tabela 6.10 – Utilização da propriedade transitiva *"faz_parte"* que permite concluir por inferência que "Bibliotecas Digitais" faz parte da disciplina "Gestão do Conhecimento"

Exemplo5: Inferência utilizando uma propriedade simétrica: a Tabela 6.11 mostra um auto-relacionamento da classe ATOR, sendo que a propriedade *"é_irmão_de"* é declarada como sendo simétrica. Neste exemplo, pode-se concluir através de inferência que, se "Pedro de Alcântara" é irmão de "Simone de Alcântara", então "Simone de Alcântara" também é irmã de "Pedro de Alcântara".

Instância da Classe ATOR	**Propriedade Simétrica**	**Instância da Classe ATOR**
"Pedro de Alcântara"	*"é_irmão_de"*	"Simone de Alcântara"

Tabela 6.11 – Utilização da propriedade simétrica "é_irmão_de" que permite concluir por inferência que "Pedro de Alcantara" e "Simone de Alcantara" são irmãos

6.3) Função de Busca

A função de busca é utilizada principalmente pelos serviços que têm necessidade de encontrar, através de uma pesquisa dentro do portal semântico, usuários, grupos ou conteúdo de seu interesse.

A função de busca pode se beneficiar da organização dos conceitos existente na ontologia construída para tornar mais efetivo o processo de pesquisa no repositório do portal semântico. Além disso, a expansão de termos de pesquisa e a utilização da linguagem de consultas RDQL são estratégias que podem fazer com que a pesquisa através do portal semântico seja mais precisa do que em uma pesquisa tradicional, conforme pode ser visto nos exemplos que se seguem.

Exemplo 1: Deseja-se encontrar grupos de usuários que se interessam em discutir o assunto "Requisitos de Software". O simples fato de os conceitos estarem associados e relacionados entre si na ontologia favorece a busca de usuários ou grupo de usuários dentro do portal. Na Tabela 6.12 estão listadas as classes da ontologia ONTO-POINT envolvidas no problema apresentado e a Tabela 6.13 lista as instâncias de classes que poderiam ser resultado desta pesquisa.

Classe	Relacionamento	Com Classe
GRUPO	*"quer_discutir assunto"*	ASSUNTO

Tabela 6.12 – Classes envolvidas no problema "grupos que se interessam em discutir o assunto 'Requisitos de Software'"

Instância da classe GRUPO	Relacionamento	Instância da classe ASSUNTO
Analistas_de_Negócios	*"quer_discutir assunto"*	Requisitos_de_Software
Pesquisadores_ES_01	*"quer_discutir assunto"*	Requisitos_de_Software

Tabela 6.13 – Instâncias de classes resultado da pesquisa "grupos que se interessam em discutir o assunto 'Requisitos de Software'"

Pode-se concluir então que os grupos *"Analistas_de_Negócios"* e *"Pesquisadores_ES_01"* estão interessados em discutir o assunto "Requisitos de Software".

Beneficiando-se da ontologia construída e da organização existente entre os conceitos, pode-se desdobrar o problema e obter um resultado mais preciso de pesquisa. A Tabela 6.14 mostra as novas classes envolvidas no problema.

Classe	Relacionamento	Com Classe
GRUPO	*"quer_discutir assunto"*	ASSUNTO
ASSUNTO	*"pode_ser_disciplina"*	DISCIPLINA
DISCIPLINA	*"é_dividida_em_disciplina"*	DISCIPLINA

Tabela 6.14 – Novas classes envolvidas no problema "grupos que se interessam em discutir o assunto Requisitos de Software'"

As Tabelas 6.15 a, 6.15 b e 6.15 c a seguir mostram as novas instâncias de classes que podem ser resultado da nova pesquisa.

a)

Instância da classe DISCIPLINA	Relacionamento	Instância da Classe DISCIPLINA
Engenharia Software – ES01	*"é_dividida_em_disciplina"*	Requisitos Software – RS01
Engenharia Software – ES01	*"é_dividida_em_disciplina"*	Projeto Software – PS01
Engenharia Software – ES01	*"é_dividida_em_disciplina"*	Qualidade Software – QS01

b)

Instância da classe ASSUNTO	Relacionamento	Instância da Classe DISCIPLINA
Engenharia de Software	*"pode_ser_disciplina"*	Engenharia Software – ES01
Requisitos de Software	*"pode_ser_disciplina"*	Requisitos Software – RS01
Projeto de Software	*"pode_ser_disciplina"*	Projeto Software – PS01
Qualidade de Software	*"pode_ser_disciplina"*	Qualidade Software – QS01

c)

Instância da classe GRUPO	Relacionamento	Instância da Classe ASSUNTO
Analistas_de_Negócios	*"quer_discutir assunto"*	Requisitos de Software
Pesquisadores_ES_01	*"quer_discutir assunto"*	Requisitos de Software
Projetistas_de_Sistemas	*"quer_discutir assunto"*	Projeto de Software
Turma_da_Qualidade	*"quer_discutir assunto"*	Qualidade de Software
Engenheiros_de_Software	*"quer_discutir assunto"*	Engenharia de Software

Tabelas 6.15 – Novas instâncias de classes resultado da pesquisa "grupos que interessam em discutir o assunto 'Requisitos de Software'"

Já que a disciplina "Engenharia de Software" é dividida em "Requisitos de Software", "Projetos de Software" e "Qualidade de Software" (Tabela 6.15-a) e todas estas disciplinas são assuntos possíveis de serem discutidos por diversos grupos (Tabela 6.15-b), pode-se concluir que todos os grupos que se interessam em discutir um assunto macro chamado "Engenharia de Software" também estão interessados no assunto "Requisitos de Software". Desta forma, analisando as instâncias apresentadas na Tabela 6.15-c, concluímos que, tanto os grupos *"Analistas_de_Negócios"* e *"Pesquisadores_ES_01"* (ambos listados como resultado da pesquisa original anterior) quanto o grupo, *"Engenheiros_de_Software"* (que aparece apenas nesta fase de desdobramento da pesquisa) se interessam em discutir o assunto "Requisitos de Software".

Exemplo 2: Deseja-se saber quem são os coordenadores dos cursos que têm a disciplina "Gestão da Informação" em seu currículo, abrangendo todas as unidades da UNIPAC. A Tabela 6.16 apresenta um exemplo de consulta para busca que poderia ser realizada para a obtenção de resultados que atendam a necessidade do usuário do portal.

Necessidade de Busca	Consulta para Busca
Deseja-se encontrar todos os coordenadores de curso, das várias unidades, cujo curso tem a disciplina "Gestão da Informação"	Select X From COORDENADORDECURSO{X} coordena CURSO {Y} possui DISCIPLINA{Z} Where Z= "Gestão da Informação"

Tabela 6.16 – Exemplo de busca para se encontrar coordenadores de curso das unidades que têm curso com a disciplina "Gestão da Informação"

No ambiente do portal UNI-POINT, o conteúdo é recuperado de acordo com a ontologia ONTO-POINT, levando em conta sua contextualização semântica. Uma das características da função de busca semântica é a expansão dos termos de busca apresentados pelo usuário a partir dos conceitos e propriedades da ontologia de domínio armazenada no portal. A idéia principal é descobrir, através da ontologia, outros conceitos relacionados e montar uma nova seqüência de palavas-chave de busca utilizando os operadores booleanos "AND" e "OR". Propriedades simétricas, transitivas e inversas também são utilizadas no sentido de expandir os termos para a busca. Entretanto, quando o termo a ser usado na pesquisa não fizer parte da ontologia, a busca se dará nos moldes tradicionais, sem o benefício advindo da estratégia de expansão dos termos. Esta estratégia cria uma nova seqüência de termos, mas continua utilizando-se de um mecanismo de busca existente para realizar efetivamente a pesquisa.

Dentro da proposta do portal UNI-POINT e utilização da ontologia ONTO-POINT, tem-se a declaração de equivalência entre o conceito DISCIPLINA e outros dois conceitos, MATÉRIA e ÁREADOCONHECIMENTO. Desta forma, qualquer pesquisa que contenha um dos termos citados acima pode ser expandida aos outros dois utilizando-se o operador OR. Na Tabela 6.17, vê-se um novo exemplo em que se compara a consulta apresentada inicialmente com uma nova consulta, agora apoiada pela noção de sinonímia da ontologia. Nesta consulta, novamente deseja-se saber quem são os coordenadores de curso, cujo curso tem a disciplina "Gestão da Informação", abrangendo todas as unidades da UNIPAC.

Consulta original	Consulta expandida
Select X From COORDENADORDECURSO{X} coordena CURSO {Y} possui DISCIPLINA{Z} Where Z= "Gestão da Informação"	Select X From COORDENADORDECURSO{X} coordena CURSO {Y} possui DISCIPLINA{Z} **OR possui MATÉRIA {Z} or possui AREADOCONHECIMENTO {Z}** Where Z= "Gestão da Informação"

Tabela 6.17 – Consulta original e consulta expandida auxiliada pela Ontologia

Exemplo 3: Mais um exemplo de busca apoiada pela ontologia, desta vez utilizando-se da linguagem de consultas RDQL, descrita no capítulo 2, pode ser visto na Figura 6.4. Esta funcionalidade só é possível para bases de dados do tipo RDF, típicas da Web Semântica e do portal semântico ora proposto. No caso, é necessário saber quais são todos os documentos publicados por membros dos grupos que discutem o assunto "Inteligência Artificial". Os possíveis resultados são mostrados na Figura 6.5.

```
SELECT ?titulo_doc, ?ator  ?grupo
WHERE (?titulo_doc, ontopoint:INSERIDOPOR, ?ator)
              (?ator, ontopoint:FORMAOGRUPO, ?grupo)
          (?grupo, ontopoint:QUERDISCUTIRASSUNTO, "Inteligência
Artificial")
 USING ontopoint FOR <http://www.unipac.br/ontopoint/ontopoint-rdf#>
```

Figura 6.4 – Exemplo de busca utilizando a linguagem RDQL para "documentos publicados por membros dos grupos que discutem o assunto Inteligência Artificial"

titulo_doc	ator	grupo
"Uso da IA na Web"	"João Carlos Alves"	Grupo da IA
"Inteligência Artificial nos dias atuais"	"Juliana Lima"	IA-001
"Web Semântica e IA"	"Sara Lucas Passos"	IA-001
"Inteligência quase natural"	"Mateus de Oliveira"	Grupo da IA

Figura 6.5 – Resultado possível da busca para "documentos publicados por membros dos grupos que discutem o assunto Inteligência Artificial"

6.4) Função de Publicação

A função de publicação utiliza ferramentas de Gestão de Conteúdo para o armazenamento de documentos e uma ferramenta de anotação que, utilizando-se da ontologia ONTO-POINT, permite a marcação semântica de seu conteúdo.

Ao considerar que determinado documento deve ser compartilhado com a comunidade, o usuário publica o mesmo no portal. É então permitido ao usuário associar metadados ao documento, ou seja, fazer anotações com termos da ontologia, conforme pode ser visto no canto inferior esquerdo (B) da Figura 6.6.

A Figura 6.6 mostra a ferramenta de Gestão de Conteúdo Plone, apresentado no Capítulo 3, onde se destaca a existência de vários documentos armazenados no portal (A).

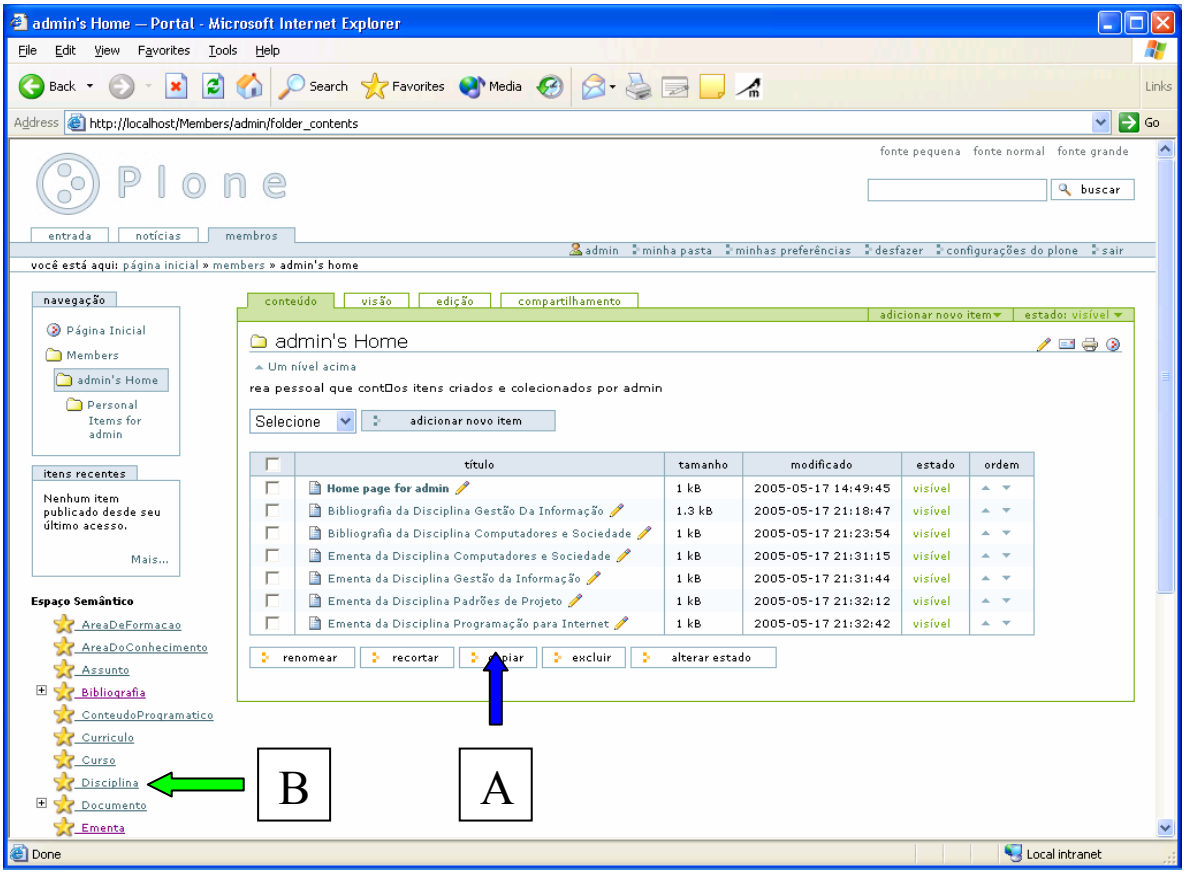


Figura 6.6 – Sistema de Gestão de Conteúdo Plone com documentos armazenados (A) e a ontologia ONTO-POINT (B).

Destaca-se, dentro da ontologia ONTO-POINT, a classe DOCUMENTO com seus atributos *título*, *subtítulo, data_inserção, URI, data_validade e descrição* entre outros. A classe DOCUMENTO, que tem como subclasses ANAIS, LIVRO, TESE, PUBLICAÇÃO-DIVERSA, PUBLICAÇÃO-PERIÓDICA e RELATÓRIO-TÉCNICO, pode ser visualizada na Tabela 6.18, que mostra os relacionamentos *"inserido_por"* e *"consultado_por"* com a classe ATOR, *"diz_respeito_a"* com a classe ASSUNTO, *"tem_autor"* com a classe AUTOR, *"tem_organizador"* com a classe ORGANIZADOR, *"tem_tradutor"* com a classe TRADUTOR e *"tem_publicador"* com a classe PUBLICADOR.

Classe	Relacionamento	Com Classe
DOCUMENTO	*"inserido_por"*	ATOR
	"diz_respeito_a"	ASSUNTO
	"consultado_por"	ATOR
	"tem_autor"	AUTOR
	"tem_organizador"	ORGANIZADOR
	"tem_tradutor"	TRADUTOR
	"tem_publicador"	PUBLICADOR

Tabela 6.18 – Classe DOCUMENTO e alguns de seus relacionamentos existentes na Ontologia ONTO-POINT

No exemplo que se segue, é mostrada a publicação de um documento do tipo Bibliografia de uma Disciplina, utilizando-se da ferramenta Espaço Semântico, também apresentada no Capítulo 3 deste trabalho, dentro do ambiente do Plone. A Figura 6.7 mostra na parte central, o documento a ser publicado (A) e no canto inferior esquerdo (B), a ontologia ONTO-POINT, que irá permitir a colocação de metadados ao referido documento.

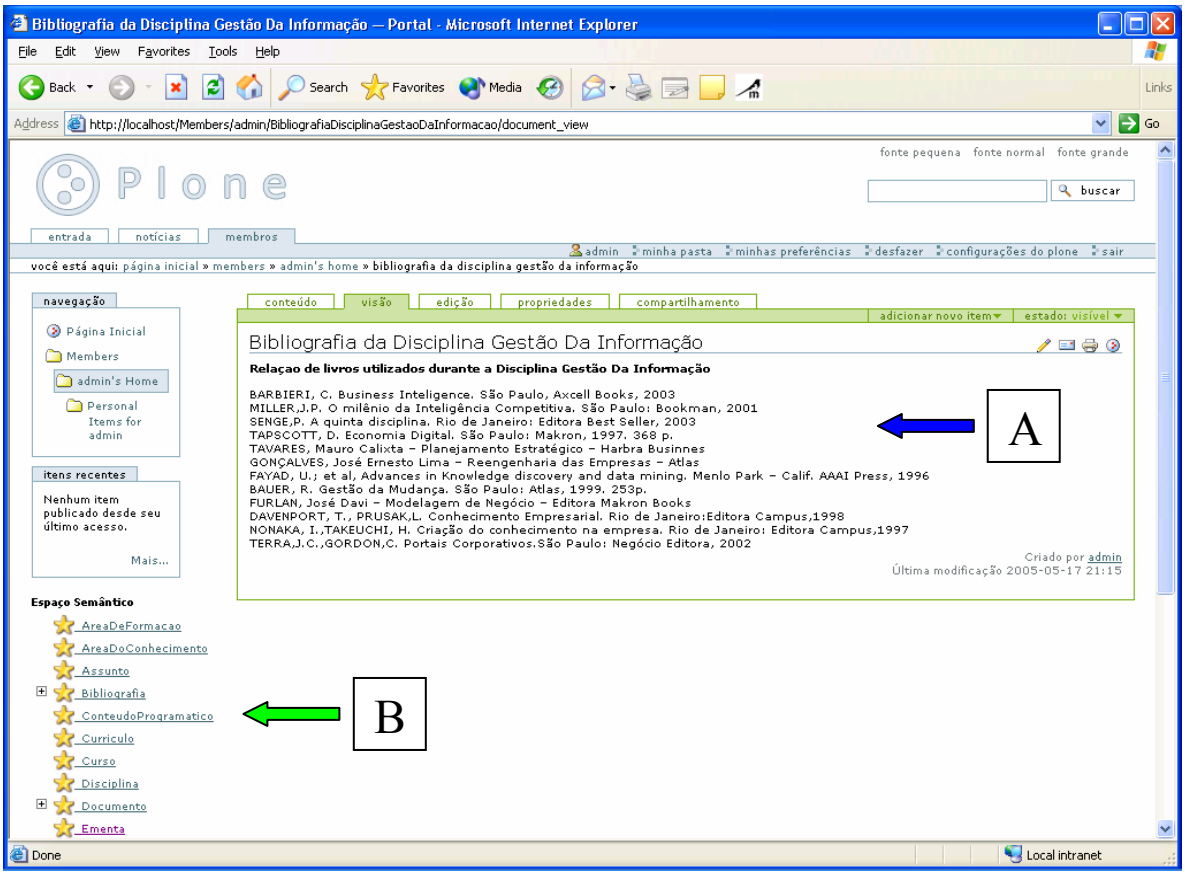


Figura 6.7 – Documento a ser publicado no portal (A) e a ontologia ONTO-POINT

A Figura 6.8 mostra a etapa seguinte de publicação do documento no portal. A ontologia é chamada e exibida em uma janela no centro da tela (A) e palavras-chave são escolhidas a partir da ontologia para serem então associadas ao documento (B). No caso apresentado, os conceitos Livro e Bibliografia foram escolhidos para serem ligados ao documento em questão.

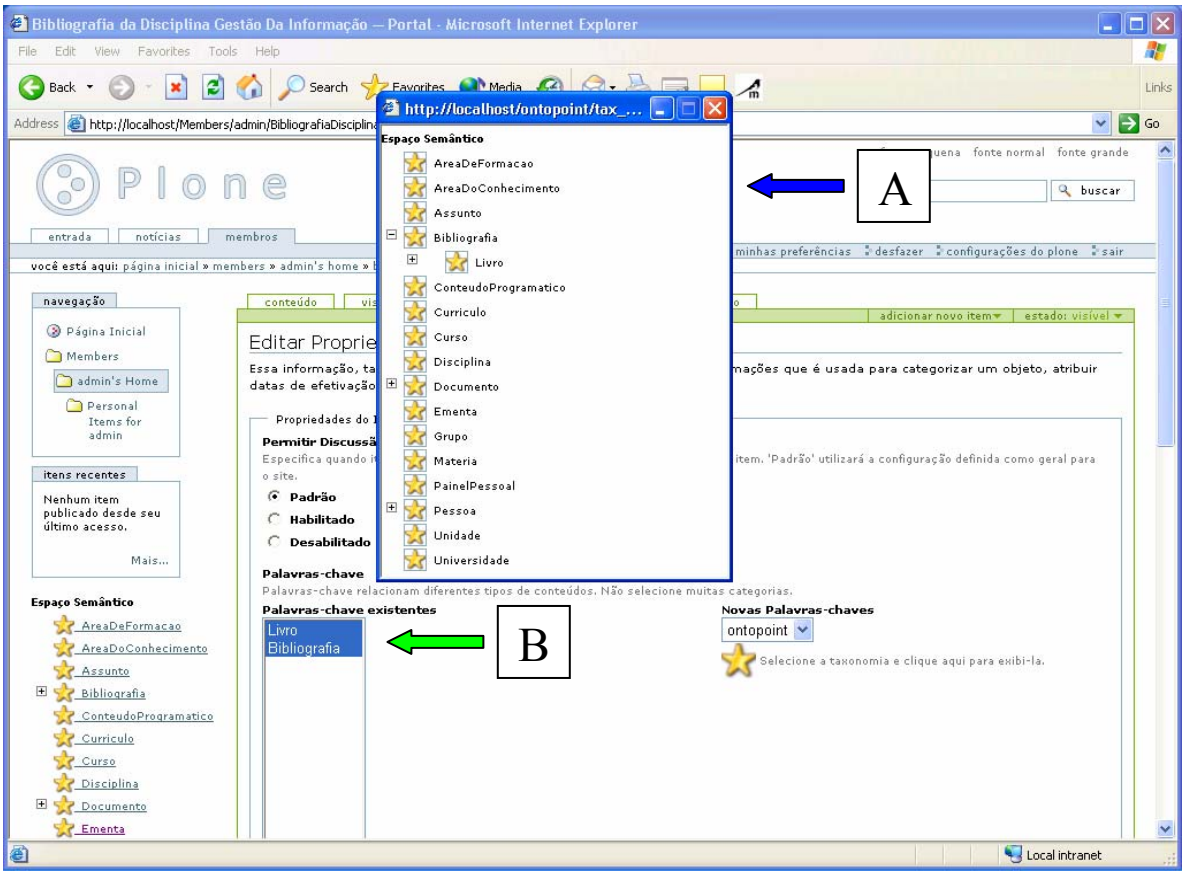


Figura 6.8 – Ontologia ONTO-POINT no centro da tela (A) com seus termos associados ao documento a ser publicado (B).

A Figura 6.9 mostra o documento publicado anteriormente, agora sendo apresentado como resultado de uma navegação (A), momento em que foram solicitados todos os documentos armazenados com palavra-chave igual a Bibliografia (B).

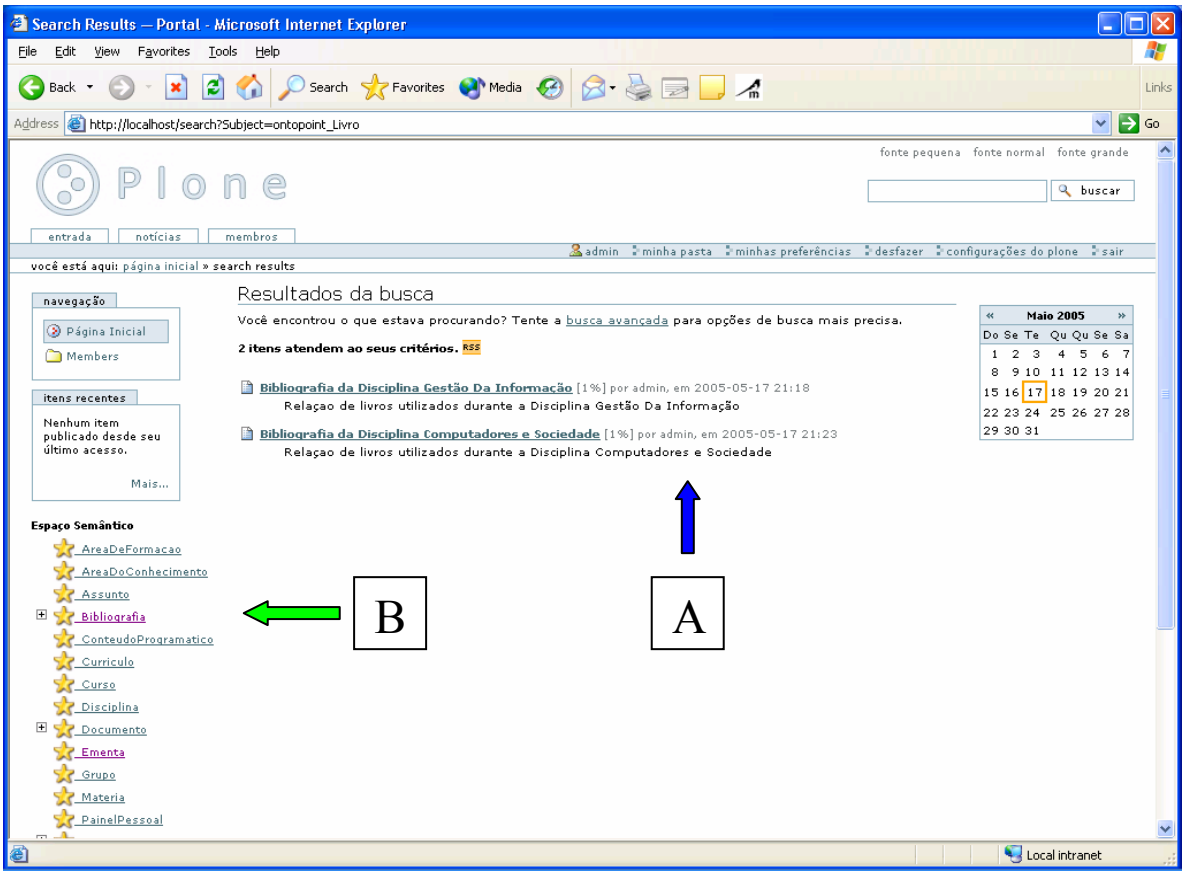


Figura 6.9 – Documento apresentado como resultado da navegação (A) através do conceito Bibliografia (B) escolhido a partir da ontologia ONTO-POINT

Outra etapa dentro da função de publicação permite que o usuário marque, com o apoio da ontologia, todos os conceitos que julgar importantes, tornando-os instâncias das classes existentes na ontologia. Desta forma, o documento poderá ser recuperado não só pelos metadados relativos ao documento propriamente dito, mas também por todas as ligações semânticas inseridas no momento de sua publicação. Assim, à medida que cada documento é armazenado e seu conteúdo é marcado semanticamente, passa a fazer parte da base de conhecimento do portal semântico, podendo ser recuperado a qualquer momento por outro usuário, tanto através de sua anotação semântica quanto através de todos os conceitos a que ele estiver ligado.

No exemplo que se segue, é mostrado o processo de criação de instância de classe e população da ontologia a partir de um documento que mostra informações a respeito do curso de Sistemas de Informação da unidade de Conselheiro Lafaiete da UNIPAC.

Na Figura 6.10, extraída do aplicativo SMORE, apresentado no Capítulo 3 deste estudo, o usuário destacou no documento, mostrado no canto superior direito do aplicativo, o nome do coordenador de curso (A) e arrastou-o em direção a classe COORDENADORCURSO (B), na janela da ontologia que fica no canto superior esquerdo. Imediatamente o sistema perguntou se é desejo do usuário a criação de uma instância para a classe selecionada (C).

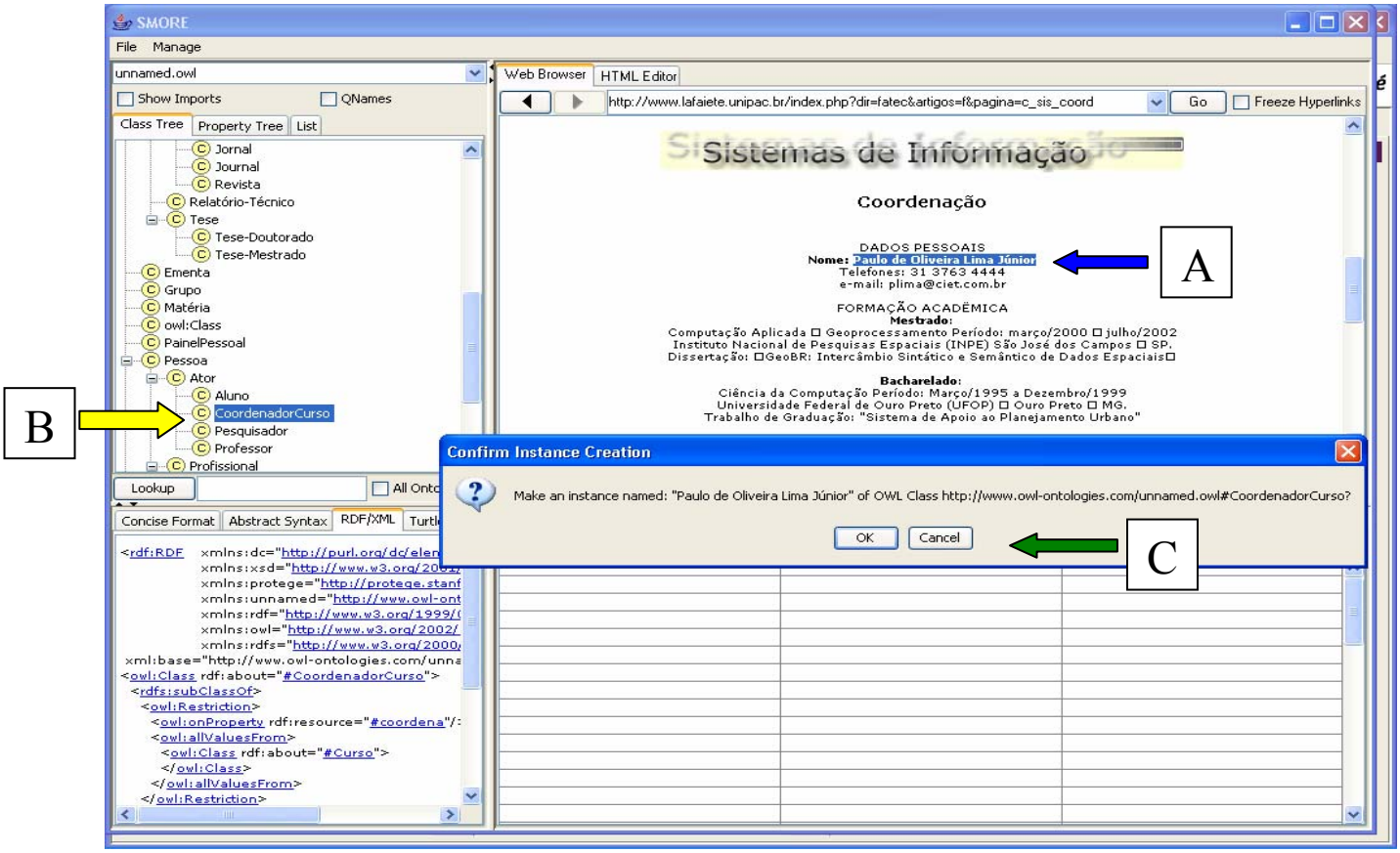


Figura 6.10 – Marcação Semântica de documento dentro da ferramenta SMORE, com destaque para criação de instância da classe COORDENADORCURSO

Na Figura 6.11 é mostrada a marcação semântica do conteúdo em estado mais avançado, com instâncias já criadas para as classes CURSO E DISCIPLINA, além da classe COORDENADORCURSO (B). Na mesma tela é apresentada, no canto inferior direito, as triplas RDF montadas seguindo o formato de sujeito, predicado e objeto (A), que serão armazenadas no repositório de metadados para posterior recuperação e navegação semântica.

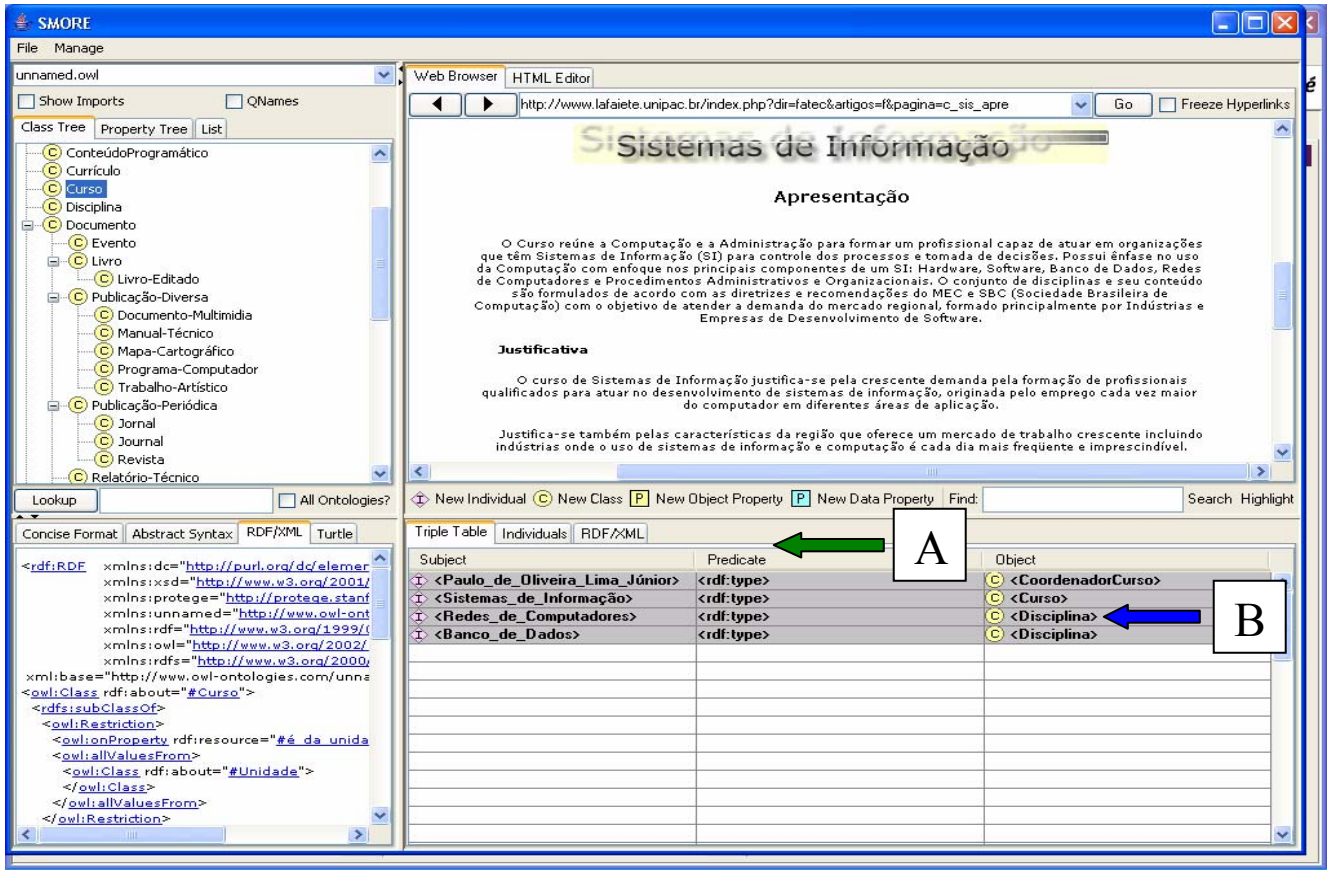


Figura 6.11 – Marcação Semântica de documento com montagem de triplas RDF (A) e com instâncias criadas para as classes (B)

Na Figura 6.12, ainda dentro do aplicativo SMORE, pode-se observar no canto inferior direito parte da ontologia montada na linguagem RDF/XML, indicando as instâncias das classes COORDENADORCURSO, CURSO E DISCIPLINA (A). No canto inferior esquerdo do aplicativo (B), podem ser visualizadas as propriedades e relacionamentos da classe em destaque, no caso a classe DISCIPLINA.

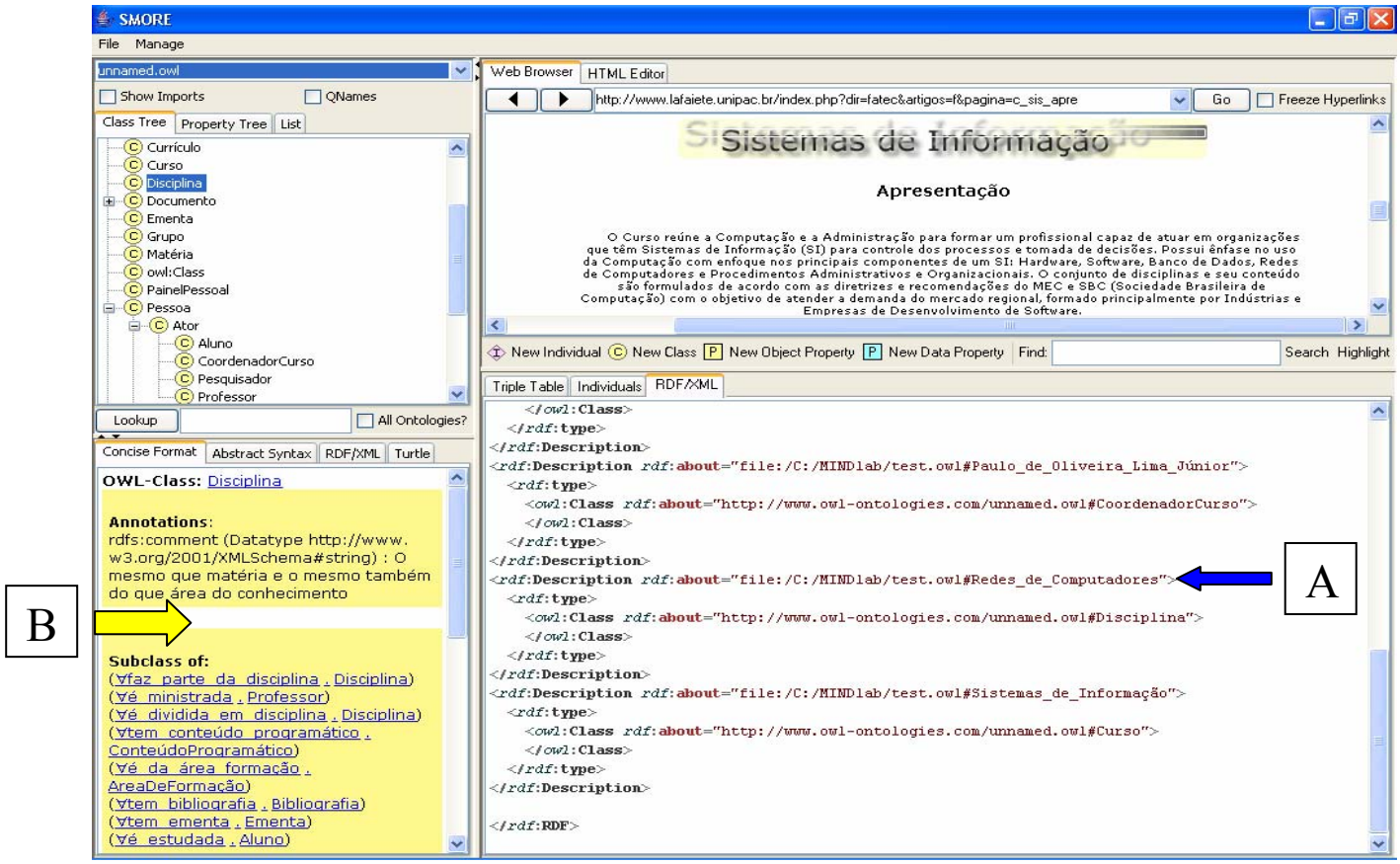


Figura 6.12 – Instâncias de classes representadas em linguagem RDF/XML (A) e relacionamentos da classe DISCIPLINA (B)

6.5) Função de Colaboração

A função de colaboração pode ser ativada por diversos serviços, quando for necessário enviar notificações a usuários ou grupos de usuários, ou ser executada isoladamente, toda vez que algum ator desejar realizar algum tipo de comunicação com outros atores do portal.

Muitas soluções de portal já vêm ou se integram facilmente com um conjunto de ferramentas de comunicação e diálogo baseadas na Web, propiciando oportunidades para que os indivíduos se encontrem, compartilhem informações e opiniões, façam apresentações ou até mesmo tomem decisões colaborativas em tempo real.

Existem dois tipos de ferramentas de colaboração: as ferramentas síncronas e as ferramentas assíncronas. Ferramentas síncronas, como por exemplo, sistemas eletrônicos de reuniões, videoconferência e ferramentas de *chat*, permitem que duas ou mais pessoas

trabalhem em conjunto ao mesmo tempo, independentemente de estarem juntas no mesmo lugar ou em lugares diferentes. As ferramentas assíncronas permitem que as pessoas trabalhem juntas em momentos diferentes. Com elas, não importa quando ou onde as pessoas estão trabalhando, pois substituem encontros e podem até torná-los desnecessários. São exemplos de ferramentas assíncronas o e-mail, fóruns de discussão e ferramentas de workflow[62].

No contexto do portal semântico UNI-POINT, algumas das ferramentas citadas acima podem ser utilizadas pelas funções de publicação e recuperação, de forma a possibilitar a notificação de usuários ou grupos de usuários toda vez que um novo conteúdo de seu interesse for publicado por outro usuário do portal. O usuário publicador também poderá receber notificação, por exemplo, toda vez que um documento postado por ele for consultado por alguma outra pessoa que tenha interesse em seu conteúdo.

A função de colaboração também apóia a formação de grupos. Auxiliada pela ontologia ONTO-POINT e pelos perfis de usuário existentes dentro do portal semântico, ela é responsável pelo disparo de uma mensagem a todos os membros de um grupo quando da filiação de um novo usuário, por exemplo.

6.6) Função de Navegação Semântica

A função de navegação semântica proposta deve adotar recursos de navegação em grafos, onde, o usuário tem condições de navegar para qualquer nodo, e situar-se facilmente sobre sua posição.

Através da ferramenta MagPie apresentada no Capítulo 3, mostra-se um exemplo de navegação semântica na Figura 6.13, onde um documento marcado e armazenado no portal propicia ao usuário acesso a informações relacionadas. As classes UNIDADE e CURSO da ontologia estão presentes na barra de ferramentas do aplicativo, destacadas com as cores verde e laranja, respectivamente (A). Ao se clicar o botão de comando verde intitulado Unidade, todas as instâncias de classe UNIDADE que representam conceitos marcados na ontologia (no exemplo, as unidades de Conselheiro Lafaiete e Barbacena) são coloridas da mesma cor verde (B). De igual forma, ao se clicar o botão de comando laranja intitulado

[62] qualquer tarefa executada em série ou em paralelo por dois ou mais membros de um grupo de trabalho (workgroup) visando a um objetivo comum

Curso, todas as instâncias da classe CURSO presentes no documento são coloridos da mesma cor laranja (C), de forma a se destacar as instâncias das classes da ontologia que estão presentes no texto.

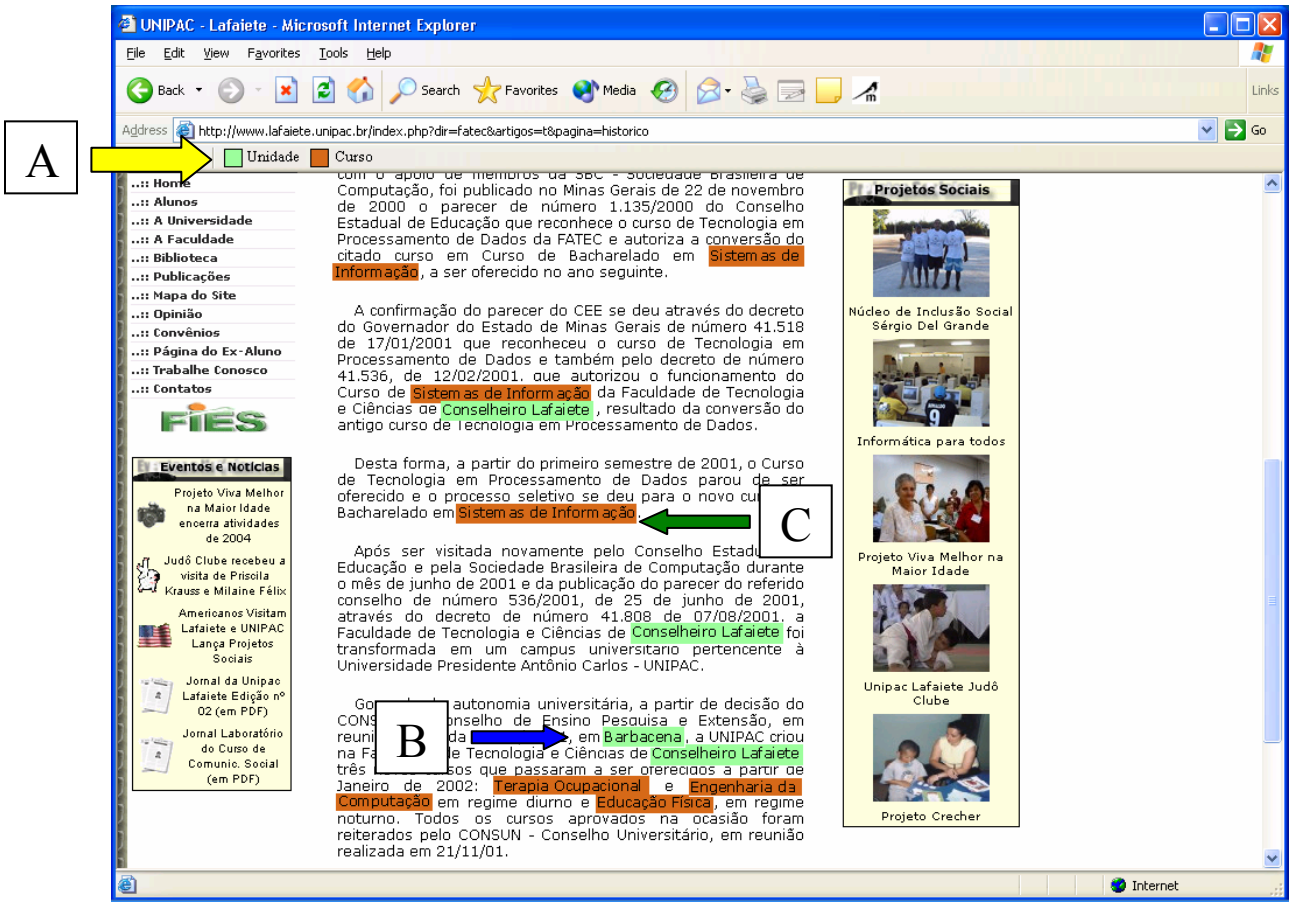


Figura 6.13 – Classes da ontologia na barra de ferramentas (A) e instâncias marcadas semanticamente para as classes Unidade (B) e Curso (C)

O relacionamento existente entre os termos coloridos e as instâncias da ontologia é utilizado para chamar serviços Web que realizam processamentos particulares. Desta forma, para cada termo marcado, o usuário pode solicitar[63] um menu dinâmico[64], que permite o acesso a informações detalhadas sobre a instância selecionada, conforme pode ser visto na Figura 6.14. Este menu apresenta as propriedades e relacionamentos da classe existente na ontologia (A), estendendo a possibilidade de navegação e aquisição de informações.

[63] através do botão direito do *mouse*.
[64] do termo em inglês menu *pop-up*

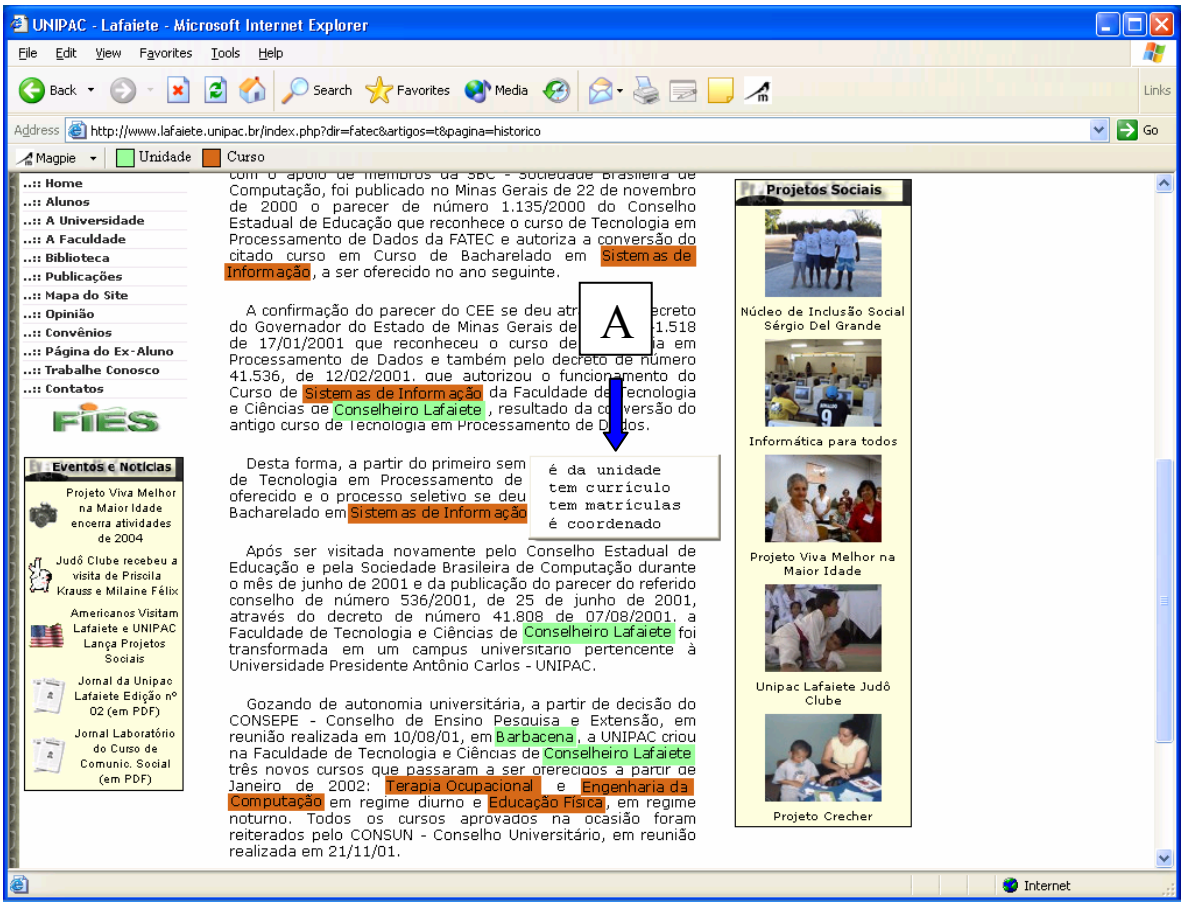


Figura 6.14 – Informações adicionais em menu dinâmico propostas para a classe Curso

6.7) Função de Navegação Taxonômica

Esta função permite a navegação na taxonomia de classes da ontologia e se beneficia do fato de que as classes estão organizadas em uma hierarquia e estão relacionadas a outras através de associações do tipo generalização-especialização, o que favorece o processo de pesquisa, como nos exemplos apresentados a seguir.

Exemplo 1: Na Tabela 6.19 pode-se ver parte da ontologia ONTO-POINT que mostra a classe PESQUISADOR e seus relacionamentos *"pertence_a_unidade"* com a classe UNIDADE e os relacionamentos *"se_interessa_por"* e *"é_especialista_em"* com a classe ASSUNTO. Através destes relacionamentos, pesquisadores de unidades distintas podem se conhecer e entrar em contato uns com os outros, caso tenham assuntos e interesses em comum. Desta forma, um pesquisador pode, através da navegação na hierarquia de classes da ontologia, descobrir pares com quem possa trabalhar em conjunto.

Classe	Relacionamento	Com Classe
PESQUISADOR	*"pertence_a_unidade"*	UNIDADE
	"se_interessa_por"	ASSUNTO
	"é_especialista_em"	ASSUNTO

Tabela 6.19 – Classe PESQUISADOR e alguns de seus relacionamentos existentes na Ontologia ONTO-POINT, que permitem a função de navegação taxonômica

Exemplo 2: Outro exemplo de navegação é apresentado a partir da Tabela 6.20, onde vemos parte da ontologia ONTO-POINT envolvendo as classes GRUPO, ATOR e ASSUNTO. Através da navegação taxonômica, é possível encontrar GRUPOS de estudo e saber quem é o ator que lidera cada grupo, quem são cada um de seus membros, de quais unidades são originados e quais são os assuntos tratados e discutidos no âmbito do grupo.

Classe	Relacionamento	Com Classe
GRUPO	*"tem_como_lider"*	ATOR
	"é_formado_por"	ATOR
	"quer_discutir_assunto"	ASSUNTO
ATOR	*Superclasse de .....*	ALUNO
		PROFESSOR
		COORDENADOR CURSO
		PESQUISADOR
ATOR	*"pertence_a_unidade"*	UNIDADE
ASSUNTO	*"é_encontrado"*	DOCUMENTOS

Tabela 6.20 – Classes GRUPO e ATOR e alguns de seus relacionamentos existentes na Ontologia ONTO-POINT, que permitem a função de navegação taxonômica

Exemplo 3: Deseja-se buscar a relação de todos os livros do curso de Engenharia da Computação da unidade de Conselheiro Lafaiete da UNIPAC.

Levando-se novamente em conta a ontologia ONTO-POINT, destaca-se a classe UNIVERSIDADE com o seu relacionamento *"tem_unidade"* com a classe UNIDADE, a classe UNIDADE com seu relacionamento *"tem_curso"* com a classe CURSO, a classe CURSO com seu relacionamento *"tem_curriculo"* com a classe CURRÍCULO, a classe CURRÍCULO com seu relacionamento *"tem_area_formação"* com a classe AREA_FORMAÇÃO, a classe AREA_FORMAÇÃO com seu relacionamento *"tem_disciplina"* com a classe DISCIPLINA e finalmente a classe DISCIPLINA com seu relacionamento *"tem_bibliografia"* com a classe BIBLIOGRAFIA que por sua vez é superclasse da classe LIVRO, conforme pode ser visto na Tabela 6.21.

Baseado nos conceitos associados da ontologia e com as informações de que a UNIVERSIDADE é a "UNIPAC", a UNIDADE é "Conselheiro Lafaiete" e o CURSO é "Engenharia da Computação", deve-se buscar todos os CURRÍCULOS deste CURSO e todas as AREA_FORMAÇÃO de cada CURRÍCULO. Como cada AREA_FORMAÇÃO contempla um bom número de disciplinas, os LIVROS de cada uma dessas DISCIPLINAS respondem à questão que foi formulada, tomando-se apenas a precaução de se eliminar as ocorrências de livros repetidas no resultado final da pesquisa.

Classe	Relacionamento	Com Classe
UNIVERSIDADE	*"tem_unidade"*	UNIDADE
UNIDADE	*"tem_curso"*	CURSO
CURSO	*"tem_curriculo"*	CURRÍCULO
CURRÍCULO	*"tem_area_formação"*	AREA_FORMAÇÃO
AREA_FORMAÇÃO	*"tem_disciplina"*	DISCIPLINA
DISCIPLINA	*"tem_bibliografia"*	BIBLIOGRAFIA
BIBLIOGRAFIA	*Super classe de .....*	LIVRO

Tabela 6.21 – Classes da ontologia ONTO-POINT e seus relacionamentos que permitem a função de navegação taxonômica

Exemplo 4: Este exemplo mostra a navegação nas classes da ontologia à procura de ementas de disciplinas na ontologia do portal. Para ilustrar este caso, é usado Plone e a ferramenta Espaço Semântico, já apresentados e exemplificados anteriormente, com a ontologia ONTO-POINT. Através de navegação na hierarquia de classes da ontologia pode-se descobrir, então, quais são as ementas cadastradas para as várias disciplinas dentro do portal. Ao se escolher a classe EMENTA (B), são mostrados ao usuário os documentos do portal que têm como metadados os objetos desta classe (B). Um possível resultado deste processo de navegação pode ser visto na Figura 6.15.

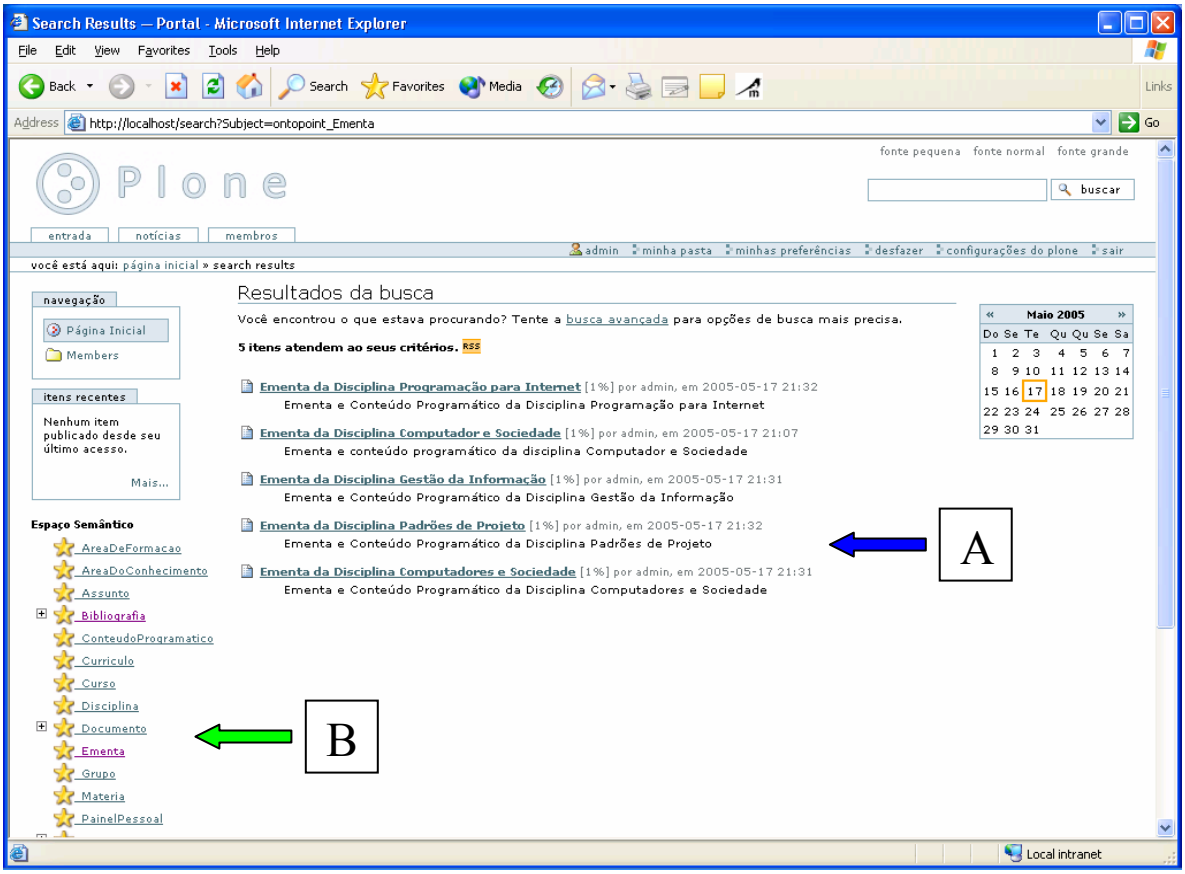


Figura 6.15 – Ementas das disciplinas cadastradas no portal (A), encontrados através de navegação na ontologia ONTO-POINT (B)

CAPÍTULO 7 - Considerações finais e trabalhos futuros

A pesquisa buscou compreender, através de revisão bibliográfica e de exemplos práticos dados no contexto de um estudo de caso, como um portal semântico i.e., baseado em uma ontologia, contribui para uma melhor gestão dos recursos de informação, em relação a portais tradicionais de informação, principalmente no que tange à reutilização e ao compartilhamento desses recursos em uma comunidade de usuários de uma instituição de ensino superior.

Segundo TARAPANOFF et al (2000), a gestão da informação se estrutura em um ciclo composto de seis processos distintos, mas interrelacionados: (1) identificação de necessidades de informação, (2) aquisição, (3) organização e armazenamento, (4) desenvolvimento de produtos e serviços, (5) distribuição, e (6) uso da informação. Sabe-se, por outro lado, que documentos e metadados, organizados por um sistema de gestão de conteúdo, favorece o processo de recuperação e promove reutilização e compartilhamento.

Este trabalho, sob o pano de fundo da especificação de um portal semântico, procurou mostrar como uma ontologia, e ferramentas baseadas em ontologias favorecem os processos constituintes desse ciclo, devendo integrar a próxima geração de sistemas de gestão de conteúdo. Para isso, foram sugeridas funções de personalização, publicação, navegação, colaboração, recuperação e busca, que permitem compartilhar e reutilizar a informação no portal.

Pode-se dizer que a personalização e a estruturação dos conceitos do domínio em uma hierarquia de classes, favorecem o processo de organização e armazenamento (3). A função de publicação, por sua vez, atende à demanda por descentralização e distribuição da informação (5), enquanto as funções de navegação e busca propiciam condições para aquisição e uso (2 e 6). A ontologia, que organiza o conteúdo, permite a navegação por uma série de informações relacionadas ao mesmo conceito, favorecendo a identificação e a satisfação de necessidades informacionais (1).

Argumentou-se durante todo o trabalho como uma ontologia, além de permitir a organização de um domínio do conhecimento, pode melhorar a recuperação da informação e permitir representação formal em lógica e mecanismos de inferência. Exemplos foram apresentados mostrando que a expressão de diversos tipos de relações (inversas, funcionais,

funcionais inversas, transitivas, simétricas etc) entre as classes permite que se realizem inferências, explicitando os relacionamentos existentes entre conceitos e viabilizando o controle da complexidade crescente da ontologia através de mecanismos de verificação formal.

PISANELLI et al (2002) afirmam que sistemas de informação, em geral, operam como 'ilhas', normalmente isoladas e independentes. Segundo o autor, "compartilhar e reusar conhecimentos entre estes sistemas não é fácil, uma vez que diferentes sistemas usam diferentes termos para se referir a um mesmo conceito e vice-versa". Armazenar metadados em repositórios de triplas RDF, além de permitir integração com outras fontes de informação, aumenta a possibilidade de elaboração de pesquisas mais precisas, através da utilização de linguagens de consulta nessas bases de dados, como a linguagem RDQL, por exemplo.

O uso de ontologias se mostrou eficiente na definição de um vocabulário comum onde o conhecimento a ser compartilhado pode ser representado. A ontologia funcionou como um vocabulário compartilhado de termos. O termo *matéria,* por exemplo, pode ter vários significados, entre eles: *substância*, *assunto* ou *disciplina;* sendo essa última acepção a adotada neste estudo. Do ponto de vista de um programa de computador, ao ser usado o termo *matéria,* nada pode ser concluído se nenhuma semântica ou significado for atribuído a ele, explícita ou implicitamente a ele no programa, pelo programador. A ontologia tem então a tarefa de estabelecer uma ligação entre a referência léxica e o conceito do mundo a que a representação se refere. Desta forma, existirá um acordo conceitual de que o termo *matéria,* neste caso específico do domínio estudado, é sinônimo de *disciplina* ou *área do conhecimento* e cumpre um certo papel se relacionando com outros conceitos da ontologia. Assim, o conceito, i.e., o acordo existente a respeito da semântica do termo permite uma comunicação dotada de significado e mais isenta de ambigüidades entre os agentes.

Pode-se concluir que usuários conseguem obter mais informações relevantes a respeito dos conceitos dentro de um domínio de aplicação. Isto pôde ser verificado em alguns dos exemplos mostrados neste trabalho. Observando-se o Exemplo 4, apresentado na Seção 6.2, conclui-se que, pelo fato de a disciplina *Bibliotecas Digitais* fazer parte de outra disciplina denominada *Gestão de Conteúdo* e esta, por sua vez, fazer parte da

disciplina *Gestão de Conhecimento*, um usuário interessado em Gestão de Conhecimento pode analisar as propriedades da disciplina Bibliotecas Digitais para obter valiosas informações a respeito do assunto que lhe interessa. O Exemplo, 1, da Seção 6.3, mostra que, pelo fato de os assuntos *Requisitos de Software, Projeto de Software e Qualidade de Software* fazerem parte de um assunto mais amplo denominado *Engenharia de Software*, usuários do portal interessados em se filiar a grupos que discutem o assunto Engenharia de Software poderão também se filiar ou realizar suas pesquisas em grupos mais específicos.

NOY e RAFNER (1997) afirmam que existe consenso sobre o benefício em se integrar ontologias, de forma a melhorar o compartilhamento e a reutilização dos conhecimentos existentes e disponíveis em cada uma delas. Deste modo, a ontologia ONTO-POINT construída neste trabalho pode ser integrada a outros sistemas de informação ou se integrar a outras ontologias existentes, relativas à área da educação superior.

Segundo GRUBER (1993), uma das metas das pesquisas em ontologias é a reutilização de componentes de conhecimento e serviços que podem ser chamados a partir das redes. Uma interpretação contemporânea desta visão são os *web services*, que podem ser instanciados a partir da Web Semântica. Em exemplos mostrados na Seção 6.6, foi apresentada a possibilidade de navegação semântica entre os termos marcados na ontologia e a instanciação de serviços Web que realizarão processamentos particulares para o conceito selecionado.

A apresentação de ferramentas semânticas, algumas delas detalhadas e exemplificadas com utilização da ontologia construída, procurou mostrar a existência de um grupo de ferramentas, que ainda estão, em sua maioria, nos laboratórios de pesquisa, mas que já tornam possível afirmar que, em um futuro bem próximo, a Web Semântica tem condições de deixar de ser uma promessa de melhor organização do conhecimento na Web e tornar-se realidade, melhorando o acesso à informação para todos os usuários da Internet.

A justificativa de um estudo que aprofunde a compreensão sobre ontologias na Ciência da Informação decorre do fato de serem a representação e a organização do conhecimento objetos privilegiados da área.

Segundo ALVARENGA (2003),

> "... no contexto da Ciência da Informação, o tratamento da informação se depara com uma tarefa complexa que assim poderia ser resumida: a despeito de todas as fragilidades dos atos de conhecer e comunicar, ..., torna-se imperativo que se encontre uma forma de se construir interfaces entre os acervos ... e seus usuários".

Este estudo também favorece a discussão de que, nos tempos atuais, está sendo exigida uma mudança de atitude dos profissionais diante da informação e do conhecimento. O ambiente digital vem ocasionando novas mudanças no trabalho de autores e demais profissionais da informação. Estes autores estão envolvidos com novas possibilidades tecnológicas, diretamente incidentes nos processos de produção, armazenagem, representação e recuperação de documentos e informações, que alteram seus processos de trabalho e produtos finais.

É fato que uma das preocupações de pesquisadores é a padronização da terminologia utilizada para organizar e recuperar a informação. Dessa forma, entende-se que este estudo se mostra útil à Ciência da Informação, em sua proposta de compreender como se dá a organização do conhecimento através da elaboração de ontologias, além da sugestão de uma arquitetura de portal semântico que provê meios para que usuários reutilizem e compartilhem informações, trabalhando em equipe, colaborando.

A proposta de implementação do portal semântico encontra alguns desafios práticos. Entre eles, a necessidade de pessoal especializado para realizar a manutenção constante da ontologia, visto que esta pode se tornar muito extensa. Existe ainda, como em todo portal, a necessidade de se considerar aspectos de certificação, que garantam a confiabilidade e a segurança dos documentos, e as questões de direitos autorais.

Como proposta de trabalhos futuros, sugere-se uma investigação em classes de ferramentas pouco aprofundadas neste trabalho, como por exemplo, ferramentas para personalização e colaboração.

Devido à constante evolução das ferramentas de código aberto, sugere-se também o desenvolvimento do portal aqui proposto através da composição e adaptação, em um único arcabouço de gestão de conteúdo, das funções semânticas pesquisadas e do repositório de metadados, além da possibilidade de se tomar como metáfora as ferramentas apresentadas neste estudo para implementar ferramentas próprias no arcabouço apresentado.

ANEXOS

ANEXO I - Classes e Relacionamentos da Ontologia ONTO-POINT

Classe	Relacionamento	Com Classe
UNIVERSIDADE	*"tem_unidade"*	UNIDADE
UNIDADE	*"é_da_universidade"*	UNIVERSIDADE
	"tem_curso"	CURSO
	"tem_ator"	ATOR
CURSO	*"é_da_unidade"*	UNIDADE
	"tem_currículo"	CURRÍCULO
	"tem_matrículas"	ALUNO
	"é_coordenado"	COORDENADOR DE CURSO
CURRÍCULO	*"pertence_ao_curso"*	CURSO
	"tem_área_formação"	ÁREA DE FORMAÇÃO
ÁREA DE FORMAÇÃO	*"é_do_currículo"*	CURRÍCULO
	"contém_disciplinas"	DISCIPLINA
DISCIPLINA	*"é_da_área_formação"*	ÁREA DE FORMAÇÃO
	"é_dividida_em_disciplina"	DISCIPLINA
	"faz_parte"	DISCIPLINA
	"tem_ementa"	EMENTA
	"tem_conteúdo_programático"	CONTEÚDO PROGRAMÁTICO
	"tem_bibliografia"	BIBLIOGRAFIA
	same as	MATÉRIA
	same as	ÁREA DO CONHECIMENTO
	"é_estudada"	ALUNO
	"é_ministada"	PROFESSOR
EMENTA	*"descreve_a_disciplina"*	DISCIPLINA
CONTEÚDO PROGRAMÁTICO	*"compõe_a_disciplina"*	DISCIPLINA
BIBLIOGRAFIA	*"é_da_disciplina"*	DISCIPLINA
MATÉRIA	*same as*	DISCIPLINA
	same as	ÁREA DO CONHECIMENTO
ÁREA DO CONHECIMENTO	*same as*	DISCIPLINA
	same as	MATÉRIA
PAINEL PESSOAL	*"é_personalizado_por"*	ATOR

GRUPO	*"quer_discutir_assunto"*	ASSUNTO
	"tem_como_lider"	ATOR
	"é_formado_por"	ATOR
ATOR	*"pertence_a_unidade"*	UNIDADE
	"é-especialista_em"	ASSUNTO
	"se_interessa_por"	ASSUNTO
	"insere"	DOCUMENTO
	"consulta"	DOCUMENTO
	"forma_o_grupo"	GRUPO
	"lidera_grupo"	GRUPO
	"utiliza_painel"	PAINELPESSOAL
ALUNO	*"está_matriculado"*	CURSO
	"estuda"	DISCIPLINA
PROFESSOR	*"ministra"*	DISCIPLINA
COORDENADOR DE CURSO	*"coordena"*	CURSO
ASSUNTO	*"tem_especialista"*	ATOR
	"é_interessante_para"	ATOR
	"é_encontrado"	DOCUMENTOS
	"pode_ser_disciplina"	DISCIPLINA
	"é_discutido_pelo"	GRUPO
DOCUMENTO	*"tem_tradutor"*	TRADUTOR
	"tem_autor"	AUTOR
	"tem_publicador"	PUBLICADOR
	"tem_organizador"	ORGANIZADOR
	"diz_respeito"	ASSUNTO
	"inserido_por"	ATOR
	"consultado_por"	ATOR
TRADUTOR	*"é_tradutor"*	DOCUMENTO
AUTOR	*"é_autor"*	DOCUMENTO
PUBLICADOR	*"é_publicador"*	DOCUMENTO
ORGANIZADOR	*"é_organizador"*	DOCUMENTO
EDITOR	*"é_editor"*	LIVRO EDITADO
LIVRO EDITADO	*"tem_editor"*	EDITOR

Tabela I.1 - Resumo das classes e relacionamentos da Ontologia ONTO-POINT

ANEXO II - Micro-especificação do Serviço 1 : "montar perfil de usuário e grupo de usuários"

Este serviço possibilita a montagem de perfis de usuários, para que eles achem os pares com quem querem compartilhar conteúdo e também para que possam ser achados por outros usuários cadastrados no portal. Permite também a montagem de grupos de usuários, cujo objetivo principal é colocar em contato mais estreito usuários que têm interesses comuns e permitir a colaboração entre eles. Estes perfis irão armazenar informações importantes sobre o comportamento e as preferências de cada ator e servirão de apoio a praticamente todos os outros serviços do portal.

- **Perfis de Usuário:**

A montagem de perfis é útil para que a página inicial do usuário seja um espaço agradável de se trabalhar e de se pesquisar, além de facilitar a navegação, as consultas, a comunicação e a colaboração entre os diversos atores, proporcionando maior performance e produtividade no uso do portal por parte de cada usuário.

Alunos, professores, coordenadores de curso e pesquisadores terão o seu próprio login[65] e senha para acesso ao portal semântico, que poderão, inclusive, ser os mesmos login e senha já usados para acessar o portal tradicional da instituição.

Cada perfil de usuário deverá conter algumas informações básicas de cada ator, como por exemplo, seus dados particulares, unidade em que trabalha (no caso de pesquisador) ou estuda (no caso de aluno), curso(s) e disciplina(s) que ministra aulas (no caso de professor) ou estuda (novamente no caso de aluno), entre outros.

Além desses dados pessoais, fazem parte do perfil do usuário os assuntos de seu interesse, sobre os quais gostaria de obter ou publicar informações a respeito e também obter ou salvar documentos sobre os mesmos.

A página principal do portal de cada usuário pode ser dividida em várias janelas, cada uma representando um serviço específico, que usa uma ou mais funções semânticas. Estes janelas podem ser selecionadas livremente de acordo com o interesse de cada ator, o

[65] código de usuário válido para entrada no portal

que propicia a personalização de cada área de trabalho. Estas informações sobre o uso de janelas selecionadas também compõem o perfil de um usuário.

É considerada também a existência do perfil do *visitante*, com apenas algumas poucas funcionalidades, quando comparados com um usuário registrado no portal, apenas sendo possível a navegação na ontologia e a visualização de informações e acesso a documentos liberados que sejam de seu interesse.

- **Perfis de Grupo de Usuários:**

Grupos de usuários podem ser montados de acordo com a conveniência e a necessidade dos usuários individuais e servem para o estudo em equipe, o desenvolvimento conjunto de trabalhos de sala de aula, artigos e apresentações ou até mesmo a realização conjunta de projetos de pesquisa entre professores localizados em unidades geograficamente distantes. Normalmente, o uso de ferramentas de colaboração se dará entre os membros de um grupo pré-estabelecido. Pessoas podem convidar outras para participar de um grupo ou um ator pode solicitar a sua entrada no grupo ao líder, normalmente, aquele que o criou.

Ao ser publicada uma notícia, um documento, ou ser colocado um *link* no portal, uma função permitirá a notificação dos usuários em geral, que estejam interessados naquele assunto a que se refere o recurso, ou emitir, de forma particular, uma notificação para um grupo em especial.

Durante a montagem de grupos de usuários, deve ser feita também a elaboração dos perfis do grupo, seguindo as mesmas regras de montagem de perfis individuais. O gerenciamento de entrada e saída de usuários no grupo deve ser realizada pelo usuário criador, ou líder, daquele grupo.

Conforme Figura II.1, o serviço "montar perfil de usuário e grupo de usuários" utiliza a função de *personalização*, mas também faz uso da função de *navegação*, que auxilia na descoberta do grupo ao qual o usuário deseja fazer parte e também a função de *colaboração*, no momento em que é necessário o envio de uma mensagem a todo o grupo notificando a filiação de mais um membro.

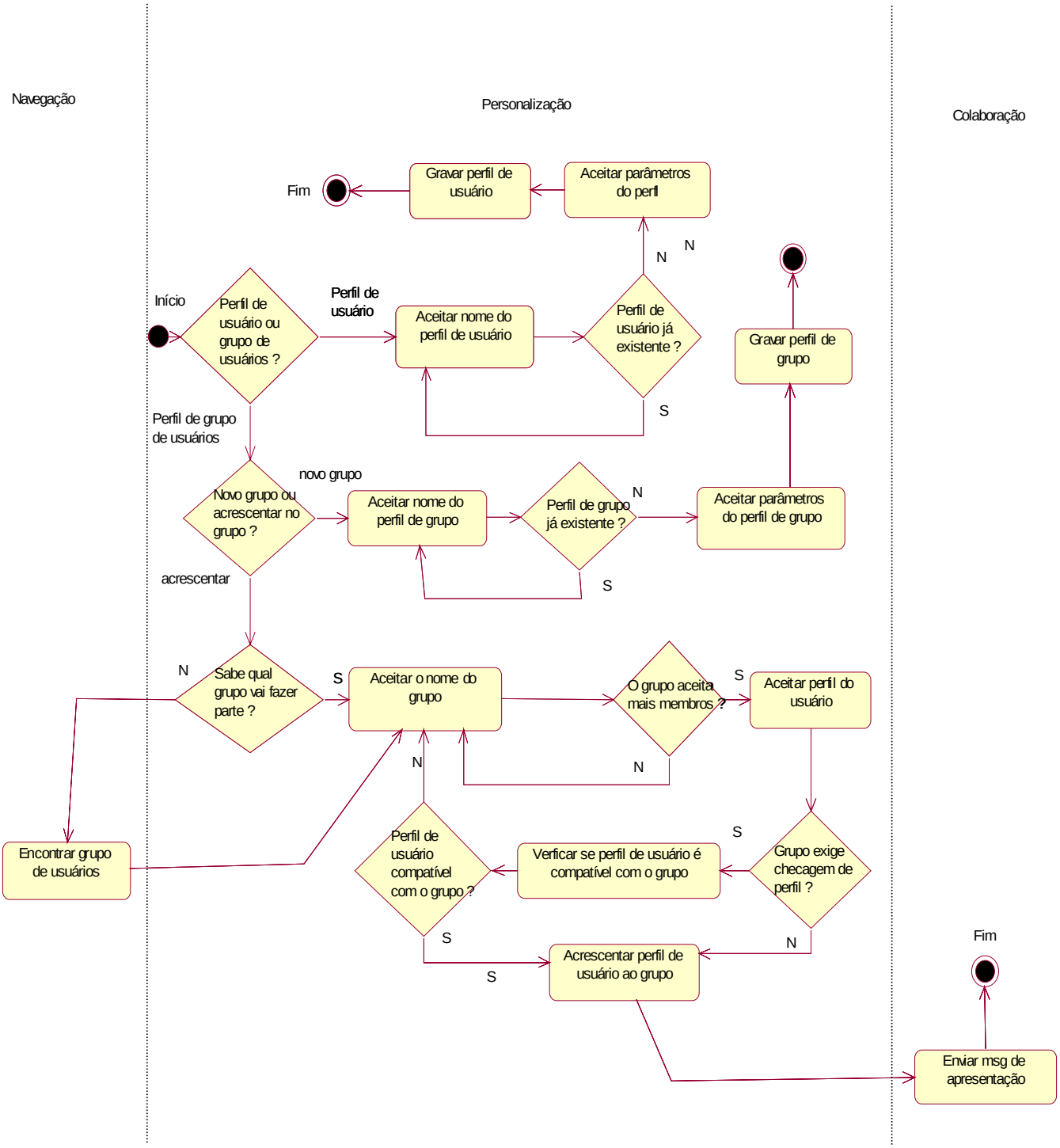


Figura II.1 – Diagrama de Atividades relativo ao serviço 1: "montar perfil de usuário e grupo de usuários"

A interface proposta para este serviço, conforme pode ser visto na Figura II.2, possibilita a gravação de perfil de usuário, perfil de grupo de usuário ou o acréscimo de um usuário a um grupo já definido.

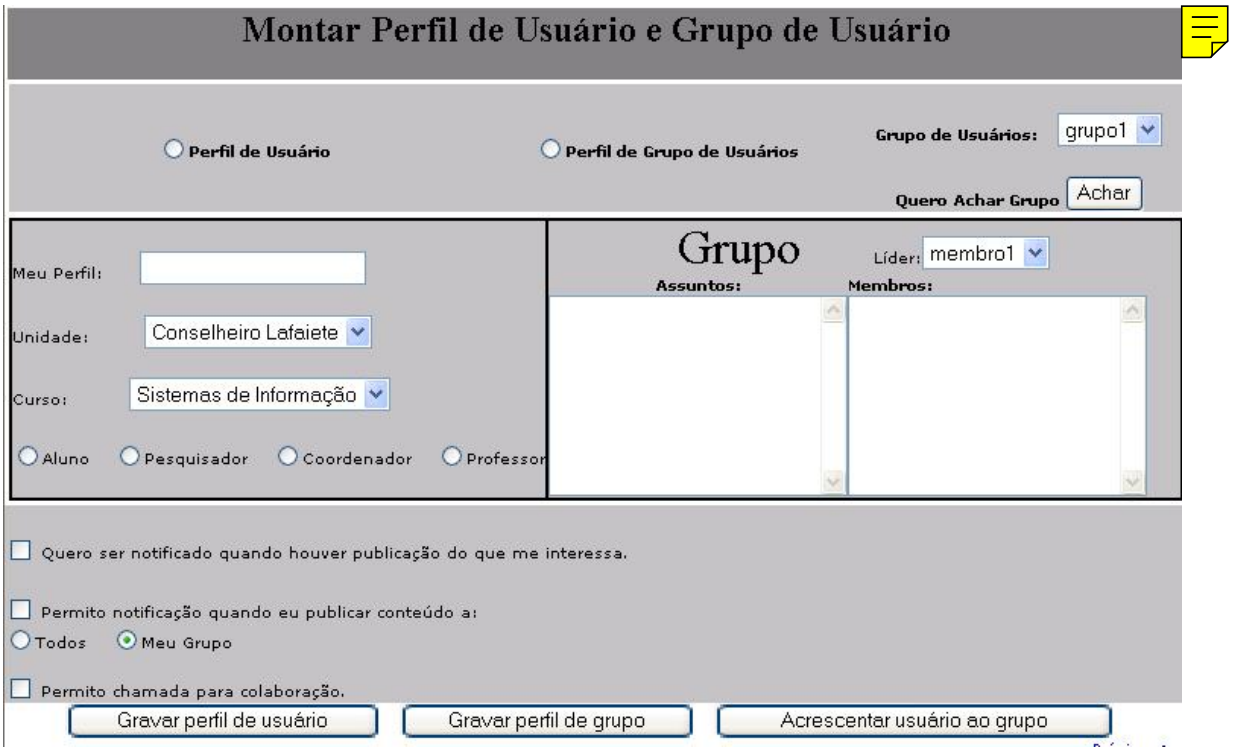


Figura II.2 – Um exemplo do que pode ser a Interface do serviço 1: "montar perfil de usuário e grupo de usuários"

Para a criação de perfis de usuário, além da unidade e do curso em que trabalha ou estuda, é necessário declarar qual é o tipo de ator: aluno, professor, coordenador ou pesquisador. Para a criação de grupo de usuários, destacam-se alguns atributos como o líder do grupo, assuntos que o grupo deseja discutir, bem como os atores que constituem os seus membros.

Na criação de novos perfis, seja de usuário ou grupo, algumas permissões devem ser declaradas, como por exemplo, se o usuário ou grupo deseja ser notificado quando algum conteúdo de seu interesse for publicado, se o usuário ou grupo permite a notificação de suas publicações e se permitem chamada para colaboração.

ANEXO III - Micro-especificação do Serviço 2 : "encontrar usuário e grupo de usuários"

O serviço de "encontrar usuário e grupo de usuários" é necessário para satisfazer a premissa básica da proposta do portal semântico, que é tornar-se um "ponto de encontro" entre os diversos atores do ambiente da universidade.

Através de navegação na ontologia do portal ou através de consulta montada pelo usuário, pode-se descobrir, por exemplo, pessoas que estudam no mesmo curso e fazem o mesmo período em unidades diferentes da universidade, ou lecionam a mesma disciplina em outra unidade, ou coordenam o mesmo curso também em unidades diferentes ou simplesmente, têm o mesmo interesse de estudo e pesquisa. Estas pessoas ou grupos, depois de descobertos, poderão ser contatados via correio eletrônico ou outra ferramenta de colaboração disponível no portal e poderão ser convidadas a participar de um assunto em discussão ou a interagir com os demais membros.

O serviço de busca de pessoas ou grupos permite o estabelecimento de contato entre os diversos atores, sendo possível, a partir daí, um compartilhamento de conhecimento tácito entre eles, o que é muito importante para o cumprimento dos objetivos propostos inicialmente pelo portal.

Conforme pode ser visto na Figura III.1, o serviço "encontrar usuário e grupo de usuários" utiliza as funções de *personalização*, *navegação* e *busca*. A primeira se encarrega do acesso e da disponibilização de perfis e as duas últimas auxiliam na descoberta do usuário ou grupo de usuários.

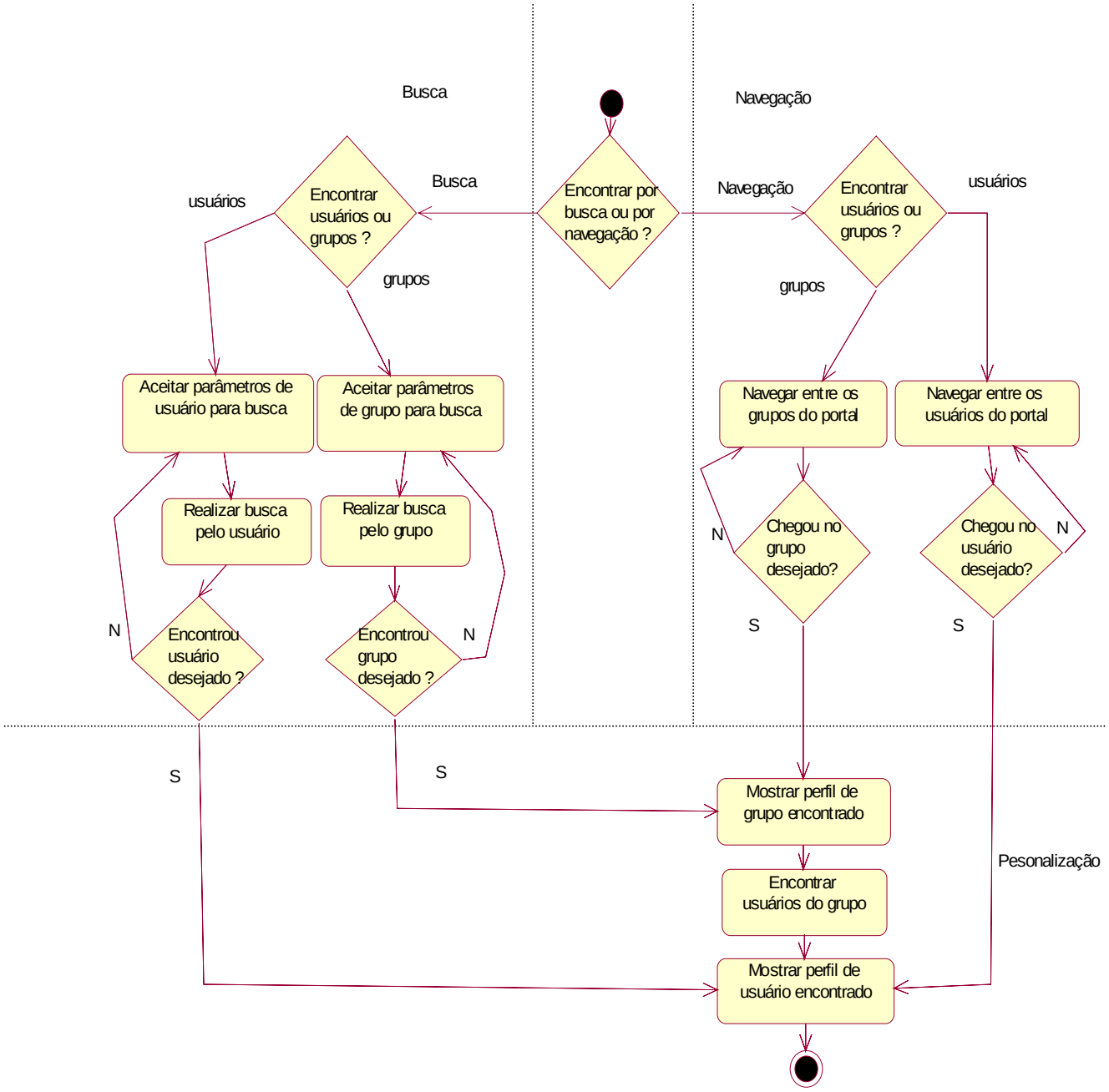


Figura III.1 – Diagrama de Atividades relativo ao serviço 2: "encontrar usuário e grupo de usuários"

A interface proposta para este serviço, conforme Figura III.2, possibilita a busca de usuário ou de grupo de usuários, através de alguns parâmetros de pesquisa.

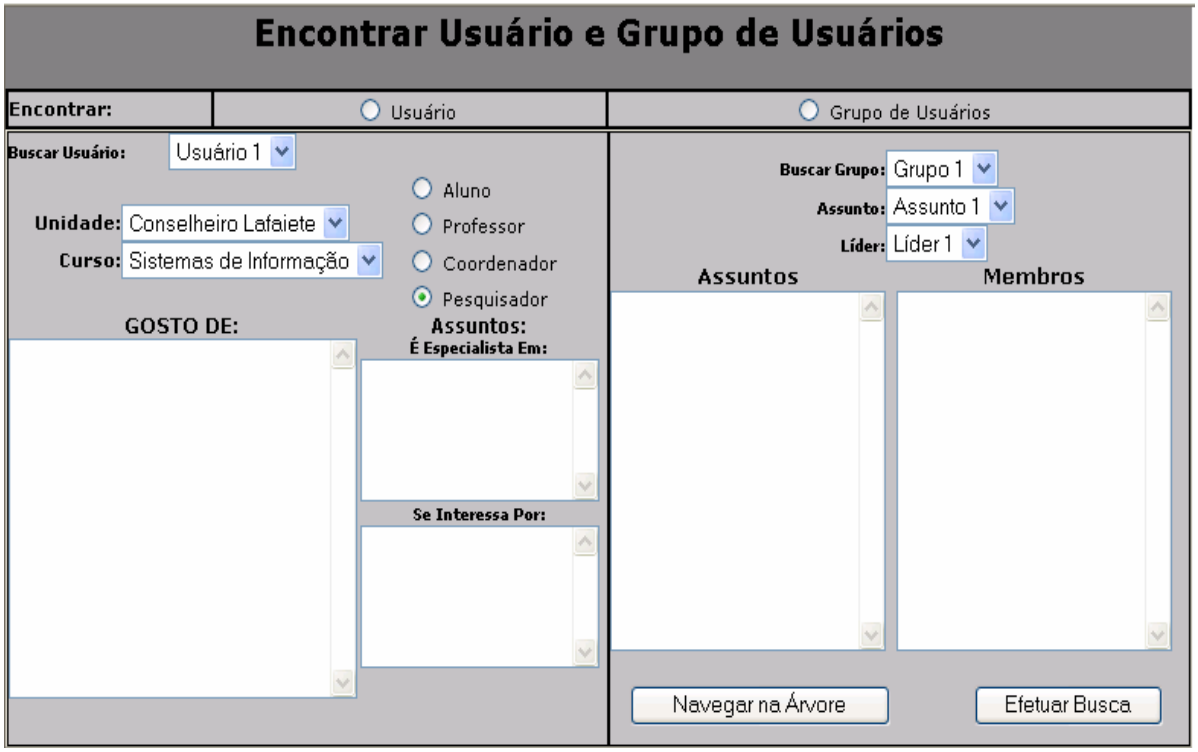


Figura III.2 – Um exemplo do que pode ser a Interface do serviço 2: "encontrar usuário e grupo de usuários"

Para se encontrar um usuário, por exemplo, pode-se efetuar a busca pelo nome do perfil, pela unidade ou curso ao qual pertence, além de existir a pesquisa pelo papel que desempenha no portal, ou seja, se o mesmo é aluno, professor, coordenador ou pesquisador. Ao ser encontrado, informações como os assuntos em que o mesmo é especialista ou pelos quais se interessa são mostradas.

Para se chegar a um grupo de usuários, por exemplo, pode-se efetuar a busca pelo nome do grupo, nome do líder ou pelo assunto discutido entre seus membros. Uma listagem de seus membros e dos assuntos discutidos em seu âmbito são mostrados quando encontrado o grupo desejado.

ANEXO IV - Micro-especificação do Serviço 3 : "permitir colaboração entre usuários"

O serviço "permitir colaboração entre os usuários", possibilita que pessoas possam trabalhar em conjunto ou, simplesmente, trocar impressões e informações.

O serviço de colaboração permite a comunicação, síncrona ou assíncrona, entre os usuários do portal. As ferramentas de colaboração também podem ser usadas por outros serviços para estender suas funcionalidades, como por exemplo, no momento da publicação de conteúdo ou o momento da recuperação de documentos.

Conforme pode ser visto na Figura IV.1, o serviço "permitir colaboração entre os usuários" utiliza a função de *colaboração*, além da função de *personalização*, quando necessário checar perfis e as funções de *busca* e *navegação*, quando necessário encontrar o usuário ou grupo para colaboração.

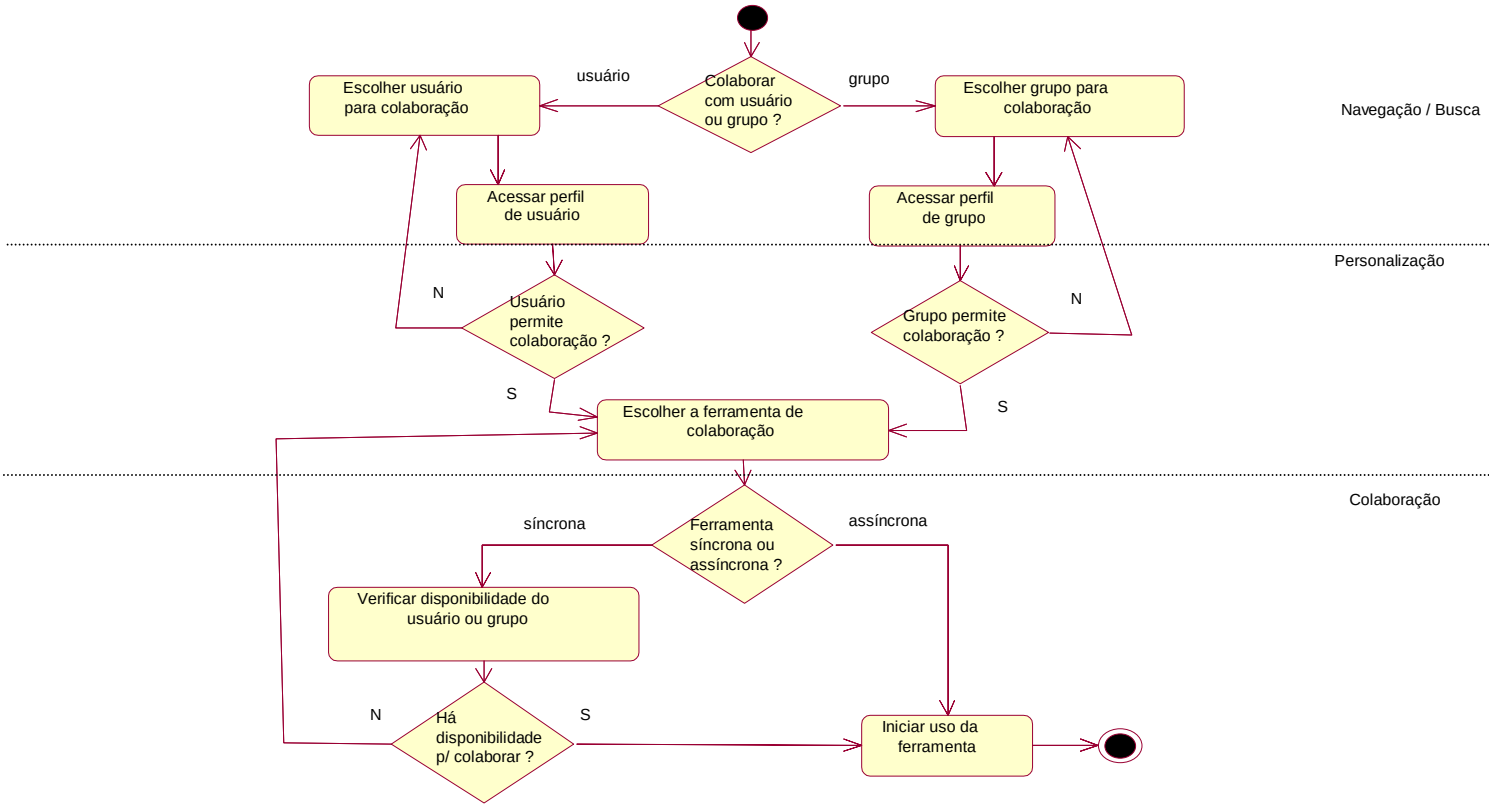


Figura IV.1 – Diagrama de Atividades relativo ao serviço 3: "permitir colaboração entre usuários"

O serviço "permitir colaboração entre os usuários", conforme Figura IV.2, vai utilizar principalmente a janela onde são apresentados os endereços eletrônicos previamente cadastrados pelo usuário do portal e a janela onde são mostradas as pessoas que estão on-line em um certo momento. A partir daí o usuário pode escolher com quem deseja manter colaboração e selecionar a ferramenta desejada para então iniciar a comunicação com ela.

A colaboração é permitida tanto para usuários individuais quanto para grupos de usuários. Em caso de seleção de ferramenta síncrona de colaboração, o perfil do usuário selecionado será acessado para verificação se o mesmo permite, naquele momento, este tipo de comunicação.

Figura IV.2 – Um exemplo do que pode vir a ser a Interface do serviço 3: "permitir colaboração entre usuários"

ANEXO V - Micro-especificação do Serviço 4 : "publicar conteúdo"

Qualquer ator, desde que tenha um perfil montado e acesso autorizado, pode publicar documentos no portal semântico UNI-POINT. A idéia é tornar mais fácil para os usuários não-técnicos entrarem com informações no sistema, permitindo assim disponibilizar estas informações de maneira rápida e eficiente para todos ou para um grupo direcionado. Todos esses documentos deverão ser marcados, para atendimento aos pré-requisitos do portal semântico, utilizando uma ferramenta de anotação ou marcação semântica.

Os documentos deverão ser associados a metadados quando da sua publicação no portal, de forma a serem facilmente recuperados por outros usuários posteriormente. Além disso, todo conteúdo, ao ser publicado, poderá ser marcado semanticamente com o mínimo de esforço por parte do usuário e as anotações armazenadas em um repositório de metadados para posterior recuperação.

Ao ser realizada essa marcação semântica, os conceitos são associados a classes existentes na ontologia, sendo permitida a inclusão de observações e outros tipos de anotações por parte do usuário. Desta forma, cada conceito marcado vira uma instância de uma classe da ontologia, abrindo-se assim a possibilidade de recuperação dos documentos através de palavras-chaves e também através da ligação entre os conceitos existente na ontologia.

Ao usuário publicador é facultada a disponibilização do material publicado a toda a comunidade ou a grupos específicos. Ferramentas de colaboração podem ser usadas para enviar mensagens a usuários ou a grupos interessados no assunto a que diz respeito o documento recém publicado. Além disso, o sistema deve favorecer o publicador com a possibilidade de ser informado toda vez que algum usuário se interessar pelo material colocado por ele no portal semântico.

Conforme pode ser visto na Figura V.1, o serviço "publicar conteúdo" utiliza as funcionalidades existentes na função de ***publicação***, além de se servir da função de *personalização,* que permite a checagem dos perfis e a função de *colaboração,* que permite a notificação quando da inclusão de determinado conteúdo no portal.

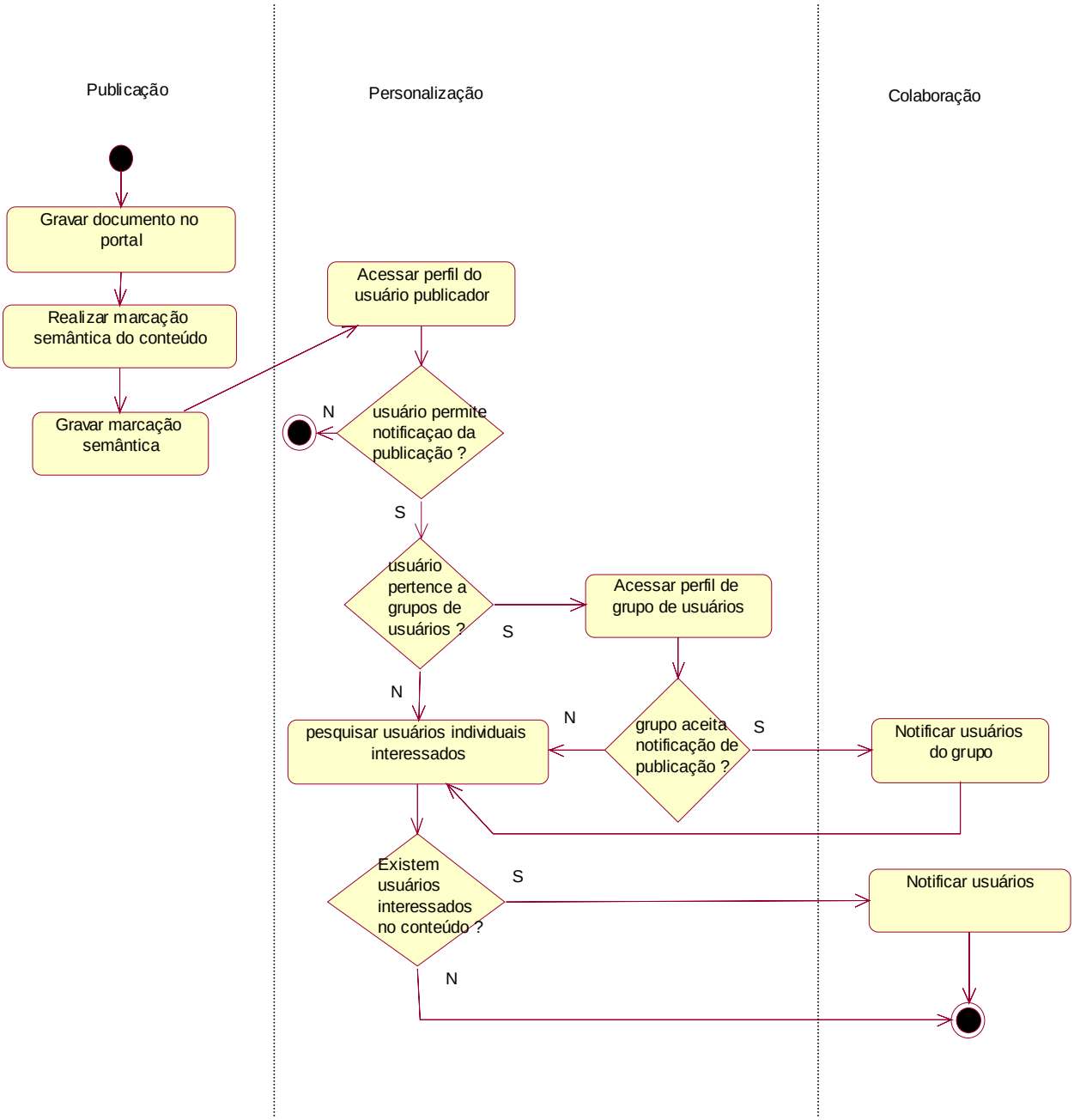


Figura V.1 – Diagrama de Atividades relativo ao serviço 4: "publicar conteúdo"

Uma interface para o serviço de publicação de conteúdo passa necessariamente pela idéia de se marcar semanticamente termos ou expressões de acordo com o vocabulário ontológico do domínio em questão. Na Figura V.2, extraído do aplicativo KIM, tem-se um exemplo do que pode ser uma marcação semântica de conceitos, onde os termos são associados a classes existentes na ontologia, sendo permitida a inclusão de observações e outros tipos de anotações por parte do usuário. A utilização de cores é uma estratégia muito usada em ferramentas deste tipo, de forma a possibilitar uma rápida associação entre o termo marcado e a classe à qual ele faz parte, pelo fato de apresentarem a mesma cor na tela de Interface.

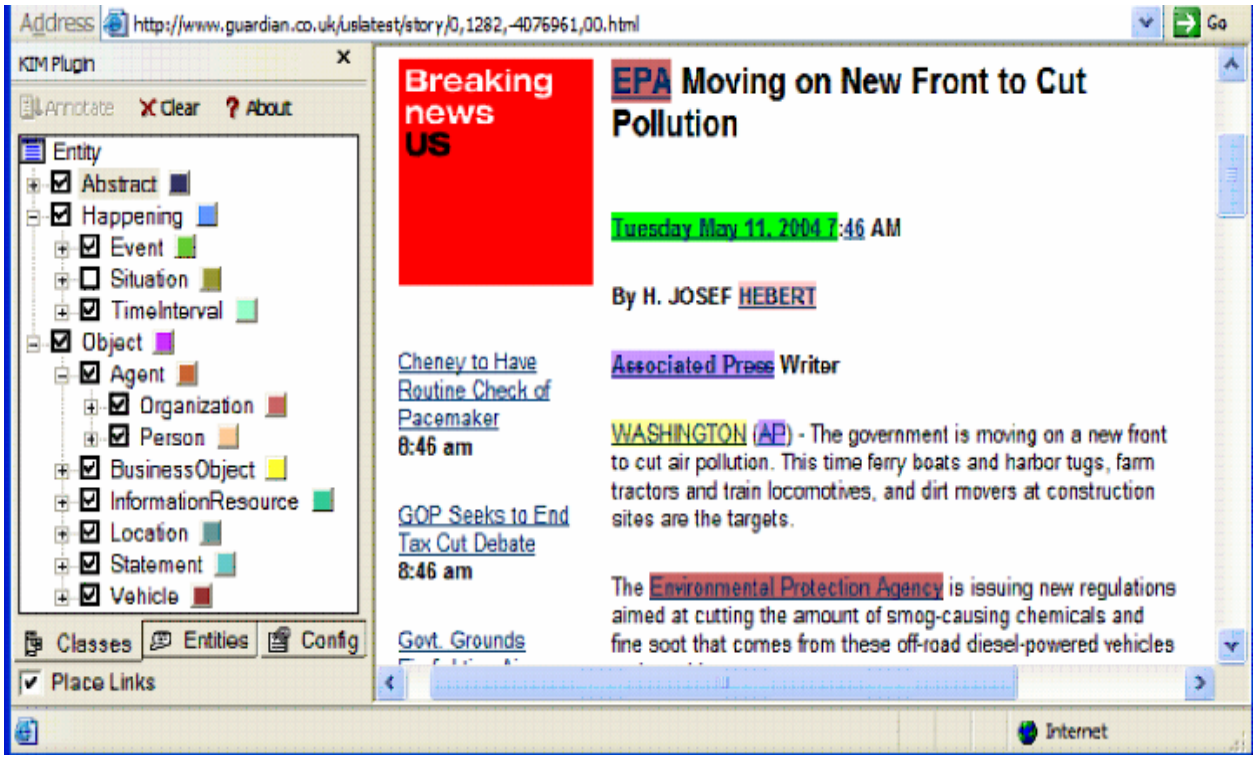


Figura V.2 – Um exemplo do que pode vir a ser a Interface do serviço 4: "publicar conteúdo

ANEXO VI - Micro-especificação do Serviço 5 : “recuperar conteúdo"

Todo o conteúdo armazenado e marcado semanticamente no portal pode ser recuperado por outros usuários que demonstrem interesse neles, através do serviço de recuperação de conteúdo.

Através de navegação nas classes da ontologia ou através de consulta montada pelo usuário e submetidas ao repositório de metadados do portal, pode-se descobrir documentos relevantes e encontrar o conteúdo desejado pelo usuário.

Os metadados, base das linguagens de marcação, como XML, por exemplo, são responsáveis por prover os contextos de significados necessários à informação, tornando os documentos mais facilmente localizáveis por mecanismos de busca e auxiliando-os a resolver diversas ambigüidades das línguas naturais que trazem ruídos ao processo de busca. Metadados também desempenham importante função ao permitir que diferentes documentos sejam agrupados mais facilmente, o que também facilita o processo de recuperação de conteúdo.

Conforme pode ser visto na Figura VI.1, o serviço "recuperar conteúdo" utiliza as funções de *navegação* ou *busca*, além da função de *personalização*, que auxilia na recuperação do perfil dos usuários e da função de *colaboração*, que auxilia na notificação do usuário publicador de que outro usuário se interessou pelo conteúdo por ele postado.

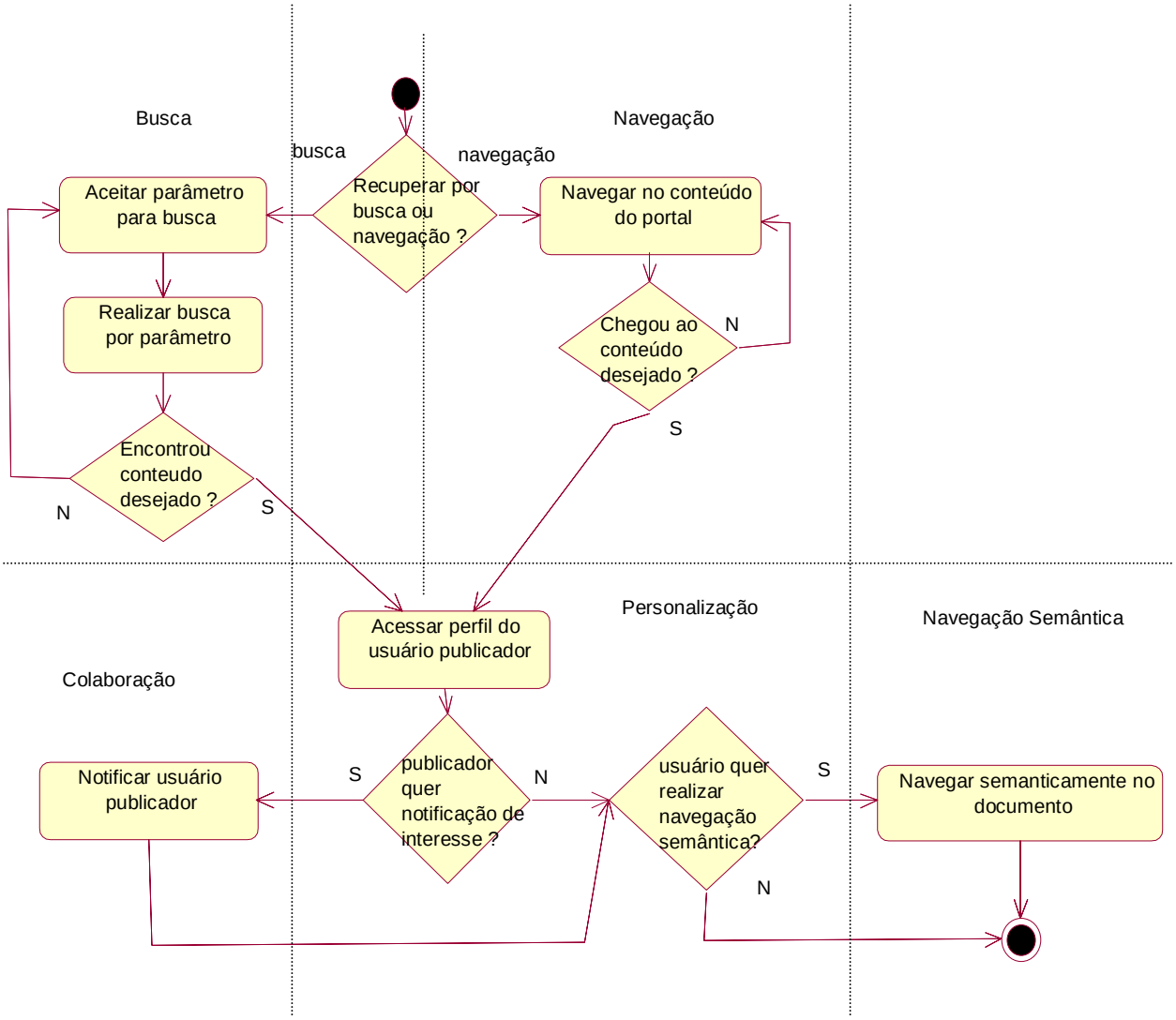


Figura VI.1 – Diagrama de Atividades relativo ao serviço 5: “recuperar conteúdo"

A interface proposta para este serviço, conforme Figura VI.2, possibilita a recuperação de conteúdo através de vários parâmetros de pesquisa.

Figura VI.2 – Um exemplo do que pode ser a Interface do serviço 5: "recuperar conteúdo"

Pode-se chegar aos documentos desejados através de busca por assunto ou por qualquer ator que tenha relação com os documentos marcados semanticamente no portal. Desta forma, pode-se procurar pelo autor, tradutor, organizador, publicador ou pelo responsável pela inserção do conteúdo do portal, além de poder procurar pela data da inserção. Vários tipos de documentos são possíveis de serem recuperados, bastando ao usuário marcar o botão de opção correspondente a anais, livros, teses, relatórios técnicos, publicações diversas e publicações periódicas, além de outros que podem ser acrescentados posteriormente na ontologia que dá suporte ao portal semântico.

ANEXO VII - Micro-especificação do Serviço 6: "navegar semanticamente"

Quando o usuário visualizar o conteúdo de um documento armazenado no portal, terá à sua disposição a opção de navegação semântica no mesmo, isto é, poderá aproveitar-se da marcação semântica existente no documento para 'saltar' de um conceito a outro da ontologia, através dos relacionamentos existentes entre eles.

Quando de sua publicação, conceitos existentes no documento postado se tornaram instâncias de classes da ontologia ONTO-POINT, através de um processo de marcação semântica. Esta classe já estava ligada a outros conceitos, através de um processo semelhante que aconteceu anteriormente. É exatamente esta ligação que favorece o processo de navegação semântica, que é a ligação de conceitos marcados semanticamente e uma posterior chamada de serviços que realizam processamentos especiais para cada termo individualmente.

Conforme pode ser constatado na Figura VII.1, o serviço "navegar semanticamente" utiliza apenas a função de *navegação semântica.*

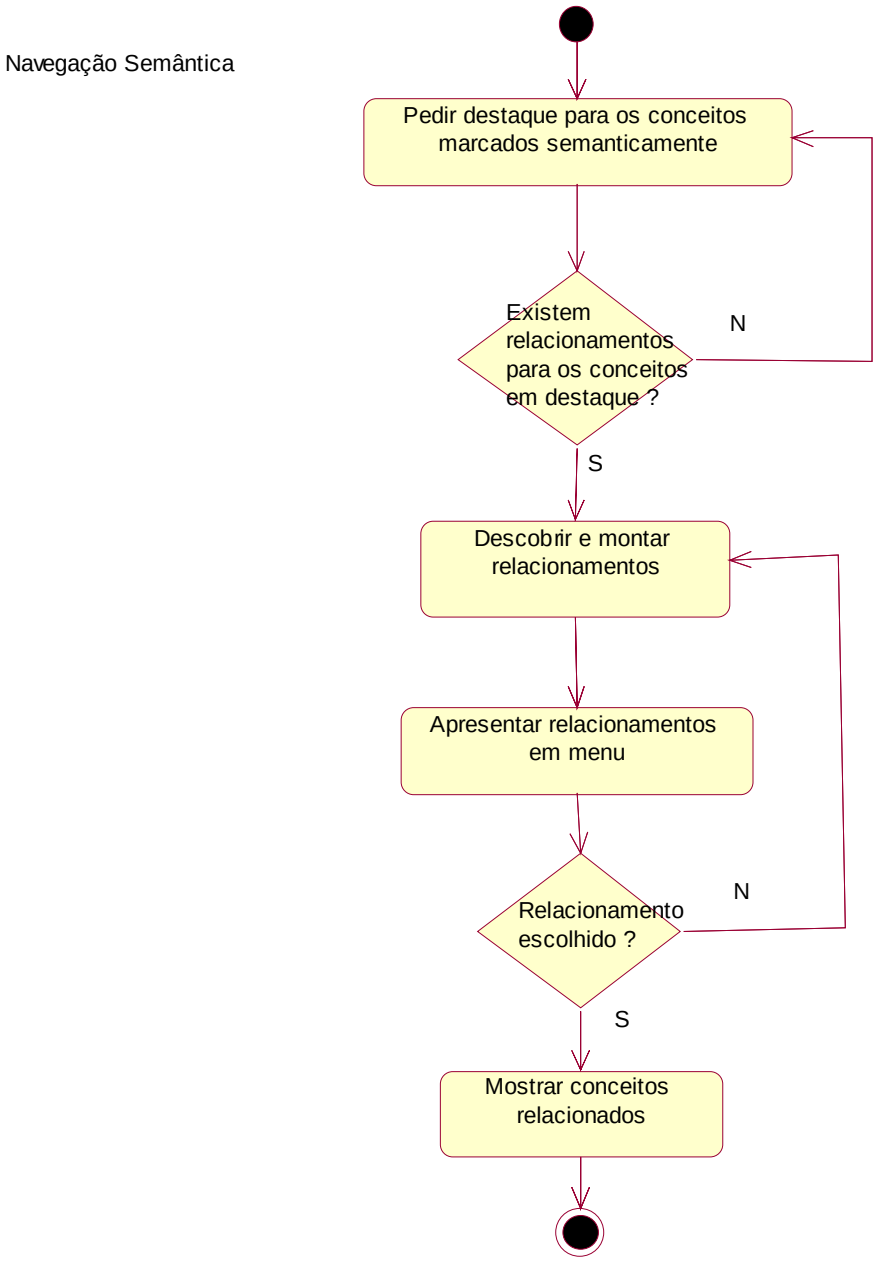


Figura VII.1 – Diagrama de Atividades relativo ao serviço 6: "navegar semanticamente"

Uma interface para o serviço de navegação semântica passa necessariamente pela idéia de se estender as propriedades e relacionamentos de um conceito da ontologia, conforme pode ser visualizado na Figura VII.2, extraído do aplicativo KIM.

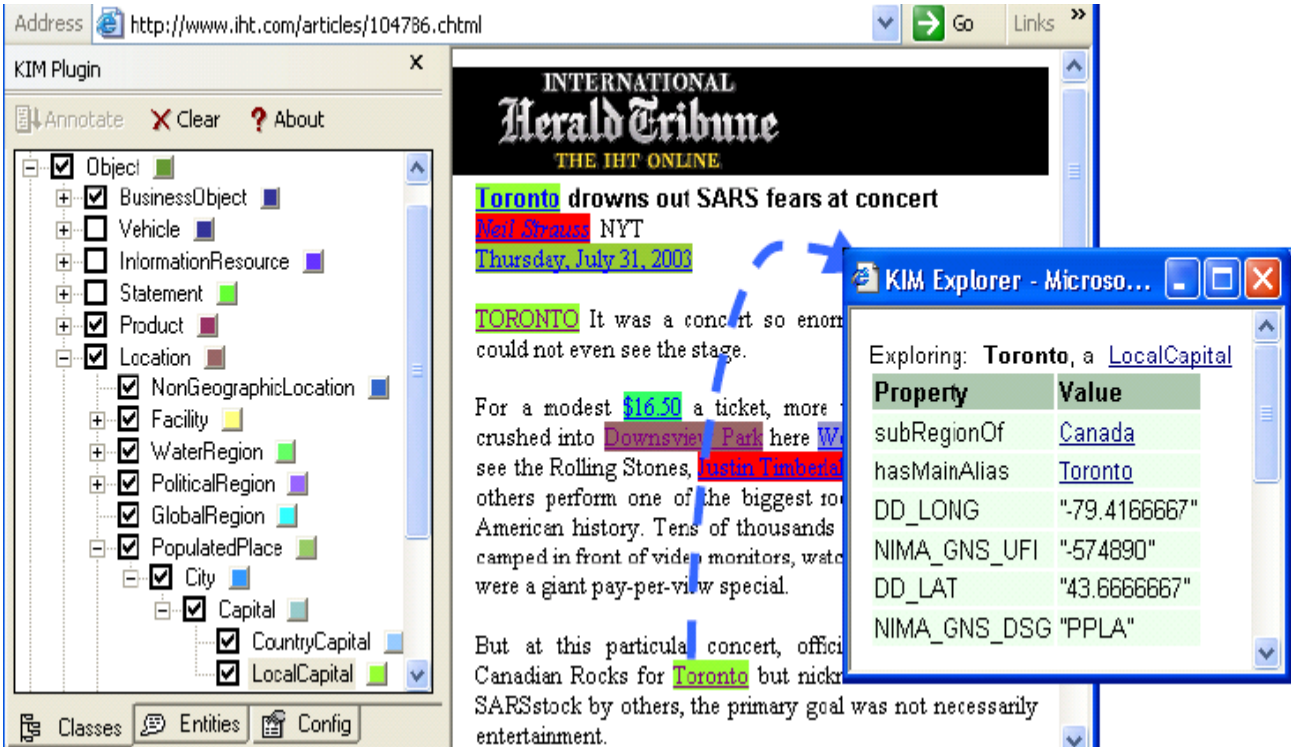


Figura VII.2 – Um exemplo do que pode ser a Interface do serviço 6: "navegar semanticamente"

A utilização de uma legenda de cores também é uma boa estratégia neste tipo de situação, com o objetivo de diferenciar os diversos termos e expressões existentes em um documento e de associá-los aos conceitos da ontologia. Além desta marcação por cores, uma interface de navegação semântica também deve estender o significado do conceito, mostrando suas propriedades e seus relacionamentos com outros conceitos existentes dentro do domínio. Desta forma, o usuário pode ser informado ao acessar um objeto que deve também acessar o outro, pois estes possuem uma relação semântica importante.

Referências Bibliográficas

ALVARENGA, L. Representação do conhecimento na perspectiva da Ciência da Informação em tempo e espaço digitais. Encontros Bibli, Florianópolis, n.15, 2003

BAWDEN, D. Information Overload. Library and Information Briefings. v.92, Jan 2001. http://litc.sbu.ac.uk/puclications/libs/libs92.pdf

BERNERS-LEE, T., HENDLER, J., LASSILA, O. The Semantic Web. A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. Scientific American, Issue 01/05/2001.
http://www.sciam.com/2001/0501issue/0501berners-lee.htm

BOOCH, G., RUMBAUGH, J., JACOBSON, I. UML – Guia do Usuário. Rio de Janeiro: Editora Campus, 2000.

BORGHOFF,U., PARESCHI,R. Information Technoloy for Knowledge Management". Berlin, Heidelberg: Springer, 1998.

BOZSAK, E. et all. KAON – Towards a large scale Semantic Web. In E-Commerce and Web Technologies, Third International Conference, EC-Web 2002, Aix en-Provence, France, September 2-6, 2002, Proceedings, Springer, Lecture Notes in Computer Science, volume 2455
http://www.aifb.uni-karlsruhe.de/WBS/dob/pubs/ecWeb2002.pdf

BRAGANHOLO, V.; HEUSER, C. XML Schema, RDF(S) e UML: uma comparação. In: IDEAS 2001 - 4th Iberoamerican Workshop on Requirements Engineering and Software Environments , Santo Domingo, Heredia, Costa Rica, 2001.
http://www.inf.ufrgs.br/~vanessa/artigos/ ideas2001.pdf

BRICKLEY, D.; GUHA, R.V. Resource Description Framework (RDF) Schema specification 1.0 27 março 2000, 1.ed., 2000.
http://www.w3.org/TR/2000/CR-rdf-schema-20000327

BROADER, A. A taxonomy of Web search, ACM SIGIR Forum, v.36, n.2, ACM Press, 2002. Online. Disponível na Internet via www.
http://doi.acm.org/10.1145/792550.792552

CAPURRO, R. What is Information Science for ?; a philosophical reflection. In: Vakkari, Perti, Cronin, Blaise. Conceptions of library and information science. Tempere, Taylor Graham, 1991

CHAVES, M.S.,VIEIRA,R., RIGO,S. Uso de Ontologias para gerenciamento e acesso a documentos na Web. In: V Oficina de Inteligência Artificial, 2001, Pelotas.
http://xldb.di.fc.ul.pt/~mchaves/pg_portugues/public/voia.pdf

DOMINGUE, J., DZBOR, M., MOTTA, E. Semantic Layering with Magpie. In Staab, S. and Studer, R. (Eds), Handbook on Ontologies in Information Systems Springer Verlag, 2004

DZBOR, M., DOMINGUE, J., MOTTA, E. Magpie – Towards a Semantic Web Browser. 2nd International Semantic Web Conference (ISWC-2003) 20-23 October 2003 Sundial Resort, Sanibel Island, Florida, USA

FENSEL, D. Ontologies: Silver Bullet for knowledge Management and Eletronic Commerce, Springer-Verlag, Berlin, 1999.
http://www.cs.vu.nl/ ~dieter

GOLBECK, J.; ALFORD, A.; HENDLER, J. Handbook of Human Factors in Web Design, chapter Organization and Structure of Information using Semantic Web Technologies, 2003.
http://www.mindswap.org/papers/Handbook.pdf

GRUBER, T.R. A Translation Approach to Portable Ontology Specifications. Knowledge Acquisition, 5, 1993.

GUARINO, N. Understanding, Building and Using Ontologies. A commentary to "Using Explicit Ontologies in KBS Development", by van Heijst, Schreiber, and Wielinga. International Journal of Human and Computer Studies , 46, 1997.
http://ksi.cpsc.ucalgary.ca/KAW/KAW96/guarino/ guarino.html

GUARINO, N. Formal Ontology and Information Systems. Proceddings of FOIS' 98, Trento, Italy, 6-8 June 1998. Amsterdan, IOS Press, pp. 3-15.
http://www.loa-cnr.it/Papers/FOIS98.pdf

HORRIDGE, M., KNUBLAUCH, H., RECTOR, A., STEVENS, R., WROE, C. A pratical guide to building OWL Ontologies using the Protégé OWL Plug-in and CO-ODE Tools. The University of Manchester, Agosto, 2004
http://www.co-ode.org/resources/tutorials/ProtegeOWLTutorial.pdf

KOCH, M., LACHER,M. "Integrating Community Services – A Common Infraestructure Proposal". In: Proceedings Knowledge Based Intelligent Enginnering Systems and Allied Technologies (KES2000). Brighton, UK: September, 2000

KNUBLAUCH, H., MUSEN,M.A. Editing Description Logic Ontologies with the Protégé OWL Plug-in. The Description Logic Workshop, Canadá, 2004.
http://protege.stanford.edu/plugins/owl/publications/DL2004-protege-owl.pdf

LACHER M.S., KOCH, M., WORNDL,W. A framework for personalizable Community Web Portals. Proceedings of Human-Computer Interaction International, New Orleans, LA, USA, August 5-10, 2001, Aug 2001.
http://www11.informatik.tu-muenchen.de/publications/pdf/lacher2001a.pdf

LASSILA, O.; SWICK, R. Resource Description Framework (RDF) model and syntax specification.1.0, 22 Feb 1999.
http://www.w3c.org/TR/REC-rdf-syntax

LOHL,S. et all. Investigação sobre a Identificação de Assuntos em Mensagens de Chat. II TIL – Workshop de Tecnologia da Informação e da Linguagem Humana. Sociedade Brasileira da Computação. 2004.
http://www.nilc.icmc.usp.br/til2004/anais_til2004.pdf

LOPEZ, M.F. Overview of methodologies for building ontologies, 1999.
http://www.ontology.org/main/presentations/madrid/analysis.pdf

MAEDCHE, A., STAAB, S., STOJANOVIC,N., STUDER,R.,SURE,Y. Semantic Portal: The SEAL Approach. In D. Fensel, J. Hendler, H. Lieberman, W. Wahlster (eds.), *Spinning the Semantic Web*, pp. 317-359. MIT Press, Cambridge, MA., 2003.
http://www.aifb.uni-karlsruhe.de/~sst/Research/Publications/semanticportal.pdf

MYERS, M. D. "Qualitative Research in Information Systems," MIS Quarterly (21:2), June 1997, pp. 241-242. MISQ Discovery, archival version, June 1997, http://www.misq.org/discovery/MISQD_isworld/. MISQ Discovery, updated version, last modified: April 14, 2005.
http://www.qual.auckland.ac.nz

MYNATT,E.D., ADLER, A.,ITO,M.,ODAY,V.L. Design for Network Communities. In: Proceedings ACM SIGCHI Conference on Human Factors in Computer Systems; 1997.

NOY, N.F. and HAFNER, C.D., "The State of the Art in Ontology Design – A Survey and Comparative Review", *AI Magazine,* 36(3), 53-74, 1997.

NOY,N.F., McGUINESS, D.L. Ontology Development 101: A Guide to Creating your first ontology. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001. http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html

PEREIRA, J.C.L., BAX, M.P. Introdução à Gestão de Conteúdos. Anais do 3o. Workshop Brasileiro de Inteligência Competitiva e Gestão do Conhecimento, 2002

PINHEIRO,W.A.,MOURA,A.M.C. Semantic Search in Portals using Ontologies. I Workshop de Web Semântica (WWS2004), Brasília, 22 de outubro de 2004 http://www.ucb.br/ucbtic/wws2004/articles/25_wws2004_final.pdf

PISANELLI, D.M., GANGEMI, A. and STEVE, G., "Ontologies and Information Systems: The Marriage of the Century?", *Proceedings of the LYEE Workshop,* Paris, 2002.

PITTS-MOULTIS, N. e KIRK, C. XML Black Book - Solução e Poder. Makron Books, 2000.

REYNOLDS, D.,; SHABAJEE, P.; CAYZER, S. Semantic Information Portals, ACM, New York, May 2004. http://www2004.org/proceedings/docs/2p290.pdf

SEABORNE, A. RDQL – A Query Language for RDF. Hewlett Packard Labs Bristol, 2003. http://www.w3.org/Submission/2004/SUBM-RDQL-20040109/

SOUZA, L. S. Relacionamento semântico de conteúdos em portais: Estudo e proposta de implementação baseados em taxonomia. 2005. Relatório Final (Iniciação Científica) – Departamento de Ciência da Computação, UFMG, Belo Horizonte, 2005.

SPYNS, P., OBERLE, D., VOLZ, R., ZHENG, J., JARRAR, M., SURE, Y., STUDER, R., MEERSMAN, R. ONTOWEB - a Semantic Web Community Portal. In Proc. Fourth International Conference on Practical Aspects of Knowledge Management (PAKM), December 2002, Vienna, Austria, 2002.
http://citeseer.ist.psu.edu/jarrar02ontoWeb.html

STAAB, S.; STUDER, R.; SURE, Y. AI for the Web – Ontology-based Community Web Portals. American Association for Artificial Intelligence, 2000. In *AAAI 2000/IAAI 2000 - Proceedings of the 17th National Conference on Artificial Intelligence and 12th Innovative Applications of Artificial Intelligence Conference, Austin/TX, USA, July 30-August 3, 2000*. AAAI Press/MIT Press, 2000.
http://www.aifb.uni-karlsruhe.de/WBS/Publ/2000/iaai_sstetal_2000.pdf

STOJANOVIC, N., MAEDCHE, A., STAAB, S., STUDER, R., SURE Y. SEAL: A framework for developing SEmantic PortALs. Proceedings of the International Conference on Knowledge capture October 2001

TARAPANOFF, K., ARAÚJO, R.H., CORMIER, P.M.J. Sociedade da informação e inteligência em unidades de informação. Ciência da Informação, Brasília, v. 29, n. 3, p. 91-100, set./dez. 2000

TERRA, J.C.,GORDON,C. Portais Corporativos.São Paulo: Negócio Editora, 2002

TERRA, J.C., BAX ,M.P. Portais Corporativos : Instrumentos de Gestão de Informação e de Conhecimento. In: A Gestão da Informação e do Conhecimento. Belo Horizonte: Editora UFMG, 2003.

USCHOLD, M., GRUNINGER, M. Ontologies: Principles, Methods and Applications. Knowledge Engineering Review, 1996

VICKERY, D.C. Knowledge Representation: a brief review. *Journal of Documentation*, v.42, n.3, sept. 1986

WILSON, T. Structure in Research Methods, 1999.
http://informationr.net/tdw/publ/ppt/ResMethods/

ZHANG, L., YU, Y., ZHOU, J., LIN, C., YANG, Y. An enhanced model for searching in semantic portals, Proceedings of the 14th international conference on World Wide Web. May 2005.