

UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
COMPUTAÇÃO

Tiago Rios da Rocha

UM MODELO DE ANOTAÇÕES LIVRES E SEMÂNTICAS
PARA BIBLIOTECAS DIGITAIS

Dissertação de Mestrado submetida à Universidade Federal de Santa Catarina como parte dos requisitos para obtenção do grau de Mestre em Ciência da Computação.

Prof. Dr. Roberto Willrich

Florianópolis, fevereiro de 2010.

Catálogo na fonte pela Biblioteca Universitária da
Universidade Federal de Santa Catarina

R672m Rocha, Tiago Rios da
Um modelo de anotações livres e semânticas para
bibliotecas digitais [dissertação] / Tiago Rios da Rocha
; orientador, Roberto Willrich. - Florianópolis, SC, 2010.
134 p.: il., tabs.

Dissertação (mestrado) - Universidade Federal de Santa
Catarina, Centro Tecnológico. Programa de Pós-Graduação em
Ciência da Computação.

Inclui referências

1. Ciência da computação. 2. Bibliotecas digitais.
3. Web semântica. 4. Anotações. I. Willrich, Roberto. II.
Universidade Federal de Santa Catarina. Programa de Pós-
Graduação em Ciência da Computação. III. Título.

CDU 681

UM MODELO DE ANOTAÇÕES LIVRES E SEMÂNTICAS
PARA BIBLIOTECAS DIGITAIS

Tiago Rios da Rocha

Esta Dissertação foi julgada adequada para a obtenção do título de Mestre em Ciência da Computação, Área de Concentração Sistemas de Computação e aprovada em sua forma final pelo Programa de Pós-Graduação em Ciência da Computação.

Prof. Dr. Mario Antonio Ribeiro Dantas
Coordenador do Curso

Banca Examinadora

Prof. Dr. Roberto Willrich (Orientador)
INE/UFSC

Prof. Dr. Renato Fileto
INE/UFSC

Prof. Dr. Ricardo Azambuja Silveira
INE/UFSC

Prof. Dr. André Luís Alice Raabe
UNIVALI

DEDICATÓRIA

*Aos meus familiares e amigos que sempre me apoiaram
e estiveram ao meu lado desde o início
na realização deste sonho.*

AGRADECIMENTOS

À minha família,
pelo incentivo e apoio
na busca do conhecimento,
em especial a minha mãe Liara Rocha
e avós Terezinha e Dalunier Rios
que me proporcionariam este momento;

À minha noiva Mayara de Oliveira Ribeiro,
por compreender minha ausência por tão longo período;

Ao meu tio Marcelo de Figueiredo Rios,
pelos conselhos de vida que me guiaram até aqui;

Ao meu orientador Dr. Roberto Willrich,
pela confiança em mim depositada,
pelos esclarecimentos e orientações,
pelas oportunidades concedidas
para que eu conseguisse seguir no mestrado;

Aos professores do mestrado,
pelos ensinamentos passados,
em especial ao professor Dr. Renato Fileto,
pela co-orientação durante
o desenvolvimento deste projeto,
e ao Dr. Alckmar Luiz dos Santos,
pelas oportunidades concedidas junto ao
Núcleo de Pesquisas em Informática, Literatura e
Linguística;

Aos meus colegas do mestrado e amigos,
pelo incentivo e amizade
em especial a Carla Lidiane Minello,
que sempre me deu forças nos momentos mais difíceis;

À Universidade Federal de Santa Catarina,
pelas condições oferecidas.

Ao Projeto SIDIE (Proc. CNPq N° 506789/2004-7),
pelo incentivo a pesquisa.

SUMÁRIO

Capítulo 1. INTRODUÇÃO	15
1.1 Objetivos.....	16
1.1.1 Objetivo Geral.....	16
1.1.2 Objetivos Específicos	17
1.2 Justificativa.....	18
1.3 Organização do Trabalho.....	19
Capítulo 2. BIBLIOTECAS DIGITAIS	20
2.1 Definição de Bibliotecas Digitais	20
2.2 Funções e Vantagens das Bibliotecas Digitais	22
2.3 Bibliotecas Digitais no Ensino.....	23
2.4 Biblioteca Digital de Literatura Brasileira (DL- LB)	25
2.4.1 Camada de Dados	26
2.4.2 Camada de Apresentação	27
2.4.3 Camada Lógica	27
2.5 Considerações Finais	27
Capítulo 3. ONTOLOGIAS.....	29
3.1 Conceito de ontologia	29
3.2 Classificações de ontologias	30
3.2.1 Classificação por nível de detalhe	31
3.2.2 Classificação por nível de dependência ...	31
3.2.3 Classificação quanto à função.....	32

3.2.4	Classificação quanto o grau de formalismo	33
3.2.5	Classificação quanto à aplicação	33
3.2.6	Classificação ao conteúdo	34
3.3	Linguagens de Especificação de ontologias	34
3.3.1	<i>Resource Description Framework</i> (RDF)	35
3.3.2	<i>Web Ontology Language</i> (OWL)	38
3.4	Considerações Finais	41
Capítulo 4.	ANOTAÇÕES DIGITAIS	42
4.1	Tipos de anotações	43
4.1.1	Anotações Não-Estruturadas	43
4.1.2	Anotações Semânticas	44
4.2	Esquema de anotação Annotea	45
4.3	Anotações em Bibliotecas Digitais	47
4.3.1	Requisitos para um sistema de anotações em DLs	48
4.3.2	Ferramentas Existentes	50
4.4	Considerações Finais	52
Capítulo 5.	SISTEMA DE ANOTAÇÕES EM BIBLIOTECAS DIGITAIS	54
5.1	Esquema de Anotação	54
5.2	Arquitetura do Sistema DLNotes	57
5.2.1	Interface e módulo de Administração	58
5.2.2	Servidor DLNotes	60
5.3	Avaliação das funcionalidades	67

5.4	Considerações Finais	68
Capítulo 6.	PROTÓTIPO DESENVOLVIDO	69
6.1	Diagrama de casos de uso.....	69
6.2	Diagrama de atividades.....	72
6.3	Diagrama de classes.....	73
6.4	Tecnologias Utilizadas.....	74
6.5	Ontologias utilizadas	79
6.5.1	Ontologia de domínio	79
6.5.2	Ontologia de anotação	81
6.6	Administração do protótipo	81
6.7	Manipulação de Anotações.....	81
6.7.1	Exibição do documento e inserção de anotações	82
6.7.2	Visualização de anotações	84
6.7.3	Criação de anotação	86
6.8	Possibilidades de uso do DLNotes no suporte ao ensino de literatura.....	89
6.9	Considerações Finais	92
Capítulo 7.	CONCLUSÕES	93
7.1	Trabalhos Futuros	95
7.2	Publicações	96
Capítulo 8.	REFERÊNCIAS BIBLIOGRÁFICAS.....	97
ANEXOS		104
ANEXO 1 – Diagrama de casos de uso.....		105
ANEXO 2 – Diagramas de atividades.....		107
ANEXO 3 – Diagrama de classes		130
ANEXO 4 – Ontologia de domínio da DL-LB		134

LISTA DE FIGURAS

Figura 1 – Arquitetura da DL-LB	26
Figura 2 - Tipos de ontologia, nível de dependência sobre uma tarefa particular. Fonte: Guarino (1997)	32
Figura 3 – Representação de uma declaração RDF	36
Figura 4 – Exemplo de uma declaração em RDF	37
Figura 5 - Exemplo de uma declaração detalhada em RDF	37
Figura 6 - (a) Ontologia apresentada em “arvore” apresentando a hierarquia entre as classes; (b) Ontologia expressa na linguagem textual de OWL.	40
Figura 7 – Ontologia de uma universidade representada por grafos dirigidos	41
Figura 8 – Esquema original do Annotea	46
Figura 9 – Extensão do esquema de anotação Annotea	55
Figura 10 – Exemplo de uma anotação utilizando o esquema Annotea estendido	57
Figura 11 - Modelo do sistema proposto	58
Figura 12 – Níveis de conhecimento no sistema proposto	63
Figura 13 – Anotação semântica no sistema proposto	66
Figura 14 – Arquitetura das tecnologias utilizadas no protótipo	78
Figura 15 – Parte da ontologia da DL-LB	80
Figura 16 – Classes da ontologia usada no esquema de anotação do protótipo	81
Figura 17 – Obra com anotações	82
Figura 18 – Ícones para diferenciar os tipos de anotações	83
Figura 19 – Menu de opções do DLNotes	83
Figura 20 – Interface de exibição de anotação livre	84
Figura 21 – Interface de exibição de anotação livre com uma seqüência de uma anotação e uma anotação resposta	85
Figura 22 – Interface de exibição de anotação semântica	86
Figura 23 – Interface de criação de anotação livre	87
Figura 24 – Interface de seleção de possíveis instâncias da base de conhecimento	88
Figura 25 – Interface de criação de anotação semântica	89
Figura 26 – Exemplo de atividade de identificação de personagens	
Figura 27 – Exemplo de atividade de identificação de figuras de linguagem	91
Figura 28 – Exemplo de atividade de identificação de ações e eventos	91

LISTA DE TABELAS

Tabela 1- Requisitos atendidos pelos sistemas comparados.....	53
---	----

LISTA DE ABREVIATURAS

AJAX	<i>Asynchronous JavaScript And XML</i>
API	<i>Application Programming Interface</i>
ASP	<i>Active Server Pages</i>
BC	Base de Conhecimento
DL	Biblioteca Digital
DL-LB	Biblioteca Digital de Literatura Brasileira
CREAM	<i>CREating Metadata</i>
CSCW	<i>Computer Supported Cooperative Work</i>
CSS	<i>Cascading Style Sheets</i>
DAML	<i>DARPA Agent Markup Language</i>
DC	<i>Dublin Core</i>
DCMI	<i>Dublin Core Metadata Initiative</i>
DILAS	<i>Digital Library Annotation Service</i>
DLESE	<i>Digital Library for Earth System Education</i>
DLMS	<i>Device Language Message Specification</i>
DLNOTES	<i>Digital Libraries Notes</i>
DOM	<i>Document Object Model</i>
GATE	<i>General Architecture for Text Engineering</i>
GPL	<i>General Public Licence</i>
HTML	<i>HyperText Markup Language</i>
IA	Inteligência Artificial
KAON	<i>KARlsruhe ONtology and Semantic Web tool suite</i>
LGPL	<i>Lesser General Public License</i>
NUPILL	Núcleo de Pesquisas em Informática, Literatura e Lingüística
OAI	<i>Open Archives Initiative</i>
OIL	<i>Ontology Inference Layer ou Ontology Interchange language</i>
OWL	<i>Web Ontology Language</i>
PDF	<i>Portable Document Format</i>
PEARL	<i>Process and Experiment Automation Realtime Language</i>
PHP	<i>Hypertext Preprocessor</i>
RAP	<i>Resource Allocation Protocol</i>
RASP	<i>Robust Accurate Statistical Parsing</i>
RDF	<i>Resource Description Framework</i>
RDFS	<i>Resource Description Framework Scheme</i>

RDQL	<i>RDF Data Query Language</i>
SAJAX	Simple Ajax Toolkit
SGBD	Sistema Gerenciador de Banco de Dados
SIDIE	Sistema de Disponibilização de Informações
para o Ensino	
SNOBASE	<i>Semantic Network Ontology BASE</i>
SIDIE	Sistema de Disponibilização de Informações
para o Ensino	
SOFA	<i>Simple Ontology Framework API</i>
SPARQL	<i>Sparql Protocol And RDF Query Language</i>
SQL	<i>Structured Query Language</i>
UML	<i>Unified Modeling Language</i>
URI	<i>Uniform Resource Identifier</i>
URL	<i>Uniform Resource Locator</i>
W3C	<i>World Wide Web Consortium</i>
XML	<i>Extensible Markup Language</i>

RESUMO

Bibliotecas Digitais (DLs) usualmente fornecem facilidades para navegação e pesquisas em suas coleções, e podem melhorar as atividades de aprendizagem via criação de conhecimento, compartilhamento de conhecimento e comunicação interativa entre os usuários. A integração de um sistema de anotação com DLs pode promover a troca de conhecimento entre instrutores e leitores e o desenvolvimento de serviços de infra-estrutura social em BDs. Ela oferece uma mudança na interação dos usuários com a DL, onde os usuários tornam-se ativos e visíveis colaboradores na criação de novos conhecimentos enriquecendo os conteúdos da DL. Esta dissertação apresenta um sistema de anotação, chamado DLNotes, que pode ser facilmente incorporado em DLs de modo a permitir anotações não estruturadas e anotações semânticas (baseadas em ontologia). DLNotes suporta também atividades de anotações e permite associar lista de discussão a cada anotação, o que é particularmente importante no contexto de aprendizagem baseada em tecnologia.

ABSTRACT

Digital Libraries (DLs) usually provide facilities for browsing and searching their collections, and can enhance learning activities through knowledge creation, knowledge sharing and interactive communication between users. The integration of an annotation system with a DL can promote the exchange of knowledge between instructors and readers. It offers a change in users' interaction with the DL, where users become active and visible contributors in the creation of new knowledge enriching the contents of the DL. Promotes the development of services for social infrastructure in DLs. This dissertation presents an annotation system, called DLNotes, which can be easily embedded in DLs in order to enable non-structured and semantic (ontology-based) annotations. DLNotes also supports supervised annotation activities and allows discussion threads to be associated with each annotation, what is particularly important for technology-based learning.

CAPÍTULO 1. INTRODUÇÃO

Bibliotecas Digitais (DLs) são ambientes que fornecem recursos para organizar coleções digitais e para construir um catálogo dos metadados que facilitam a organização e a descoberta de informações nessas coleções. O uso de DLs em aprendizagem eletrônica (*e-learning*) traz uma série de vantagens em relação a bibliotecas tradicionais, tais como (JAYAPRAKASH, 2006): (a) o acesso às informações independentemente de sua localização geográfica; (b) o acesso simultâneo aos documentos por parte de diversos usuários, vinte e quatro horas por dia, sete dias na semana; (c) a facilidade de atualização, impressão e cópia dos textos digitais; (d) os documentos digitalizados não sofrem danos pelo uso ou má condições de armazenamento; e (e) o fornecimento de informações atualizadas aos usuários.

Em *e-learning*, as coleções digitais das DLs são tratadas como conteúdos (ou objetos) de aprendizagem relacionada aos diferentes domínios do conhecimento, tais como física, química, matemática, biologia, literatura, etc. Os objetos dessas coleções representam conhecimentos explícitos que podem ser usados dentro de ambientes de aprendizagem para a produção ou agregação de novos conhecimentos.

Anotações são uma forma natural que as pessoas utilizam para gravar comentários e idéias sobre um documento. Por exemplo, quando pessoas lêem, elas freqüentemente sublinham ou escrevem notas na margem de um documento em papel. Estas atividades podem ser facilmente levadas para as DLs via sistemas de anotação. Existem diversos sistemas de anotações colaborativos, (por exemplo, ComMentor (HECK, 1999), DEBORA (NICHOLS et al. 2000) e CREAM (DAVIES et al, 2006)). Eles permitem que um conjunto de usuários crie, revise e compartilhe anotações de maneira colaborativa. O uso deste tipo de sistema tem como principal vantagem o enriquecimento de coleções. Este enriquecimento pode ser resultado de várias atividades de aprendizagem, tais como reflexões e análise de textos, as quais podem ser feitas de forma colaborativa. As informações adicionadas podem ser usadas por professores, alunos e pelo próprio sistema para aprimorar os mecanismos de busca personalizada. Com a inclusão de sistemas de anotação, espera-se que a DL não somente forneça informações aos usuários (no papel de leitores passivos), mas também permita que os usuários (no papel de leitores ativos, críticos ou

mesmo autores) criem informações sobre as obras que lêem e consigam organizar, classificar e reusar tal informação.

Existem diversos tipos de anotações. O tipo mais simples é a anotação não estruturada, onde os usuários podem selecionar e marcar uma parte do texto para a qual desejam criar uma anotação e então escrever um texto livre referente àquela marcação. Outra forma de anotação é a anotação semântica, que identifica precisamente entidades reconhecidas no texto, através de referências a suas descrições formais em uma ontologia (KIRYAKOV et al., 2004). A ontologia, por sua vez, é uma conceitualização explícita, formal e compartilhada de um universo de discurso (GUARINO, 1998).

As anotações semânticas podem ser utilizadas pelo sistema de gerência das coleções para a criação automática de relações entre elementos e para a implementação de sistemas avançados de recuperação de informações. Este tipo de anotação já vem sendo utilizado em DLs, como a proposta de (KHELIF et al., 2007).

É sabido que as grandes mudanças ocorridas na educação e na teoria pedagógica estão ligadas às transformações ocorridas nos meios de comunicação. Com a disseminação da Internet, modelos inéditos de educação baseada em *e-learning* vêm sendo desenvolvidos, gerando grande quantidade de projetos de pesquisa e *softwares* no contexto de Informática na Educação. A utilização de sistemas de anotações em DLs, assim como o uso de outras ferramentas de aprendizagem devem trazer para os educandos possibilidades de desenvolver habilidades na sua formação, estimulando a interatividade e a criação de inteligências coletivas.

O uso destes sistemas no contexto do ensino tem como foco promover o surgimento de inteligências coletivas onde o aluno participe ativamente na construção do conhecimento pessoal e coletivo, se envolva em processos colaborativos numa comunidade de aprendizagem e trabalhe com problemas abertos, onde escolha assuntos e defina o nível em que é capaz de desenvolver.

1.1 Objetivos

1.1.1 Objetivo Geral

O presente trabalho propõe um sistema para apoiar a criação colaborativa e o uso compartilhado de anotações em DL. Este sistema é

chamado de DLNotes e seu principal propósito é fornecer suporte para atividades de aprendizagem colaborativas em DL.

DLNotes tem como objetivo suportar atividades de *e-learning* através da colaboração e criação do conhecimento compartilhado entre os usuários da DL. Através de seu uso, usuários compartilham seus conhecimentos e trocam experiências a fim de construir colaborativamente um novo conhecimento comum, para isto, DLNotes proporciona aos usuários de uma DL lerem seus documentos possibilitando de forma privada criarem anotações em determinados lugares/peças dos documentos e o uso de anotações públicas, onde todas informações de anotações podem ser usadas de forma pública e proporcionar discussões online sobre questionamentos importantes sobre determinadas partes dos documentos, interpretações em grupo e intermediação de informações, podendo ser adicionados alguns links em determinadas partes do documento.

O DLNotes emprega ontologias do domínio específico da DL. Com base na ontologia de domínio, este sistema permite gerar anotações colaborativas segundo uma conceitualização padronizada. Com o DLNotes, os usuários podem identificar no texto dos documentos da DL instâncias de conceitos definidos em uma ontologia de domínio, ligar tais instâncias a marcações efetuadas no texto e relacionar essas instâncias em uma base de conhecimento. Assim, o sistema permite a criação de uma base de conhecimento acerca do domínio específico tratado pela DL.

Além de suportar a troca de conhecimento através do *e-learning*, o trabalho tem por objetivo utilizar o conhecimento criado pelos usuários para que pessoas e *softwares* consigam processar e interpretar as informações (aplicações, agentes de *software*). O sistema tem como objetivo trabalhar com dados interoperáveis para que sempre possa ser adicionado mais informações a base de conhecimento e outras aplicações possam utilizar tais informações. Outro objetivo é que o sistema seja de fácil instalação em diferentes DL, e que o sistema possa ser utilizado em DLs de diferentes domínios (por ex. biologia, computação, literatura).

1.1.2 Objetivos Específicos

Os objetivos específicos deste trabalho são os seguintes:

- Descrever teorias de bibliotecas digitais, sua importância e utilização;
- Descrever as definições de ontologias e as diferentes linguagens de representação de ontologias;
- Descrever os tipos de anotações digitais, ferramentas existentes bem como suas arquiteturas, funcionalidades e usos;
- Modelar um sistema de anotações em DLs levando em consideração as necessidades e oportunidades levantadas nos itens anteriores;
- Avaliar a proposta através do desenvolvimento de um protótipo para ser utilizado em uma DL;

1.2 Justificativa

As DLs facilitam o acesso e a leitura de uma grande quantidade de textos, com o decorrer dos últimos anos seu uso tem sido freqüente por muitos usuários. Anotações é uma forma natural que pessoas utilizam para criarem marcações ou comentários sobre partes de um texto, elas podem ser também utilizadas em textos digitais a fim de proporcionarem aos usuários de uma DL lerem seus documentos possibilitando de forma privada criarem anotações em determinados lugares/peças dos documentos para simplesmente sinalizar partes para uma leitura posterior, ou ainda criar comentários, gravar idéias, algum dado a mais no documento ou algum problema de trabalho (resposta a algum exercício). Além disso, proporcionam um compartilhamento de conhecimento, discussões online sobre questionamentos importantes sobre determinadas partes dos documentos, interpretações em grupo e intermediação de informações, podendo ser adicionados alguns links em determinadas partes do documento.

Os usuários podem identificar no texto dos documentos da DL instâncias de conceitos definidos em uma ontologia de domínio, ligar tais instâncias a marcações efetuadas no texto e relacionar essas instâncias em uma base de conhecimento. Assim, o sistema permite a criação de uma base de conhecimento acerca do domínio específico tratado pela DL. Tais anotações podem auxiliar na análise, classificação e recuperação das obras manipuladas na DL, manipulando conhecimento sobre conteúdo das coleções digitais.

Estes sistemas podem suportar atividades de *e-learning* através da colaboração e criação do conhecimento compartilhado entre os

usuários da DL. Através de seu uso, usuários compartilham seus conhecimentos e trocam experiências a fim de construir colaborativamente um novo conhecimento comum. Podem também, contribuir para o *e-learning* em um modelo de interação professor-aluno (formando - formando) numa perspectiva colaborativa, onde ambos contribuem para a construção compartilhada de conhecimento. O *e-learning* é frequentemente identificado como uma extensão da sala de aula na Internet, neste sentido, os sistemas de anotações podem abarcar situações de apoio tutorial ao ensino em um modelo de interação entre professor-aluno (formador - formando), em que o professor, sugere recursos e interage *on-line* com os alunos, fomentando debates e estimulando a colaboração. Podem contribuir para o ensino em um cenário de formação a distância permitindo concretizar abordagens pedagógicas baseadas na interação freqüente entre professor/aluno e na adoção de estratégias de trabalho colaborativo envolvendo alunos e professores.

1.3 Organização do Trabalho

O restante desta dissertação está organizado da forma que segue. O capítulo 2 apresenta Bibliotecas Digitais, sua definição, seu uso, vantagens e desvantagens. O capítulo 3 apresenta Ontologias. Sua definição, classificações, linguagens de ontologias. A definição de anotação, seus tipos e sistemas de anotações são explanados no capítulo 4, onde também são apresentados alguns sistemas de anotações em DLs e estes são comparados seguindo uma série de requisitos que se fazem necessários para um sistema de anotações em DLs. O capítulo 5 apresenta a proposta de um sistema de anotação em DL, chamado DLNotes, sua arquitetura e funcionalidades. O capítulo 6 apresenta o sistema DLNotes, sua modelagem, tecnologias utilizadas em sua implementação e o protótipo em execução. Finalmente, o capítulo 7 apresenta as conclusões e propostas para trabalhos futuros.

CAPÍTULO 2. BIBLIOTECAS DIGITAIS

Muitos exemplos de bibliotecas digitais (DL) podem ser encontrados na bibliografia atual. Seu rápido crescimento tem sido incentivado pela rápida publicação de materiais digitais e por projetos de digitalização de grande quantidade de conteúdos que originalmente eram disponibilizados somente de outras formas. DL é um termo genérico, que serve para diversas atividades. Biblioteca virtual, biblioteca eletrônica, biblioteca sem paredes e alguns outros termos são usados com uma conotação similar (ALENCAR, 2004). Neste capítulo serão apresentadas definições sobre este termo, suas diferentes interpretações, vantagens e desvantagens. Esta seção apresenta também a importância do uso de DLs no ensino. Por fim, é apresentada a Biblioteca Digital de Literatura Brasileira (DL-LB).

2.1 Definição de Bibliotecas Digitais

Uma biblioteca tradicional é aquela onde a maioria dos itens de seu acervo é constituída de documentos em papel, livros, revistas, etc. Para TEDD (2005), uma biblioteca tradicional não é somente constituída pelo depósito de documentos, mas também por uma quantidade de técnicas e serviços que são mantidos para atender seus usuários. Bibliotecas tradicionais utilizam técnicas como políticas de aquisições de coleções, que por sua vez atende as necessidades de seus usuários. Além disso, contempla múltiplas abordagens que facilitam ao usuário encontrar o que está procurando, por exemplo, um livro pode ser encontrado através de seu autor, título, assunto, etc. Esta organização envolve a disposição física dos documentos, que trata do seu armazenamento em prateleiras, e salas. Bibliotecas tradicionais disponibilizam ainda serviços que podem incluir uma ampla variedade de questões informativas e serviços especiais para os usuários. Círculos de leitura para crianças e entrega de livros. Finalmente, o espaço físico de bibliotecas tradicionais pode oferecer ambientes para estudo e trabalhos, disposições para reuniões, e assim por diante (TEDD, 2005).

O surgimento de DL pode ser considerado uma evolução de bibliotecas tradicionais, em virtude do aumento da quantidade de informações geridas pela biblioteca tradicional e a difícil atualização e recuperação de informação. Para tanto, é primordial uma preliminar definição de uma biblioteca tradicional.

Alguns autores entre 1945 e 1985 como Bush e Licklider, foram os visionários de DLs, eles imaginavam a necessidade de haver mudanças na biblioteca tradicional e a criação de uma biblioteca que utilizasse recursos tecnológicos para melhorar o acesso a informação, tinham uma noção que as bibliotecas no futuro trariam inovação, diferenças na estrutura, processamento, e acesso. Apesar das visões futurísticas, nas próximas duas décadas nada evoluiu. No final dos anos 80, as DLs eram mal uma parte das bibliotecas, e informática. Apenas por volta de 2000, a pesquisa e os desenvolvimentos práticos tiveram o interesse global (HAUBEN, 2004, LYNCH, 2005).

Existem várias definições para DLs, o termo pode significar desde coleções de textos publicados de forma eletrônica até projetos de digitalização de coleções concentrados em desenvolver a infra-estrutura necessária para efetivar o uso generalizado da informação em rede.

LESK (1997) define DL como “uma coleção organizada de informações digitais, com mesma estrutura e coleta de informações dos arquivos e bibliotecas tradicionais, com a representação digital que os computadores tornaram possíveis”.

AMÂNDIO (2007), as DL são consideradas como “um repositório de informação digital, dotadas de serviços integrados, devidamente organizados e descritos de modo a que essa informação se torne acessível mediante métodos de pesquisa”.

ATKINS (1997, p. 1), afirma que:

“Biblioteca digital não é uma mera equivalência a uma coleção digitalizada com ferramentas de administração de informações. É, certamente, um ambiente que reúne coleções, serviços, e pessoas para apoiar o ciclo completo de criação, disseminação, discussão, colaboração, utilização, nova autoria, preservação de dados, informações, e conhecimento. Os desafios e as oportunidades, que motivam uma iniciativa de pesquisa avançada sobre biblioteca digital, são associados com essa ampla visão do ambiente de biblioteca digital. Outras pesquisas sobre o assunto também irão explorar e ajudar a motivar os investimentos nas redes avançadas e na computação de última geração”.

Para SARACEVIC & DALBELLO (2001, p. 5), BORGMAN (1999) fornece uma definição completa de DL, uma definição que pode ser considerada uma ponte entre a comunidade de pesquisa e a comunidade pratica:

“Bibliotecas digitais são uma série de recursos eletrônicos e capacidades técnicas associadas para criar, buscar, e usar informações. (...) elas são uma extensão e aumento dos sistemas de recuperação e armazenamento de informações que manipula dados digitais em algum meio (...). O conteúdo de biblioteca digitais inclui dados, (e) metadados (...) Bibliotecas digitais são construídas, coletadas e organizadas por (e para) uma comunidade de usuários, e suas capacidades funcionais suportam a necessidade da informação e os usos desta comunidade” (SARACEVIC & DALBELLO, 2001, p. 5, *apud* BORGMAN, 1998, 1999).

Como afirmado anteriormente por BORGMAN (1999), AMÂNDIO (2007) e ATKINS (1997), DLs podem ser consideradas como uma extensão dos sistemas de recuperação e armazenamento de informações que manipula dados digitais, os conteúdos geridos por uma DL pode ser em diversos formatos, desde textos, sons, e imagens estáticas e dinâmicas. Os conteúdos de DLs incluem ainda dados e metadados, que descrevem vários aspectos dos dados (por exemplo: formato, criador, proprietário, direitos de reprodução, etc.) e metadados que consistem em *links* ou relacionamentos, entre outros, dados ou metadados internos ou externos à DL (BORGMAN, 1999). AALBERG (2003, p. 9), afirma que o grande desafio dos sistemas de bibliotecas digitais é ser capaz de lidar com um conjunto heterogêneo de documentos dentro de um mesmo ambiente.

2.2 Funções e Vantagens das Bibliotecas Digitais

DLs podem ter diversas funções (ALENCAR 2004): (a) Colecionar - no sentido de que DL disponibilizam coleções de documentos; (b) Reunir - no sentido de reunir documentos, para disponibilizá-los em formatos digitais; (c) Preservar Materiais - diz respeito ao cuidado e a possibilidade de conceder maior durabilidade aos documentos devido ao formato digital; (d) Usar - facilitar o uso e permitir a utilização das mesmas coleções e; (e) Armazenar - a

possibilidade que as DLs têm de guardar grande quantidade de diferentes materiais.

A utilização de DLs, trás diversas vantagens como (PROCÓPIO, 2004, p. 28): (a) Acesso 24 horas por dia, 7 dias por semana, 365 dias por ano; (b) Permite os mesmos dispositivos de direitos de propriedade dos livros impressos; (c) Permite adicionar mais títulos ao acervo já criado, sem a necessidade de investimentos em espaço físico, infra-estrutura ou dispêndio operacional com funcionários; (d) Retiradas, devoluções e recolocações automáticas nas prateleiras digitais; (e) Proteção contra roubo ou danificação de documentos; (f) Permite atender mais usuários com menos livros; (g) Fornece relatórios detalhados para analisar a utilização da biblioteca em níveis sem precedentes, melhorando a qualidade das decisões de aquisição; (h) O mecanismo de busca permite pesquisa de palavras em um livro ou em uma coleção inteira de livros; (i) É a solução mais adequada pra atender a alunos de cursos à distância, ou iniciativas de Inclusão Digital, que necessitam de acesso a uma biblioteca completa; (j) Suporte completo a todos os recursos de anotação e pesquisa.

Pode-se ser citadas as seguintes desvantagens do uso de DLs (PROCÓPIO, 2004, p. 28): (a) O excesso de informação cria redundância e perda de tempo; (b) Inexistência de infra-estruturas necessárias; (c) Perigo relacionado com os Direitos autorais e; (d) A complexidade dos sistemas informáticos pode levar à info-exclusão.

2.3 Bibliotecas Digitais no Ensino

Como definido anteriormente por BORGMAN (1998) e ATKINS (1997), DLs são recursos de informação que devem ser acessados e usados por uma comunidade de usuários, esta comunidade de usuários muitas vezes é formada por pesquisadores e estudantes que tem o mesmo interesse. Atividades de colaboração podem aproximar uma equipe de usuários e facilitar a interação científica (AGOSTI, 2004). "oferecem novas oportunidades para colaboração e comunicação que são inviáveis em bibliotecas tradicionais" (NICHOLS, 2000). A tecnologia disponibilizada em DLs permite novos meios de trabalhar com os objetos da biblioteca. Especialmente para trabalhos em grupo ou para usuários que possam contribuir de alguma forma para a DL ao invés de simplesmente usufruir das coleções.

Colaboração pode ser definida como uma atividade voluntária, estabelecida em fatores de motivação intrínseca, apela à autonomia e combina processos de trabalho individual e coletivo (MEIRINHOS, 2006, p. 3). Para a obtenção do objetivo comum através da colaboração, negocia-se e orienta-se a interação visando o objetivo comum, através do consenso onde pretende-se que cada membro, individualmente, seja responsável por atingir o objetivo do grupo. A colaboração em ambientes virtuais é um fenômeno cada vez mais freqüente, pode ser vista como uma estratégia pedagógica, mais também como uma filosofia de vida (HENRI, 2003 apud MEIRINHOS, 2006, p. 3).

A colaboração em DLs propicia uma maior interação e envolvimento da comunidade de usuários, pois ela apresenta-se como um processo que facilita a criação de comunidades e como um meio de partilha e construção de conhecimento no seio da comunidade potencializando processos de aprendizagem colaborativa (MEIRINHOS, 2006).

A definição de aprendizagem colaborativa, por sua vez, é norteadada de divergências, tais divergências surgem do dissentimento nas definições de aprendizagem. Para alguns autores como BOUTON & GARTH (1983) *apud* HILTZ (1998), a aprendizagem pode ser vista em uma abordagem passiva, onde o estudante aprende em assimilar o conhecimento de forma individual; Outros autores como MEAD (1934) *apud* HILTZ (1998), assumem que a aprendizagem se dá por um processo social que ocorre através da comunicação com outros, assumindo então, uma abordagem ativa. Dentre as diversidades de definições, aprendizagem colaborativa, para HILTZ (1998), “refere-se a métodos de ensino que incentivem os alunos a trabalharem juntos em tarefas acadêmicas”, ele sustenta que a aprendizagem colaborativa é fundamentalmente diferente da aprendizagem tradicional onde o professor é tomado como única fonte de informação. HILTZ (1998) completa que na aprendizagem colaborativa a instrução é centrada no aluno e não no professor, e a construção do conhecimento através da aprendizagem colaborativa é facilitada pela interação, avaliação e cooperação entre eles.

Da importância da aprendizagem colaborativa no fator social, MEIRINHOS (2006, p. 7) sustenta que:

“Aprendizagem colaborativa, alicerçada na atividade social, funciona como catalisador de dinâmicas de desenvolvimento da comunidade.

(..) A criação de um ambiente de maior envolvimento colaborativo pode ser indutor de dinâmicas sociais. (..) As relações de colaboração criam confiança, um ingrediente essencial para o intercâmbio e desenvolvimento de idéias”.

No contexto eletrônico de DLs, aprendizagem colaborativa, em alguns casos, pode ser tratada como aprendizagem eletrônica (*e-learning*). Aprendizagem eletrônica pode ser brevemente definida como “conteúdo instrucional ou experiências de aprendizagem entregues ou habilitadas por tecnologias eletrônicas” (JAYAPRAKASH, 2006, p. 2),

A razão de usar DLs no *e-learning* é que ela pode armazenar e gerenciar grandes quantidades de diversificados conteúdos digitais (textos, áudios, imagens e vídeos). Assim, fornecer um ambiente para reunir acervos, serviços e pessoas no suporte ao ciclo de vida completo de criação, disseminação e preservação de dados, informações e conhecimento (JAYAPRAKASH, 2006, p. 6).

2.4 Biblioteca Digital de Literatura Brasileira (DL-LB)

O projeto SIDIE (Sistema de Disponibilização de Informações para o Ensino) (PENEDO, 2005), teve como objetivo a construção de uma infra-estrutura de digitalização, organização e disponibilização de conteúdos digitais voltados ao ensino e pesquisa. O sistema constitui-se de um sistema de DL distribuída, composto de diversas DL, operando de forma integrada, formando um grande centro de documentação, direcionado ao ensino. Foi implementada usando a plataforma Web com o uso de plataformas de código aberto LAMP (Linux, Apache, MySQL e PHP).

Como parte do projeto SIDIE, foi desenvolvida a Biblioteca Digital de Literatura Brasileira (DL-LB)(WILLRICH et al., 2006). A DL-LB visa a disponibilização de uma grande coleção de obras literárias brasileiras de domínio público, assim como informações sobre escritores brasileiros e portugueses. Ela é uma DL voltada para o ensino, suporta serviços personalizados de recuperação de informações e distribuição seletiva de informações. Estes serviços utilizam como base um modelo de perfil do usuário que mantém informações pessoais, e preferências gerais e específicas ao tema tratado pela DL. A DL-LB suporta o

protocolo de comunicação OAI-PMH¹ e o conjunto de metadados DCMI².

A arquitetura da DL-LB, apresentada na Figura 1, segue um modelo a três camadas.

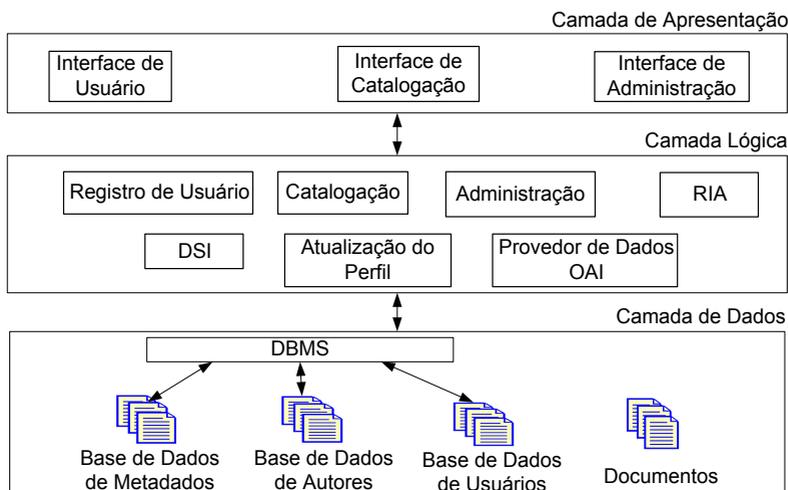


Figura 1 – Arquitetura da DL-LB

Fonte: (WILLRICH et al., 2006, p. 4)

2.4.1 Camada de Dados

A Camada de Dados é a base de dados, ou seja, a própria fonte dos dados e é composta por (PENEDO, 2005, WILLRICH et al., 2006):

- **Base de Dados de Metadados:** Mantém os valores de metadados utilizados para indexar e descrever as obras literárias disponibilizadas pela DL-LB. Foi adotado um grupo de metadados Dublin Core³ (DC) para descrever as obras literárias e metadados específicos para caracterizar obras literárias, como gênero literário, personagens e fatos históricos associados;

¹ OAI-PMH - *Open Archives Initiative*. Mais informações em: <http://www.openarchives.org/>

² DCMI - *Dublin Core Metadata Initiative*. Mais informações em: <http://dublincore.org/>

³ Dublin core - <http://dublincore.org/>

- **Base de Dados de Autores:** mantém as informações sobre os autores de documentos (nome completo, filiações, biografia, etc.);
- **Base de Dados de Usuários:** mantém o perfil de cada usuário registrado na DL-LB;
- **Documentos:** repositório que mantém obras literárias disponibilizadas na forma de páginas HTML.

2.4.2 Camada de Apresentação

A Camada de Apresentação é uma parte do código da aplicação que interage com o usuário. Esta camada é dividida em três partes (PENEDO, 2005, WILLRICH et al., 2006):

- **Interface de Usuário:** permite aos usuários realizar as operações como manutenção de perfil, busca, navegação e visualização dos documentos da coleção;
- **Interface de Administração:** disponibiliza aos administradores um conjunto de ferramentas que permitem o gerenciamento de usuários, metadados, e personalização de serviços;
- **Interface de Catalogação:** utilizada pelos colaboradores responsáveis pelo catálogo de novos autores e documentos, bem como pela inserção de arquivos digitais.

2.4.3 Camada Lógica

A Camada Lógica é responsável pelo acesso à camada de dados para recuperar, modificar e remover os dados, e enviar os resultados à camada de apresentação. Esta camada contém os módulos que implementam as funcionalidades oferecidas pela DL (PENEDO, 2005), WILLRICH et al., 2006).

2.5 Considerações Finais

Este capítulo procurou apresentar uma visão geral de DLs. Foram apresentadas as diferentes conceitualizações em torno da definição de DL, uma evolução em relação às tradicionais. Definições estas existentes desde que visionários imaginavam o uso de recursos tecnológicos para facilitar o acesso as informações na década de 1940. Este capítulo apresentou as funções e vantagens do uso de DLs e seu uso

no ensino, enfatizando e justificando sua importância no processo de ensino-aprendizagem. Por fim, foi apresentada a DL-LB, utilizada no contexto de ensino e pesquisa do projeto SIDIE.

O próximo capítulo trata de ontologias, utilizada neste trabalho para promover atividades de colaboração em DLs, interoperabilidade de informações, entre outros.

CAPÍTULO 3. ONTOLOGIAS

Nesse capítulo serão apresentadas as definições de ontologias, quais suas principais vantagens, classificações, áreas de aplicação e as principais linguagens de representação.

3.1 Conceito de ontologia

O termo ontologia, criado por Aristóteles, é originário da Filosofia (*ontos+logo*) que trata da natureza do ser (GUARINO, 1998, p. 2, GUIMARÃES, 2002, p. 49). No contexto da filosofia, filósofos tentavam descobrir o que é um ser, e quais são as características comuns de todos os seres. Para a comunidade de Inteligência Artificial (IA) e gestão de conhecimento, ontologia se refere a conceitos e termos usados para descrever alguma área do conhecimento. Ela pode ser assumida como um conjunto de entidades e suas relações, restrições, axiomas e vocabulário.

GRUBER (1993, p. 2) define ontologia como “uma especificação explícita de uma conceitualização”. Por sua vez, STUDER et al. (1998) complementa a definição dada por GRUBER (1993) explicitando o significado de algumas palavras: o termo conceitualização refere-se a ao conjunto de conceitos pertencentes ao domínio de interesse; o termo explícito significa que o tipo de conceitos e as restrições de utilização são explicitamente definidos, por exemplo, no domínio médico, os conceitos são doenças e sintomas, as relações entre elas são causas e uma restrição é que a doença não pode ser provocada por ela mesma; o termo formal significa que a ontologia deve ser passível de ser processada por máquina, excluindo a linguagem natural; compartilhada refere-se a fato que ela deve ser aceita por um grupo de pessoas, e não restrito a alguns indivíduos.

STUDER et al. (1998) afirma que ontologias podem ter várias formas, desde que tenha um vocabulário de termos e seus significados sejam de comum acordo com os participantes.

FREITAS (2003) conclui que a importância de uma ontologia é “descrever e representar uma área do conhecimento”, ele afirma ainda que ontologia é o sistema de representação do conhecimento de um determinado domínio.

Ontologias são utilizadas em áreas como gestão do conhecimento, comércio eletrônico, processamento de linguagens

naturais, web semântica, recuperação da informação, informática na educação, entre outros. É importante destacar as vantagens do seu uso (GUIMARÃES, 2002): (a) Fornecem um vocabulário para representação do conhecimento, evitando assim interpretações ambíguas do vocabulário; (b) Ontologias permitem o compartilhamento de conhecimento, ou seja, ela pode ser compartilhada e usada por aplicações do mesmo domínio; (c) Diferentemente da linguagem natural em que as palavras podem ter semântica totalmente diferente conforme o seu contexto, ontologias fornecem uma descrição exata do conhecimento.; (d) É possível fazer o mapeamento da linguagem da ontologia sem que com isso seja alterada a sua conceitualização, ou seja, uma mesma conceitualização pode ser expressa em várias línguas; (e) Pode ser possível estender o uso de uma ontologia genérica de forma a que ela se adeqüe a um domínio específico.

Para complementar as vantagens das ontologias apontadas por GUIMARÃES (2002), pode-se citar aquelas indicadas por STUDER (1998): (a) Ontologias são interessantes para facilitar a comunicação entre pessoas nas organizações. Elas fornecem os termos, seus significados, suas relações e restrições de um determinado domínio. (b) Em áreas como recuperação de informação e banco de dados, ontologias podem ser usadas para interoperabilidade de recursos de informação heterogêneos.

FREITAS (2003) além das vantagens citadas acima classifica como vantagem a possibilidade de tradução entre diversas linguagens e formalismos de representação do conhecimento.

3.2 Classificações de ontologias

Não existe um consenso entre os autores para classificação dos tipos de ontologias. Em alguns casos, autores distintos utilizam a mesma nomenclatura na classificação, mas significados distintos, enquanto outros autores utilizam diferentes nomenclaturas e mesmo significado. Em outros casos ocorre uma sobreposição das categorias propostas pelos autores. As ontologias podem ser classificadas de acordo com o grau de formalidade e níveis de detalhe do seu vocabulário, quanto sua estrutura e ao assunto da conceitualização, quanto sua função e aplicação, entre outras. As características gerais de uma ontologia são apresentadas na seção seguinte, através de uma breve revisão da literatura.

3.2.1 Classificação por nível de detalhe

GUARINO (1998) classifica ontologias quanto ao nível de detalhe em refinadas e não refinadas, definidas a seguir:

- Ontologias refinadas (*off-line*): este tipo de ontologia precisa de uma linguagem de alta expressividade e tem um grande número de axiomas. Para GUARINO (1998), ontologias deste tipo deveriam ser usadas *off-line* e somente para referência;
- Ontologias não-refinadas (*on-line*): Tem um número mínimo de axiomas e seu objetivo é ser compartilhada por usuários que concordem sobre uma determinada visão de mundo. Segundo GUARINO (1998), este tipo de ontologia tem mais chance de ser compartilhada, justificando o motivo de mantê-la *on-line*.

Por exemplo, as ontologias simples como enciclopédias podem ser mantidas *on-line*, enquanto teorias esclarecendo o significado dos termos da enciclopédia podem ser mantidas *off-line*.

3.2.2 Classificação por nível de dependência

GUARINO (1997), nesta dimensão, classifica ontologias entre ontologias de nível superior, ontologias de domínio, ontologias de tarefas e, ontologias de aplicação. A Figura 2 exhibe os diferentes tipos de ontologias conforme seu nível de dependência e a relação entre estas ontologias. Os conceitos de uma ontologia de domínio ou de tarefa devem ser especializações dos termos introduzidos por uma ontologia de nível superior, e os conceitos de uma ontologia de aplicação devem ser especializações dos termos das ontologias de domínio ou de tarefa.

O nível de capacidade de reuso e o nível de detalhamento da ontologia também estão especificados na Figura 2. A ontologia classificada como nível superior apresenta capacidade de reuso maior do que as outras, pois definem conceitos genéricos, em contrapartida a ontologia do tipo de aplicação tem o nível de especificidade superior que as outras classificações. Logo abaixo são detalhados os quatro tipos de ontologia.

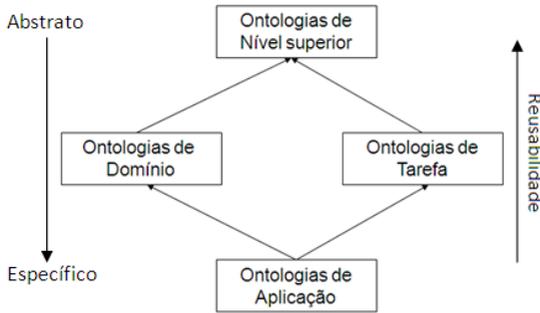


Figura 2 - Tipos de ontologia, nível de dependência sobre uma tarefa particular. Fonte: Guarino (1997)

Ontologias de nível superior ou ontologias de alto nível (*top-level ontologies*) descrevem conceitos muito gerais, como espaço, tempo, assunto, objeto, evento, ação, entre outros, os quais são independentes de um problema ou domínio em particular.

Ontologias de domínio (*domain ontologies*) e ontologias de tarefas (*task ontologies*) descrevem, respectivamente, o vocabulário relacionado a um domínio genérico (como medicina ou automobilismo) ou uma tarefa ou atividade genérica (como realizar o diagnóstico ou vender algo) pela especialização de termos introduzidos na ontologia de alto nível (*top-level*).

Ontologias de aplicação (*application ontologies*) descrevem conceitos dependendo ambos de um domínio e de uma tarefa em particular, o que são sempre especializações de ambas as ontologias relacionadas. Estes conceitos sempre correspondem a papéis desempenhados por entidades de domínio enquanto realizando certa atividade, como unidade substituível ou componente disponível.

Ontologias de representação constituem um tipo especial de ontologias de meta-nível (*meta-level ontologies*) descrevendo a classificação de primitivas usadas por uma linguagem de representação do conhecimento (como conceitos, atributos e relações).

3.2.3 Classificação quanto à função

MIZOGUCHI et al. (1995) classifica ontologias de acordo com sua função. Os autores identificam três tipos:

- Ontologias de domínio: são ontologias reutilizáveis e fornecem um vocabulário sobre conceitos do domínio em questão, sobre os relacionamentos sobre estes conceitos, sobre as atividades e regras dessas atividades;
- Ontologias de tarefa: fornecem um vocabulário sistematizado de termos utilizados na solução de problemas, especificando as tarefas;
- Ontologias genéricas ou gerais: incluem um vocabulário relacionado a coisas, eventos, espaço, tempo, comportamento, casualidade, funções, etc.

3.2.4 Classificação quanto o grau de formalismo

USCHOLD et al. (1996) classificam ontologias de acordo com o grau de formalidade utilizado para especificar o vocabulário de termos e seus significados. Os autores identificam quatro classes de ontologias:

- Ontologia informal: o vocabulário que é expresso em linguagem natural;
- Ontologia semi-informal: o vocabulário é expresso em linguagem natural de forma restrita e estruturada;
- Ontologia semi-formal: cujo vocabulário é expresso em linguagem definida formalmente;
- Ontologia formal: Os termos são definidos com semântica formal, teoremas e provas.

3.2.5 Classificação quanto à aplicação

USCHOLD et al. (1999) classificam as ontologias quanto a sua função no processo de desenvolvimento de sistemas computacionais. Os autores identificam três classes de ontologias:

- Ontologia de autoria neutra: que enfatiza a reutilização de dados, possibilitando que um aplicativo seja escrito em uma única linguagem e, depois, convertido para uso em diversos sistemas, reutilizando-se as informações;
- Ontologia de especificação: uma ontologia usada para documentação no desenvolvimento de software;

- Ontologia de acesso comum à informação: que torna a informação inteligível quando o domínio é expresso em um vocabulário inacessível, proporcionando conhecimento compartilhado dos termos.

3.2.6 Classificação ao conteúdo

VAN HEIJST et al. (2002) classificam as ontologias, quanto ao seu conteúdo, em:

- Ontologias terminológicas: especificam termos usados para representar o conhecimento em um domínio, por exemplo, os léxicos;
- Ontologias de informação: especificam uma estrutura de banco de dados, por exemplo, os esquemas registros;
- Ontologias de modelagem do conhecimento: especificam conceitualizações do conhecimento, têm estrutura interna semanticamente rica, e são específicas para o uso no domínio do conhecimento que descrevem;
- Ontologias de aplicação: contêm as definições necessárias para modelar o conhecimento em um software;
- Ontologias de domínio: expressam conceitualizações específicas para um determinado domínio do conhecimento;
- Ontologias genéricas: similares às ontologias de domínio, mas com conceitos considerados genéricos;
- Ontologias de representação: explicam as conceitualizações que suportam os formalismos de representação do conhecimento.

3.3 Linguagens de Especificação de ontologias

São inúmeras as linguagens disponíveis para a construção de ontologias: Ontolingua/KIF, CycL, Loom, Flogic, RDF(S), SHOE, XOL, OIL, DAML+OIL, OWL, etc. As seções abaixo apresentam algumas das principais linguagens de representação de ontologias existentes.

3.3.1 *Resource Description Framework* (RDF)

A RDF⁴ é um padrão, desenvolvido pelo grupo “*RDF Core Working Group*” como parte do projeto “*W3C Semantic Web Activity*” do W3C⁵ que fornece uma estrutura que permite que dados sejam compartilhados e reusados por aplicações, empresas e comunidade. RDF é uma linguagem desenvolvida para identificação de recursos semânticos na Web (W3C-RDF, 2004).

Seu desenvolvimento foi motivado pelos seguintes usos (W3C-RDF, 2004): (a) fornecer informações sobre recursos Web aos sistemas que os usam; (b) ser largamente usados por aplicações que exigem modelos de informações abertos (por exemplo, descrever processos organizacionais, anotar recursos Web, etc.); (c) Fazer com que aplicações processem informações fora do ambiente particular em que foi criado; (d) Possibilidade de combinar dados de diversas aplicações para descobrir novas informações e; (e) Possibilitar o tratamento automatizado de informações na Web por agentes de software.

O RDF é um modelo simples de dados independente da sintaxe de serialização, sendo facilmente manipulado e processado por aplicações, possui semântica formal que estabelece uma base confiável para raciocínio sobre os significados de suas expressões permitindo definição de regras de inferências, permitindo representar e compartilhar dados semânticos na Web (LASSILA, 1999). Usa vocabulário extensível baseado em URI⁶ e uma sintaxe baseada em XML⁷ onde é possível fazer afirmações sobre recursos, facilitando operações na internet (W3C-RDF, 2004).

O modelo de dados do RDF é caracterizado por prever três tipos de objetos (W3C-RDF, 2004):

- Recursos (*Resource*): Todas as coisas que estão sendo descritas por expressões RDF são chamadas de recursos. Recursos podem ser páginas Web (por exemplo, “<http://www.ufsc.br>”), partes dela, elementos XML de uma página, uma figura, e até mesmo objetos

⁴ RDF – *Resource Description Framework* – Mais informações em: <http://www.w3.org/TR/2004/REC-rdf-mt-20040210/>

⁵ W3C - *World Wide Web Consortium* – Mais informações em: <http://www.w3.org/>

⁶ URI - *Uniform Resource Identifier* - Um URI representa um endereço único para cada recurso na Web, sendo um identificador único para cada um deles.

⁷ XML - *Extensible Markup Language* – Mais informações em: <http://www.w3.org/XML/>

que não acessíveis pela Web (por exemplo, um CD). Os recursos são identificados por URIs, tudo pode ter um URI, a extensibilidade de URIs, permite a introdução de identificadores para qualquer coisa imaginável, por exemplo, uma pessoa chamada “Maria” pode ser representada pelo recurso “http://www.inf.ufsc.br/~maria”;

- Propriedades (*Properties*): São aspectos específicos, características, atributos ou relações utilizadas para descrever recursos. Cada propriedade tem um significado específico, define seus valores permitidos, os tipos de recursos que podem descrever, e sua relação com as outras propriedades. As definições dessas características são feitas através de RDF *Schema*;

- Declarações (*Statements*): Informação estruturada composta de sujeito (recurso), predicado (propriedade) e objeto (valor da propriedade). O objeto pode ser outro recurso ou um dado primitivo (p.e uma *string*).

Uma forma de representar declarações RDF é por meio de um grafo dirigido (fig. 3). Nesse tipo de diagrama, as elipses representam recursos, os retângulos com bordas arredondadas representam literais, e as setas representam propriedades. A seta sempre parte do sujeito e aponta para o objeto da declaração RDF. A propriedade também é representada por uma URI (não apresentado na figura 3 por motivos de simplificação).

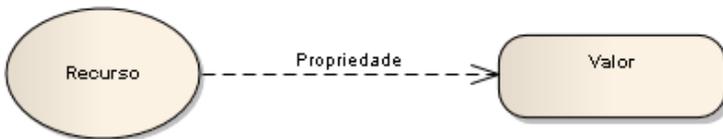


Figura 3 – Representação de uma declaração RDF

A figura 4 apresenta um exemplo de uma declaração RDF, onde é especificada a seguinte sentença:

"A obra Dom Casmurro tem Personagem Bentinho", ou de maneira mais geral "<recurso> TEM<propriedade><objeto>".

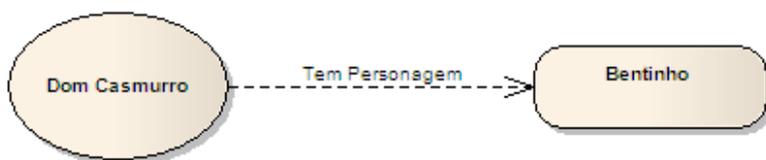


Figura 4 – Exemplo de uma declaração em RDF

É possível detalhar mais o objeto acrescentando diversas características a este. Por exemplo, a figura 5 representa a seguinte sentença:

“A Obra Dom Casmurro Tem Personagem identificado pelo recurso Bentinho que por sua vez, tem o nome Bento Santiago, é casado com Capitu e vivem em Rio de Janeiro”.

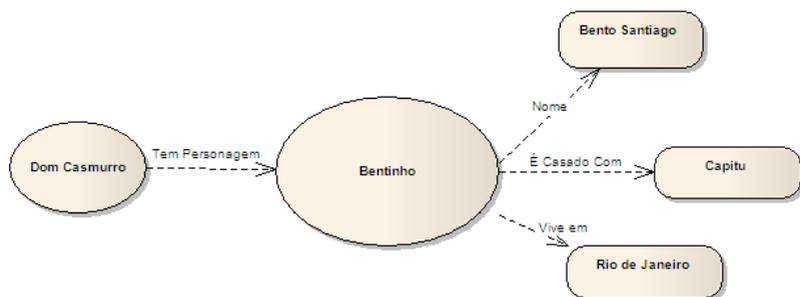


Figura 5 - Exemplo de uma declaração detalhada em RDF

O modelo de dados RDF possui mecanismos para a declaração de propriedades, no entanto, não possui mecanismos para a definição de relacionamentos entre propriedades e declarações, para isso definiu-se o RDF *Schema*. RDF *Schema* provê um conjunto de primitivas de modelagem para definição de ontologias e uma maneira padrão de codificá-las em XML (W3C-RDFS, 2004). É o mecanismo o qual é possível fornecer um sistema básico de tipos que pode ser usado nos modelos RDF.

Através do vocabulário proposto na especificação de RDF *Schema* é possível descrever conceitos e relacionamentos entre os conceitos (GUIMARÃES, 2003).

Os conceitos são descritos por meio de classes, a definição de classe é similar a definida em orientação a objetos onde uma classe é uma estrutura que abstrai um conjunto de objetos (indivíduos no contexto de RDF *Schema*) com características similares. Um exemplo de classe em um contexto automotivo poderia ser Carro, a classe Carro pode descrever características restritas a carros, como Modelo, Marca, Ano, Velocidade, e Combustível. Estas características são chamadas de propriedades no contexto de RDF *Schema*. Propriedades são as relações a serem construídas entre as classes e seus indivíduos (relação binária) ou para o indivíduo (função, característica). Por exemplo, a propriedade Ano, é uma propriedade que descreve uma característica do indivíduo da classe Carro, no entanto, a propriedade Combustível, pode ser uma relação binária com indivíduos da classe Combustível.

Com RDF *Schema* é possível fazer inferências sobre as informações contidas em um modelo RDF (GUIMARÃES, 2002), ou seja, ontologias concentram-se em métodos de classificação, em ênfase na definição de classes, subclasses e a relação com suas instâncias (indivíduos). O uso de inferências é caracterizado pela descoberta de novos relacionamentos com base nesses dados contidos, e algumas informações adicionais na forma de um vocabulário, por exemplo, um conjunto de regras.

3.3.2 *Web Ontology Language (OWL)*

A OWL⁸ é uma especificação de uma linguagem para ontologias desenvolvida pela W3C. A expressividade apresentada pela linguagem OWL é superior para representação de ontologias que XML, RDF, RDFS. (GONÇALVES, 2004). Ela foi projetada para o uso por aplicações que precisam processar o conteúdo da informação ao invés de apenas apresentá-la aos usuários. Isto, pois OWL fornece um vocabulário adicional (W3C-OWL, 2004).

A OWL provê mecanismos para a representação de classes, suas hierarquias e relacionamentos assim como RDFS. Restrições em propriedades (cardinalidade, igualdade, herança), características das propriedades (transitiva, funcional), anotações e instâncias (W3C-OWL, 2004).

⁸ *Web Ontology Language* - <http://www.w3.org/2004/OWL/>

OWL é dividida em três sub-linguagens (OWL *Lite*, OWL DL e OWL *Full*). Estas três sub-linguagens vão aumentando em expressividade (W3C-OWL, 2004):

- OWL *Lite*: é uma versão simplificada da OWL oferecendo classificação hierárquica e restrições simples;
- OWL DL: é utilizada por aqueles usuários que querem a máxima expressividade e ao mesmo tempo manter a computabilidade (garante-se que todas as conclusões sejam computáveis) e decidibilidade (todas as computações terminarão em tempo finito). Ela é um subconjunto da OWL *Full*, faz uso da lógica descritiva (*description logics*) de onde vem a extensão DL.
- OWL *Full*: mantém a expressividade, porém sem garantia de decidibilidade, uma vez que instâncias e propriedades podem ser especificadas como classes, diferentemente da OWL DL.

A figura 6 ilustra um simples exemplo de uma ontologia, que representa um domínio de uma universidade. Na coluna da esquerda da figura 6 a ontologia é mostrada como uma árvore, onde todas as relações entre os nodos são subclasse (*SubClassOf*), por exemplo, a classe “Pessoa”, tem duas subclasses, “Empregado” e “Estudante” que por sua vez, também possuem outras subclasses, como visto na figura. A coluna direita da figura 6, apresenta a mesma ontologia de uma maneira textual, é possível visualizar as definições de classes da ontologia, por exemplo, a primeira linha está definindo uma classe chamada “Academico”, que é subclasse da classe “Empregado”.

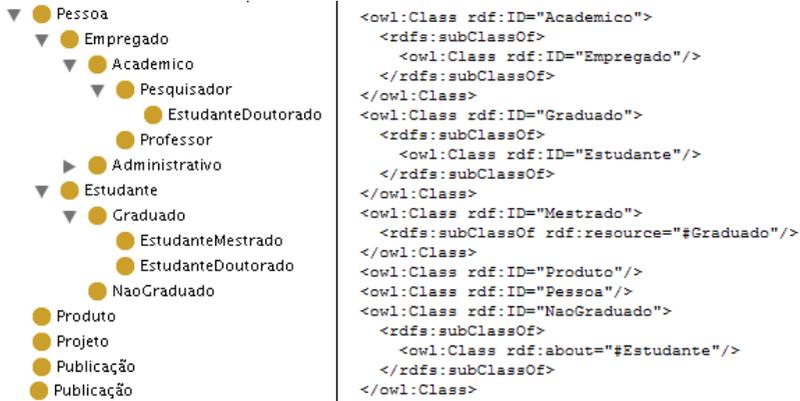


Figura 6 - (a) Ontologia apresentada em “arvore” apresentando a hierarquia entre as classes; (b) Ontologia expressa na linguagem textual de OWL.

As ontologias em geral são representadas por um grafo, não por uma arvore. Isto pois com grafos é permitida relações entre os nodos, podendo um nodo ter mais que um pai (superclasse). Neste exemplo, a classe “EstudanteDoutorado” é uma subclasse de duas classes, “Pesquisador” e “Graduado”. A visualização desta ontologia utilizando grafos dirigidos é apresentada na Figura 7.

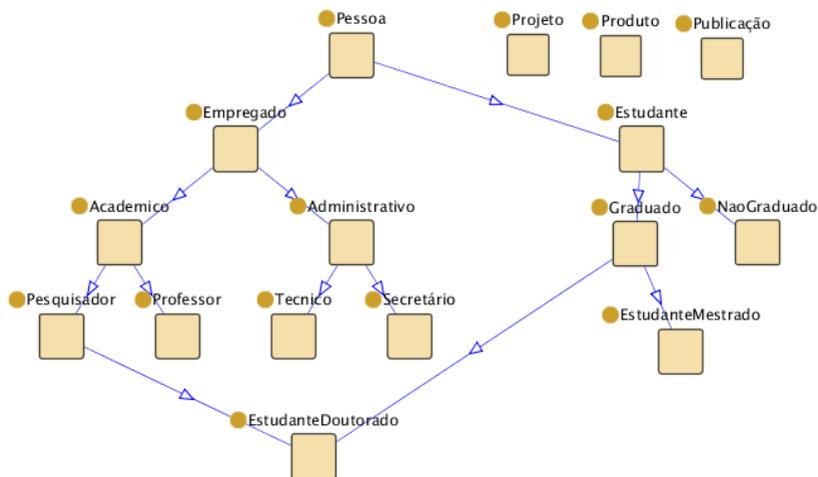


Figura 7 – Ontologia de uma universidade representada por grafos dirigidos

3.4 Considerações Finais

Este capítulo procurou apresentar uma visão geral de Ontologias. Foram apresentadas as definições em torno do tema e explanado seus usos e vantagens. Foram apresentados os diferentes tipos de classificações, logo, foram apresentadas algumas das principais linguagens de especificação de ontologias. A linguagem tratada no protótipo desenvolvido neste trabalho foi RDF/RDFS, podendo ser estendido para linguagem OWL, esta, com um poder de expressividade superior que RDF/RDFS.

O próximo capítulo apresenta os diferentes tipos de ferramentas de anotações, expondo algumas ferramentas existentes na literatura comparando com os requisitos definidos como necessários em um sistema de anotações em DLs.

CAPÍTULO 4. ANOTAÇÕES DIGITAIS

Anotações são uma forma natural que as pessoas utilizam para gravar comentários e idéias sobre um documento. Por exemplo, durante a leitura de um documento, pessoas freqüentemente sublinham ou escrevem notas na margem do documento em papel, estas anotações podem ser levadas para documentos digitais através de anotações digitais. MARSHALL (1998), afirma que utilizar anotações em conteúdos digitais é uma necessidade, ele ressalta a importância deste uso afirmando que em alguns casos, um livro anotado tem um valor maior do que um livro sem anotações, um exemplo disso, são os livros de direito que podem ser comprados com as anotações de um determinado autor.

Anotações tem sido freqüentemente proposta como uma técnica para que os usuários adicionem conteúdos, e assim, compartilhem idéias dentro de sistemas de informação (NICHOLS et al., 2000). Anotação digital é definida por WALLER (2003) como “um comentário sobre um recurso digital acessado como um todo, ou o conteúdo de um recurso, que pode ser acessado e armazenado digitalmente”.

Uma anotação pode ser aplicada em diferentes recursos, textos digitais, imagens, páginas da Web, bem como outras formas multimídia e até mesmo em DLs, são elementos da criação que se tornam fragmentos do documento. Anotações digitais podem ser aplicadas em diferentes níveis de granularidade como, partes do texto de um documento, parágrafos, partes de uma imagem, cenas de um vídeo, etc. (WALLER, 2003).

Anotações podem ser interpretadas em muitas maneiras: como um *link*, como uma construção de caminho, como comentário, como uma marca em torno de um texto, como uma descentralização de autoridade, como um registro de leitura e interpretação, ou como uma memória para comunidade (MARSHALL, 1998).

Esta seção apresenta as formas de representação e armazenamento de anotações, apresenta o esquema de anotação proposto pelo W3C, órgão responsável por padrões de criação e interpretação dos conteúdos da Web. Define os requisitos necessários para um sistema de anotações em DLs, por fim apresenta algumas comparações dos sistemas de anotações em DLs com os requisitos estipulados.

4.1 Tipos de anotações

Da mesma forma que as ontologias, existem diversas formas de classificar os tipos de anotações. Uma destas formas é quanto ao armazenamento da anotação, que podem ser intrusivas ou não-intrusivas. Anotação intrusiva é aquela armazenada no próprio documento, mas de forma a distingui-la do documento original e não corrompê-lo. A anotação não-intrusiva, por outro lado, é armazenada fora do documento anotado, usando ponteiros (por exemplo, via URI) para referenciar o (trecho de) documento a que se refere, sem causar quaisquer alterações no documento original.

A seguir é descrito duas classes de anotação de acordo com o modo de representação da anotação: anotações não-estruturadas e anotações semânticas.

4.1.1 Anotações Não-Estruturadas

Nos sistemas de anotação não estruturados, o usuário pode selecionar uma parte do texto o qual deseja criar uma anotação e então escrever um texto livre referente aquela marcação. Neste caso, as anotações geradas não seguem nenhum formalismo explícito. A principal desvantagem deste fato é a impossibilidade de uma interpretação automatizada do conteúdo das anotações e do que ela esta descrevendo. Este tipo de anotação é utilizado principalmente quando o usuário deseja criar um link para outro recurso em uma determinada parte do texto como, anotar uma possível interpretação ou uma marca para recordar posteriormente. As anotações não-estruturadas podem ser compartilhadas com outros usuários.

Existem vários sistemas de anotações não-estruturados, por exemplo, ComMentor (HECK, 1999) e Annotea (W3C-ANNOTE, 2002).

ComMentor desenvolvido pelo *Stanford Integrated Digital Library Project* em 1994 (HECK, 1999) e CoNote desenvolvido pela *Cornell University* em 1995 (HECK, 1999), são sistemas de anotações semelhantes, ambos foram criados para criar anotações na Web. ComMentor é implementado como um navegador (baseado no NCSA x Mosaic 2.4). Nestes sistemas as anotações podem ser públicas, privadas ou pertencer a um determinado grupo. Ambos suportam anotações em qualquer lugar na página e anotações de anotações, ou seja, nestes dois sistemas de anotação, um usuário pode criar uma anotação sobre uma

anotação criada por outro usuário criando assim uma sequência de discussão. As informações são armazenadas em servidores, para a inclusão das anotações nos documentos Web, o tráfego de informações deve passar por um *Proxy* que interceptará as informações e adicionará os *links* nos textos dos documentos Web. Ambos os sistemas não tiveram continuidade.

4.1.2 Anotações Semânticas

Os sistemas de anotações semânticas permitem ao usuário, ou um sistema automatizado, identificar itens de informações no texto e criar uma associação com os conceitos ou instâncias de uma ontologia. Esta associação elimina a ambigüidade e distingue a denotação da palavra ou expressão dos demais significados possíveis. O objetivo básico é a extração de conhecimento de uma coleção de documentos, seja na Web, em uma biblioteca digital, ou qualquer coleção de documentos. Por exemplo, quando encontrarmos as palavras “São Paulo” em um texto, esta pode representar diversos significados, como o nome de um estado, cidade, clube de futebol, ou ícone religioso.

O processo de anotação semântica gera uma base de conhecimento que pode ser interpretada pelos usuários e processada por sistemas computacionais, possibilitando a criação de sistemas de busca semântica, ou sistemas de criação e uso de metadados avançados.

A principal vantagem dos sistemas de anotação semântica é a possibilidade de formalmente anotar e ligar entidades e relacionamentos em documentos textuais, além de permitir a indexação e recuperação de documentos no que diz respeito às entidades/relações referidas. As anotações semânticas permitem novas aplicações para busca semântica, realces (*highlighting*), geração de metadados mais avançados, ligação de texto não organizado e base de conhecimento formal. Como principal desvantagem, é que anotações semânticas necessitam de manutenção constante, sob o risco de tornarem-se obsoletas.

Os sistemas de anotações semânticas podem ser classificados em três categorias dependendo do nível de interação com o usuário (TENIER et al, 2006): supervisionados, semi-automáticos e não supervisionados (automáticos).

Sistemas de anotações semânticos supervisionados são totalmente baseados na intervenção humana. Para tal, estes sistemas oferecem interfaces que mostram simultaneamente a ontologia e o

documento a ser anotado, possibilitando que o usuário marque instâncias de conceitos na ontologia. Como exemplos deste tipo de sistema pode-se citar Mangrove (UREN et al, 2006), é um sistema que permite o usuário, a partir de um navegador, fazer anotações semânticas em conteúdos Web. O objetivo é fazer com que os usuários anotem suas páginas com uma série de informações semânticas como por exemplo, membros universitários, projetos de pesquisa, eventos, calendário, etc.

Sistemas de anotação semi-automáticos são centrados no usuário para realização de algumas tarefas, e outras são realizadas de forma automatizadas, como extração e integração de informações. Um sistema que fornece um mecanismo para anotações automáticas de textos em diferentes domínios, desde que seja disponibilizada uma série de dados para treinamento. As informações disponibilizadas podem ser criadas manualmente pelo usuário. CREAM (DAVIES et al, 2006) é um exemplo de sistema de anotação semi-automático. Ele integra os conceitos de marcação fornecidos pelo Onto-O-Mat (que define metadados relacionais) com as marcações semânticas fornecidas pelo Amilcare.

Nos sistemas de anotação automáticos ou não supervisionados, o foco é manter o usuário fora do processo de anotação. Estes sistemas rastreiam a web e exploram a redundância de informação para encontrar anotações. Como exemplo, Armadillo (UREN et al, 2006) associa técnicas de extração de informação com um método de integração de informações estatísticas que confirmam a validade da informação encontrada. Magpie (DAVIES et al, 2006) é um sistema que anota páginas Web com metadados em uma forma inteiramente automática sem a intervenção manual para combinar o texto com instâncias da ontologia. O Sistema SemTag (UREN et al, 2006) executa anotações semânticas em grande escala na ontologia TAP, que contém cerca de 65 mil instâncias sobre diferentes domínios (música filmes, autores, esportes, etc).

4.2 Esquema de anotação Annotea

Annotea (W3C-ANNOTEa, 2002) é um protocolo de anotação criado pela W3C que modela anotações como metadados. No Annotea, as anotações são recursos Web de primeira classe e podem ser armazenadas em um ou mais servidores. O servidor de anotação armazena as anotações em uma base RDF. Os metadados são modelados através de um esquema RDF que fornece informações como data de

criação da anotação, nome do autor, etc. Este sistema assim como os anteriores, possibilita a criação de anotações públicas e privadas. As anotações privadas são armazenadas localmente no computador do usuário enquanto, as anotações remotas são armazenadas em um servidor. Annotea também permite a criação de anotações a partir de anotações e as anotações podem ser de diferentes tipos (comentário, Exemplo, Explicação, etc.).

A figura 8 apresenta o esquema RDF definido pelo Annotea.

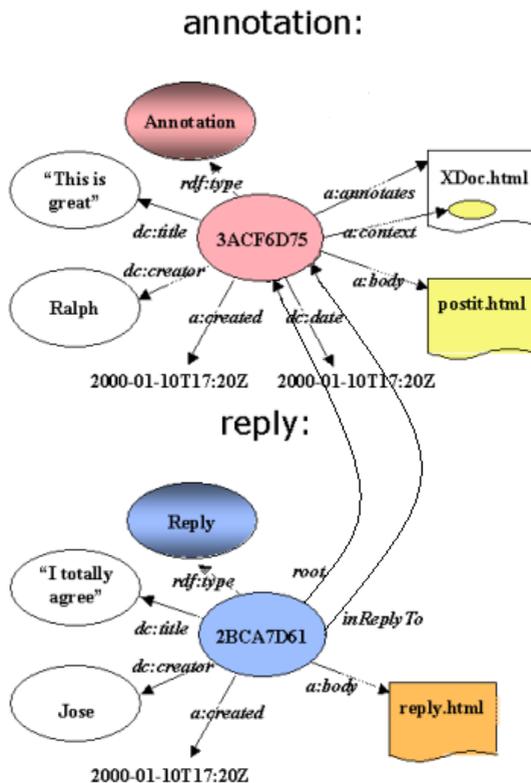


Figura 8 – Esquema original do Annotea

Fonte: KOIVUNEN & SWICK, 2003

A figura 8 exibe na parte superior o esquema de anotação definido pelo Annotea para a criação de anotação do tipo “*Annotation*”, na parte inferior da figura 8, o esquema adotado para anotações do tipo “*Reply*” (anotação resposta). Definido a seguir:

- A propriedade “*type*” indica a intenção do usuário ao fazer a anotação. Para as anotações do tipo “*Annotation*” o valor para esta propriedade podem ser os seguintes: “*Advice*”, “*Change*”, “*Comment*”, “*Example*”, “*Explanation*”, “*Question*”, “*SeeAlso*” e a própria “*Annotation*”. Anotações do tipo “*Reply*” podem receber os seguintes valores: “*Comment*”, “*Agree*”, “*Disagree*”, “*SeeAlso*” e o próprio “*Reply*”;
- A propriedade “*annotates*” refere-se ao documento anotado;
- A propriedade “*context*” especifica o lugar onde a anotação é referenciada no texto;
- A propriedade “*body*” refere-se ao conteúdo da anotação, o próprio texto descrito pelo usuário. Em alguns casos, seu valor pode ser um recurso que identifica a localização do conteúdo ou o próprio conteúdo.

Annotea usa múltiplos esquemas RDF a fim de padronizar suas propriedades, algumas propriedades são descritas pelo padrão Dublin Core:

- Propriedade “*dc:title*”, recebe o título da anotação;
- Propriedade “*dc:creator*”, define o autor da anotação;
- Propriedade “*dc:created*”, define o dia e hora em que a anotação foi criada.
- Propriedade “*dc:date*”, contém o dia e hora em que a anotação foi modificada;

Anotações do tipo “*Reply*” apresentam ainda duas propriedades não existentes em anotações do tipo “*Annotation*”:

- Propriedade “*root*”, indica a primeira anotação da seqüência de anotações;
- Propriedade “*inReplyTo*”, indica a anotação anterior a qual esta responde.

4.3 Anotações em Bibliotecas Digitais

A principal razão de utilizar bibliotecas digitais em *e-learning*, é a capacidade de armazenar e controlar grandes quantidades de

conteúdos digitais (por exemplo, textos completos, materiais de cursos, imagens, vídeos, áudios, etc.), utilizar varias ferramentas eletrônicas para encontrar informações utilizando avançada engenharia de busca, e avançada tecnologia de intercomunicação.

4.3.1 Requisitos para um sistema de anotações em DLs

O uso de anotações digitais pode expandir os recursos de informações geridos pelas DLs. A fim de suportar a expansão de informações, a colaboração entre os usuários e o uso de recursos de anotações, é proposto neste trabalho, uma série de requisitos para um sistemas de anotações em DLs, listados a seguir:

(i) O sistema de anotação deve adotar uma abordagem de integração flexível para ser facilmente integrado a DLs. É importante quando o objetivo é criar um sistema de anotação reutilizável em DLs diferentes, e não criar uma DL para apoiar a criação de anotações. Portanto, um sistema de anotação deve fornecer uma API oferecendo suas funcionalidades para a DL, minimizando as alterações na DL.

(ii) O sistema de anotação deve ser integrado em navegadores existentes e não deve exigir a instalação de *softwares* para os usuários. Isto é importante para simplificar o uso de sistemas de anotações e proporcionar que os usuários visualizem a mesma interface, não importando a máquina em que estão usando.

(iii) Anotações livres e anotações semânticas devem ser suportadas. Anotações livres permitem aos usuários associarem novos conteúdos (por exemplo, comentários, questionamentos, referências, etc.) para determinadas partes do documento que será anotado. Por sua vez, anotações semânticas permitem criar instâncias de conceitualizações explícitas, e seus relacionamentos contidos nos documentos.

(iv) O sistema de anotação deve permitir aos leitores dos documentos (usuários da DL) criarem anotações semânticas supervisionadas, a fim de partilharem seus conhecimentos. As anotações semânticas supervisionadas tem um custo elevado, são propensas a erros, e usualmente necessitam do especialista do domínio (Uren et al, 2006). Apesar destes inconvenientes, acredita-se que a anotação semântica semi-automática permite aos usuários criarem suas próprias anotações, a fim de partilharem seus conhecimentos.

(v) Para AGOSTI (2004), a criação de novos recursos de informação é suportada de duas maneiras, pela adição de informação aos recursos existentes e pela criação de novos recursos de informações a partir das anotações criadas. Deve ser possível um usuário incluir novas anotações a partir de anotações criadas por outros usuários, criando uma seqüência de anotações. Suportando, a partir disto, a comunicação e discussão entre os usuários da DL. A ordem seqüencial destas anotações deve ser mantida.

(vi) Diferentes níveis de anotações coexistem no mesmo documento: O Nível privado, onde as anotações deste nível são acessíveis somente pelo autor da anotação; O Nível público, onde as anotações podem ser compartilhadas por todos os usuários da DL. Deste modo, uma política de acesso deve ser associada às anotações, que podem ser públicas e privadas. A política de acesso deve indicar direitos de leitura e gravação. Além disso, as anotações públicas devem passar por um processo de revisão, para que anotações inválidas sejam retiradas.

(vii) O sistema de anotação não pode alterar o documento em si. As anotações devem ser armazenadas em um repositório separado do documento original.

(viii) O sistema de anotação deve ser facilmente usado pelos usuários da DL, que em geral não são especialistas em editores de ontologia.

(ix) O sistema de anotação deve armazenar as anotações livres e semânticas em sua base de conhecimento de forma que possibilite outros sistemas trabalharem com o conhecimento armazenado.

(x) O sistema deve permitir a transferência dos dados que descrevem os documentos armazenados na DL para a base de conhecimento. Esta transferência é importante, pois informações como autor, data de criação, título, etc., estão armazenadas na DL. Estes dados podem ser utilizados também para o enriquecimento da base de conhecimento.

(xi) O sistema deve possibilitar a cada usuário ter uma base de conhecimento privada, para suas próprias relações semânticas. O uso de base de conhecimento privada é útil para que cada usuário possa estabelecer as relações semânticas de seu conhecimento da forma que

melhor lhe convier. Posteriormente o usuário poderá disponibilizar o conhecimento armazenado para a base de conhecimento pública, ou vice-versa.

(xii) O sistema deve suportar atividades de aprendizagem eletrônica. Para dar tal suporte, o sistema deve possibilitar que instrutores (especialistas) proponham atividades de anotações a seus alunos. Uma atividade de anotação é identificada por um identificador único e uma lista de tipos de anotações possíveis que os alunos são autorizados a criarem em seus contextos. Por exemplo, o instrutor pode criar uma atividade de anotação que é criar somente anotações semânticas, e que as anotações que são criadas poderão somente utilizar parte da ontologia. O instrutor deve ser capaz de identificar que partes da ontologia fazem parte da atividade.

(xiii) As anotações podem ser ligadas no documento de duas formas: todo documento ou única. A primeira significa que a anotação ocorrerá em todas as vezes que o texto selecionado aparecer no documento. A segunda diz respeito a ligar a anotação para o texto selecionado somente naquela localização do documento.

4.3.2 Ferramentas Existentes

Existem diversas tecnologias desenvolvidas para dar suporte ao *e-learning*, mas existem poucos trabalhos que forneçam mecanismos para melhoria e enriquecimento das coleções de DLs. Esta sessão apresenta algumas destas. Os tipos de anotações adotadas em geral são textos livres, mas existem alguns trabalhos que utilizam anotações semânticas para DL.

NICHOLS et al. (2000) propõe um sistema, desenvolvido em Java, para anotação não-estruturada colaborativa em uma biblioteca digital de imagens obtidas de obras do Renascimento, chamado DEBORA. O usuário pode criar anotações selecionando determinados pontos da imagem. O usuário ainda tem a possibilidade de restringir a visualização de suas anotações por outros usuários ou compartilhá-la com um grupo de usuários. O sistema gera anotações ancoradas em posições na imagem da obra e não com o conteúdo da obra em si. Como já apresentado, as anotações não-estruturadas (livres) impossibilitam a busca semântica do conteúdo, portanto, não está de acordo com os requisitos (iv) e (iii). O software é desenvolvido em plataforma em que é exigida a instalação de software e apresenta uma interface de difícil

interação com usuários não especialistas, não seguindo os requisitos (ii) e (iii), respectivamente. As informações geridas pelo DEBORA são armazenadas em banco de dados, não possibilitando que outras aplicações utilizem este conhecimento armazenado, não suporta atividades de aprendizagem eletrônica e a discussão entre os usuários não é possibilitada, não atendendo os requisitos (ix), (xii) e (v), respectivamente.

ARKO et al. (2005) propõe um sistema de anotação estruturada para a biblioteca DLESE (*Digital Library for Earth System Education*). Este sistema não se baseia em uma ontologia. Nele, são definidos metadados de anotação que mantêm informações sobre o recurso educacional. Este conjunto de metadados de anotação é o mesmo para qualquer conceito sendo anotado, não é possível criar mais de uma anotação sobre um recurso já anotado, impossibilitando a discussão entre os usuários, não atendendo o requisito (v). As informações são armazenadas em banco de dados, e podem ser distribuídas via o protocolo OAI-PMH, DLESE. Possui um módulo de busca onde é possível encontrar documentos a partir dos metadados criados pelos usuários. DLESE não suporta anotações em partes do texto do documento, somente anotações sobre o documento, não atendendo a maioria dos requisitos estipulados na seção anterior ((iv), (iii), (v), (xi), (xii), (xiii)).

DiLAS (*Digital Library Annotation Service*) (AGOSTI et al. 2004, 2005) é um serviço de anotação que tem o objetivo de ser usado em diferentes sistemas de gerenciamento de bibliotecas digitais (DLMSs – *Digital Library Management Systems*). Ele define um conjunto de APIs para permitir o acesso aos serviços de anotação por diferentes DLMSs, este sistema adota somente o tipo de anotações não-estruturadas, as informações são armazenadas em banco de dados o que impossibilita a interoperabilidade entre diversas aplicações e não possui uma interface que faça a manipulação das anotações com os documentos, não atendendo os requisitos (iv), (iii), (ix), (x), e (xii).

MeatAnnot (KHELIF et al, 2007) é um sistema para geração de anotação semântica semi-automática desenvolvida para auxiliar biólogos a anotar seus documentos e facilitar a tarefa de recuperação de informação. Ela permite estender automaticamente a ontologia de biologia a partir das anotações semânticas extraídas de documentos textuais. MeatAnnot utiliza várias técnicas de processamento de linguagem natural (por exemplo, GATE, RASP), para extrair

informações do texto, instancias dos conceitos e relacionamentos na ontologia, e para gerar as anotações RDF do documento. Utiliza um sistema chamado MeatSearch que permite biólogos a fazerem inferências sobre as anotações. Os pontos fortes são a boa qualidade de extração de anotação automática e a possibilidade de fazer raciocínio sobre as anotações. A principal limitação deste trabalho é a dependência de trabalhar em um domínio específico de biologia (não atendendo o requisito (i)), a utilização somente por especialistas da área, não atendendo o requisito (viii) e a impossibilidade de criar anotações de forma manual de acordo com o nível de conhecimento do usuário e o não suporte a colaboração, não atendendo os requisitos (iii), (v), (vi), (xi), (xii).

Colate (Collaboratory for Annotation Indexing and Retrieval of Digitized Historical Archive Material) (BROKS et al, 2002) tem o objetivo de proporcionar a colaboração científica entre especialistas na área cultural. Produtores incluem documentos no sistema após isto, a catalogação por metadados, indexação e anotação é gerada colaborativamente entre os usuários. É possível a inclusão de anotações não-estruturadas em partes dos conteúdos e inclusive anotações aninhadas proporcionando a discussão. O usuário pode criar anotações de diferentes tipos como interpretação, argumentação, etc. Este software foi criado para um domínio específico, o que não proporciona o uso em diferentes DLs, não seguindo o requisito (i), não é possível a criação de anotações semânticas, foi implementado na plataforma Java sendo que necessária a instalação, não possui uma política de acesso as anotações, não atendendo os requisitos (iv), (ii), (iii) e (vi).

4.4 Considerações Finais

As anotações podem possuir diferentes interpretações, desde um simples *link*, construção de um caminho ou interpretação ou memória para a comunidade através de diferentes aplicabilidades nos diversos meios (texto, vídeo, etc.).

Este capítulo tratou das formas de representação e armazenamento de anotações, e o padrão utilizado pela W3C para o suporte a ambos os tipos de anotações.

Esta sessão diferenciou as conceitualizações de anotações livres e anotações semânticas e apresentou algumas ferramentas que proporcionam seu uso.

Após a exposição das diversas ferramentas de anotações foram definidos os requisitos necessários para um sistema de anotações em DLs. Logo, foram apresentados alguns sistemas de anotações para DLs e feito um comparativo com os requisitos definidos. O estudo comparativo pode ser resumido na Tabela 1, onde são apresentados os sistemas de anotações para DLs e quais requisitos são atendidos por cada sistema.

Tabela 1- Requisitos atendidos pelos sistemas comparados

Sistemas / requisitos	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII
DEBORA	X					X	X	X		X	X		X
ARKO	X	X				X	X	X	X	X			
DILAS	X	X			X	X	X	X			X		X
MEATANNOT		X		X			X	X	X	X			
COLATE				X	X		X	X	X	X	X	X	X

Com base na análise dos sistemas de anotações para DLs existentes, é notório que não há até o momento um sistema de anotações para DLs que atenda todos os requisitos expostos neste capítulo, para isto o próximo capítulo propõe um sistema de anotação para DLs, chamado DLNotes.

CAPÍTULO 5. SISTEMA DE ANOTAÇÕES EM BIBLIOTECAS DIGITAIS

Esta seção apresenta a proposta de um sistema para o suporte e manipulação de anotações em DL, chamado DLNotes. O objetivo é proporcionar aos usuários, identificarem passagens importantes no texto, desde simples marcações até anotações referentes a uma conceitualização explícita e formal. Além disso, a colaboração entre as atividades dos usuários proporcionará o compartilhamento de conhecimento através da reflexão, discussão e tomada de decisão em grupo.

Baseado no conhecimento adquirido através da troca de experiências entre os usuários, o sistema DLNotes permite a criação e manutenção de uma base de conhecimento, organizada segundo uma ontologia de domínio específico da DL tratada. A base de conhecimento criada pode servir de apoio para que estes dados possam ser processados tanto manualmente por usuários, quanto automaticamente por *softwares*, a fim de revelar novos relacionamentos possíveis entre porções de dados, ou no suporte a outros tipos de serviços, como por exemplo, serviço de busca semântica.

As seções que seguem neste capítulo, apresentam a proposta de um sistema de anotações em DL.

5.1 Esquema de Anotação

DLNotes adota uma extensão do esquema Annotea permitindo representar anotações não-estruturadas e anotações semânticas. A extensão do esquema Annotea é apresentado na figura 9. As setas direcionadas com cor azul são as propriedades incluídas ao esquema original.

Diversas propriedades apresentadas na figura 9 foram detalhadamente definidas na seção 4.2 onde foi apresentado o esquema de anotação definido pelo Annotea. A seguir serão definidas as propriedades incluídas ao esquema original:

- Propriedade “*access_control*” especifica se a anotação é pública ou privada. Anotações públicas são acessíveis e compartilhadas com todos usuários, por sua vez, anotações privadas são acessíveis somente ao usuário autor da anotação;
- Propriedade “*wholeDoc*”, indica se a anotação pertence a todas ocorrências do documento ou somente a uma localização indicada pela propriedade “*context*”;
- Propriedade “*anchor*”, esta propriedade armazena o texto selecionado pelo usuário quando ele cria a anotação. Note que a propriedade “*context*” especifica o lugar da anotação. A propriedade “*anchor*” é necessária para a criação dos links das anotações que podem aparecer em diferentes lugares no documento. (por exemplo, a anotação referente ao texto “Capitu” aparece em vários momentos em um documento, e todas, refere-se a mesma instância);
- Propriedade “*Activity*” indica a qual atividade de anotação ela pertence, esta propriedade é utilizada no contexto de *e-learning*. Uma atividade de anotação é definida com um identificador único e uma lista de tipos de anotações que os usuários autorizados podem criar. Utilizando este conceito, o instrutor pode definir uma série de atividades de anotações e submeter aos alunos. Esta propriedade é usada no contexto de anotação para possibilitar a recuperação das anotações de uma atividade de anotação específica.
- Propriedade “*review*”, criada para o controle de anotações inconsistentes, esta propriedade indica se a anotação foi aceita, rejeitada ou ainda não foi revisada.

Além disto, foi definido um novo tipo de anotação, chamado “*semantics*” para dar suporte à criação de anotações semânticas. Este novo tipo de anotação semântica possui uma propriedade chamada “*semantics*” que relaciona a anotação a uma instância de uma classe definida na base de conhecimento.

A figura 10 apresenta um exemplo de uma anotação não-estruturada usando o esquema de anotação proposto. Esta anotação é

identificada por #40309154541, é um comentário privado criado por Tiago Rocha no documento <http://www.dl.ufsc.br/text/01149.html>. O lugar desta anotação é especificado pela propriedade “*context*” e o texto selecionado foi “José”. O título desta anotação é “Título da anotação” e seu conteúdo é “Comentário da anotação”. Esta anotação foi criada durante a atividade de anotação “*Activity1*”.

```
<?xml version="1.0" encoding="UTF-8" ?>
<r:RDF
  xmlns:r="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:a="http://www.w3.org/2000/10/annotation-ns#"
  xmlns:d="http://purl.org/dc/elements/1.1/"
  xmlns:DINotes= "http://dlnotes.inf.ufsc.br/annotation-ns#"
  xmlns: DINotesSemantic= "http://dlnotes.inf.ufsc.br/annotation-ns/semantic#"

  <r:Description rdf:about="http://dlnotes.inf.ufsc.br/annotation-ns#40309154541"
    <r: type resource="http://www.w3.org/2000/10/annotation-ns#Annotation />
    <r:type resource="http://www.w3.org/2000/10/annotationType#Comment"/>
    <a:annotates r:resource="http://www.dl.ufsc.br/text/01149.html"/>
    <a:context>#xpointer(/html[1]/body[1]/ul[5]/li[4]/a[1])</a:context>
    <d:title>Título da anotação</d:title>
    <d:creator>Tiago Rocha</d:creator>
    <a:created>2009-04-03T12:10Z</a:created>
    <a:body>conteúdo da anotação</a:body>
    <DINotes:privacy>private</DINotes:privacy>
    <DINotes:anchor>José</DINotes: anchor>
    <DINotes: wholeDoc>>false</DINotes: wholeDoc>
    <DINotes:activity>Activity1</DINotes: activity>
  </r:Description>
</r:RDF>
```

Figura 10 – Exemplo de uma anotação utilizando o esquema Annotea estendido

5.2 Arquitetura do Sistema DLNotes

Sistemas de anotações Web podem ser divididos em três grupos baseado em sua implementação (BRUSH, 2002): baseado em servidor, em *proxy* ou em uma extensão do navegador Web. Em abordagens baseadas em servidor, somente documentos do servidor podem ser anotados; Abordagens baseadas em *proxy*, os documentos a serem anotados são acessados por um *proxy*, tipicamente um URL especial, que fornece capacidades de anotação; Por fim, abordagens baseadas em uma extensão do navegador Web, fornecem recursos de anotação ao navegador Web ao estendê-lo com a instalação de um *software*.

DLNotes destina-se em anotar documentos HTML que compõe uma coleção disponibilizada em um servidor Web. A adoção de uma abordagem baseada em servidor é a mais natural, pois, não requer a instalação de um software ou uma configuração especial no navegador do usuário. Além disso, simplifica o compartilhamento de anotações entre os usuários da DL.

A Figura 11 ilustra a arquitetura global do sistema DLNotes. Seus componentes principais são descritos nas subseções a seguir.

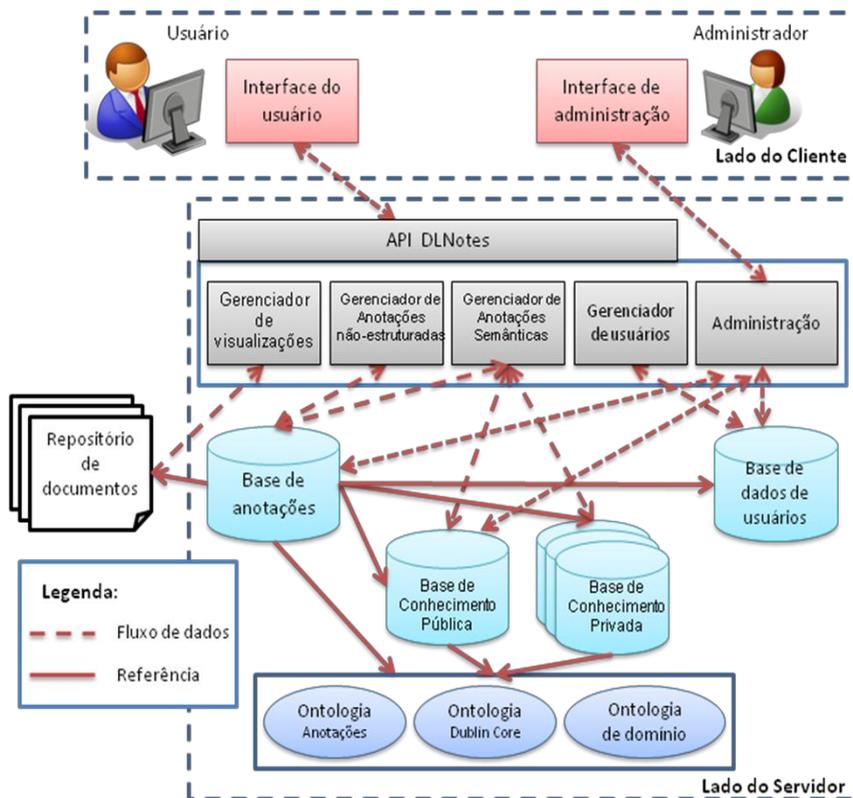


Figura 11 - Modelo do sistema proposto

5.2.1 Interface e módulo de Administração

O papel de administrador no sistema DLNotes é dividido em dois tipos: administrador geral, responsável pela administração completa

do sistema e; Administrador especialista, responsável por gerenciamentos específicos, detalhado a seguir.

A interface e o módulo de administração possuem funcionalidades que permite os administradores executarem as seguintes funções:

- Criar e gerenciar atividades de anotações. Para isto, os administradores devem indicar quais atividades serão permitidas no contexto de uma atividade de anotação, caso seja incluída nesta atividade, anotações do tipo semântica, os administradores devem indicar quais classes da ontologia de domínio pertencem a atividade. Por fim, selecionar os usuários participantes da atividade;
- Verificar e manter a consistência e segurança das anotações. Estas funcionalidades permitem aos administradores fazerem uma validação das anotações criadas pelos usuários a fim de rejeitar as anotações inconsistentes e tornar públicas as anotações aprovadas;
- Gerenciar usuários. Esta atividade é somente disponibilizada ao administrador geral do sistema, ele manterá os cadastros dos usuários, podendo cancelar a conta de determinados usuários e manter (incluir/retirar usuários) a lista de usuários especialistas.

A Interface de administração permite adaptar o sistema de anotação ao domínio da DL. Somente o administrador geral pode executar tais funções. Esta adaptação envolve os passos descritos abaixo:

- Importação da ontologia de domínio: o administrador indica a ontologia de domínio a ser considerada pelo sistema de anotação. A ontologia de domínio pode ser composta por diversas ontologias;
- Seleção dos conceitos “anotáveis”: o administrador escolhe, dentre todos os conceitos da ontologia de domínio, quais podem ser utilizados na criação de anotações semânticas. Funcionalidade semelhante à de criar atividades de anotações, mas esta é utilizada em um contexto global, ou seja, por todos usuários que não estão em uma atividade de anotação.
- Geração da base de conhecimento: a base de conhecimento é gerada e armazenada em forma persistente no banco de dados. Para tal, o administrador informa os dados para acesso do sistema

ao banco de dados, para que a base de conhecimento seja criada e armazenada em forma persistente, as informações necessárias para isto são informações como nome de usuário, senha e nome da base de dados.

- Povoamento da base de conhecimento com metadados da DL: esta etapa tem o objetivo de povoar a base de conhecimento a partir das informações já adquiridas e classificadas pela DL. Duas opções são suportadas: importação dos metadados DC utilizando o protocolo OAI-PMH ou usando aplicação padrão para a coleta direta dos metadados.

5.2.2 Servidor DLNotes

Como apresentado na Figura 10, o servidor DLNotes contém uma API e vários módulos que implementam suas funcionalidades, detalhados a seguir.

5.2.2.1 API DLNotes

A API DLNotes oferece métodos necessários para criação e visualização das anotações separados em diversos módulos, detalhados a seguir.

No momento que o usuário solicita a apresentação de um documento à DL, é então invocado um método que inicia o uso da API DLNotes. Para esta invocação, é necessário alterar o código fonte da DL, de maneira a alterar o link de acesso ao documento digital para a API DLNotes, no momento desta chamada, dois argumentos devem ser passados, a URL do documento a ser anotado e a atividade de anotação, caso exista. Após esta invocação, é necessário que o usuário acesse o sistema através de sua conta, caso não possua, será direcionado para o módulo gerenciador de usuários, onde poderá se cadastrar.

A partir da identificação do usuário, o servidor DLNotes consulta a base de usuários para verificar as preferências de visualizações das anotações, esta informação está armazenada no perfil do usuário. O sistema classifica três níveis de visualizações (Sem visualização das anotações, visualização somente de suas anotações, visualização de suas anotações e as anotações públicas de outros usuários). Em seguida, é criada a interface de anotação apresentando o documento solicitado pelo usuário, e inseridas dinamicamente as anotações no documento.

Durante a sessão de visualização do documento e das anotações na interface do usuário, o servidor DLNotes interage com a interface através de sua API. Foi definido uma série de métodos que permitem criar, modificar e deletar anotações não-estruturadas e semânticas.

Modulo gerenciador de usuários e base de dados do usuário

O módulo gerenciador de usuários implementa todas funcionalidades relacionadas com a base de dados do usuário, incluindo registro de usuários, autenticação de usuários e atualização de perfil. A base de dados do usuário mantém as informações cadastrais e de perfil de cada usuário da DL. Neste perfil são mantidas as informações de autenticação do usuário, sua identificação no sistema, e suas preferências e privilégios em termos de anotações. As preferências indicam se o usuário deseja visualizar as anotações pessoais e/ou públicas, e se ele deseja criar anotações. O campo privilégio indica se o usuário pode gerar ou não anotações públicas sem passar por um processo de validação por parte dos administradores. Este privilégio pode ser dado, por exemplo, para professores e especialistas no domínio da DL.

Modulo gerenciador de visualizações

Este módulo altera o código HTML do documento a fim de incluir os *links* das anotações. Este módulo é invocado após a autenticação do usuário. Neste momento, é verificado as preferências de visualização no perfil do usuário, se necessário, é incluído dinamicamente os *links* das anotações no documento. Os *links* são compostos pelo próprio texto da anotação e um ícone que distingue os diferentes tipos de anotações.

Cada *link* da anotação pode ser incluído no texto de duas formas: em todas ocorrências de uma determinada *string* (frase/texto) no documento, ou somente em um determinado local do documento.

Quando o usuário clica sobre o *link* de uma determinada anotação, uma janela é exibida sobre o texto do documento contendo os detalhes da anotação recuperados dos módulos gerenciadores de anotações não-estruturadas e semânticas.

Módulos gerenciadores de anotações não-estruturadas e semânticas

Estes módulos são responsáveis pela criação, recuperação, alteração e exclusão de anotações não-estruturadas e semânticas.

O usuário pode criar uma anotação sobre qualquer parte do documento. A criação de anotações deve ser facilitada para usuários não experientes. Um possível método simples para o sistema identificar a intenção que o usuário tem de criar uma anotação, é utilizar o clique do *mouse*, por exemplo, se após o usuário selecionar a *string* (frase/texto) a qual deseja criar uma anotação, clicar com o botão direito do *mouse*, o sistema pode identificar esta ação e exibir uma interface que permite a geração de anotação. Durante o processo de criação de anotação, o usuário deve selecionar o tipo de anotação que ele deseja criar (anotação não-estruturada ou semântica).

A anotação não-estruturada permite que o usuário faça alguma marcação sobre o texto selecionado, ou ainda, inclua algum comentário, questionamento ou referências extras.

Anotação semântica é o ato de ligar instâncias encontradas no texto do documento, a seu referido conceito contido em uma ontologia, como ilustrado na Figura 12.

Uma ontologia é composta por conceitos (classes) e propriedades (relação binária ou função, característica, detalhadas na seção 3.3.1). No exemplo da figura abaixo, o conceito “Pessoa” tem duas propriedades (relação binária): “é_pai” e “é_filho”. A propriedade denominada “é_pai” é uma propriedade que relaciona a própria instância com outra instância do conceito “Pessoa”, por exemplo, a instância do conceito “Pessoa” identificado por “Pedro Antunes” possui uma relação binária do tipo “é_pai” com a instância do conceito “Pessoa” identificado por “Maria Antunes”. A segunda propriedade da Figura 12 é semelhante, porem, apresenta a ordem inversa entre sujeito e objeto.

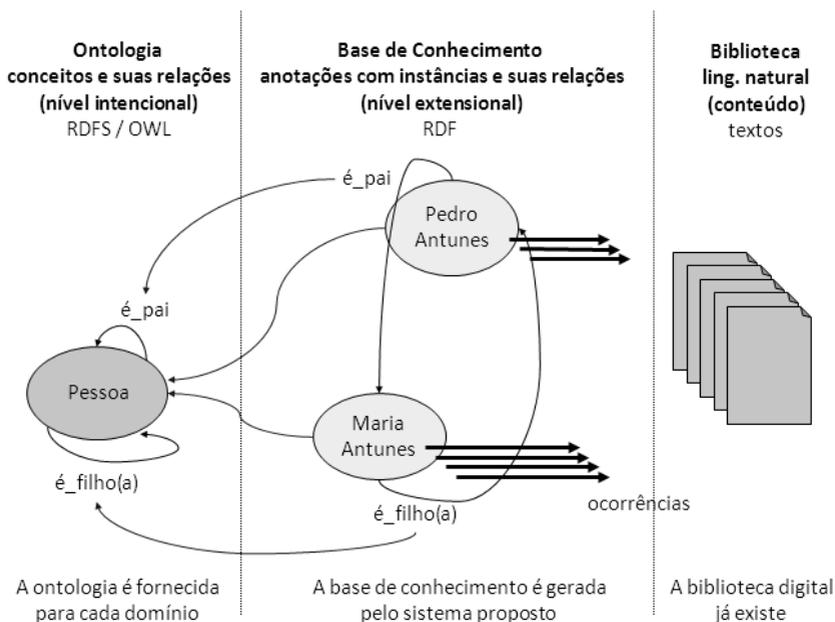


Figura 12 – Níveis de conhecimento no sistema proposto

O processo de criação de uma anotação semântica pode ter a seguinte seqüência: no momento que o usuário solicita a criação de uma anotação semântica sobre uma *string* selecionada, o módulo responsável procura por alguma possível instância referente à *string* selecionada na base de conhecimento, exibe para o usuário a lista das possíveis instâncias, caso exista, o usuário pode identificá-la, e alterá-la, ou então criar uma nova instância referente à *string* selecionada

Por exemplo, o usuário seleciona a *string* “Catarina” em um documento e solicita a criação de uma anotação semântica, automaticamente o sistema busca todas as instâncias desta *string* na base de conhecimento e apresenta ao usuário. Alguns conceitos candidatos poderiam ser: “Santa Catarina – é um Estado”. “Catarina Silveira – é um Personagem”, “Universidade Federal de Santa Catarina é uma Instituição de ensino”. Se a *string* “Catarina” a qual o usuário deseja anotar é alguma exibida pelo sistema, o usuário pode selecioná-la e visualizar o que há cadastrado sobre aquela instância, completar e alterar os dados contidos, caso não seja nenhuma das instâncias listadas, o usuário poderá identificar que a *string* “Catarina” pertence à outra classificação e então será direcionado para criar uma nova anotação

referente à *string* “Catarina” podendo instanciá-la há algum termo da ontologia de domínio disponível.

Concluída a etapa de criação de anotação, o módulo responsável salva as informações na base de conhecimento. Se a anotação for do tipo pública (disponível para ser visualizada por todos usuários), ela não é visualizada por outros usuários até que o especialista faça a verificação de consistência e segurança da anotação. Se a anotação for do tipo privada, ela será armazenada em uma base de conhecimento privada do usuário. A base de conhecimento privada permite ao usuário estabelecer relações do texto com os conceitos da ontologia de maneira que somente ele visualize estas relações, podendo posteriormente transferir esta base de conhecimento privada para a base de conhecimento pública.

Ontologias

O DLNotes é um sistema de anotações baseado em ontologias, onde as ontologias são representadas usando RDFS (RDF *Schema*). Três ontologias são definidas:

- Ontologia de anotações: esta ontologia representa o esquema de anotação proposto pelo projeto Annotea e a extensão proposta na seção 5.1;
- Ontologia Dublin Core (DC): é uma ontologia predefinida usada para anotar os documentos da DL. No sistema, cada documento é representado por uma instância da classe “*dc:documents*” com seus elementos DC. Algumas propriedades definidas nesta ontologia também são utilizadas pela ontologia de anotações (“*dc:title*”, “*dc:creator*”, “*dc:date*” e “*dc:created*”);
- Ontologia de domínio: especifica os conceitos do domínio e suas relações específicas tratadas pelo DLNotes durante a criação de anotações semânticas. A ontologia de domínio pode ser composta por várias ontologias. O desenvolvimento desta ontologia deve ser feito por especialistas da área de domínio.

Base de anotações, bases de conhecimento pública e privadas

DLNotes armazena as anotações em uma base de anotações separado dos documentos da DL. Com a criação de anotações semânticas, usuários contribuem para a composição de uma base de conhecimento sobre o domínio tratado pela DL. Esta base de

conhecimento contém instâncias das classes e suas propriedades especificadas na ontologia de domínio.

A base de conhecimento pública mantém todas as propriedades geradas pela importação de metadados da DL. Além disso, contém instâncias e propriedades das classes criadas durante a geração de anotações semânticas públicas.

As bases de conhecimento privada contém instâncias e propriedades das classes criadas durante a geração de anotações semânticas privadas. Neste caso, cada usuário tem sua base de conhecimento.

No sistema proposto, a ontologia e as bases de conhecimento fornecem um vocabulário, especificando conceitos, instâncias e suas relações. Para anotações semânticas, o sistema permite associar fragmentos do documento com conceitos ontológicos ou instâncias. Se um novo termo ou entidade nomeada é identificado por um usuário, este pode criar uma nova anotação semântica, onde será criada uma instância na base de conhecimento. Por exemplo, na Figura 14, o usuário pode criar uma anotação semântica sobre a entidade nomeada Romeu e Julieta criando duas instâncias do conceito Personagem na base de conhecimento. A figura 14 exibe também as diferentes localizações das informações geridas pelo sistema. A descrição do conceito “Personagem” está na ontologia de domínio, e a descrição do conceito “Semantic” (anotação semântica), está na ontologia de anotações. A base de anotações mantém as informações referente as anotações e as informações quanto a instâncias semânticas estão armazenadas na base de conhecimento.

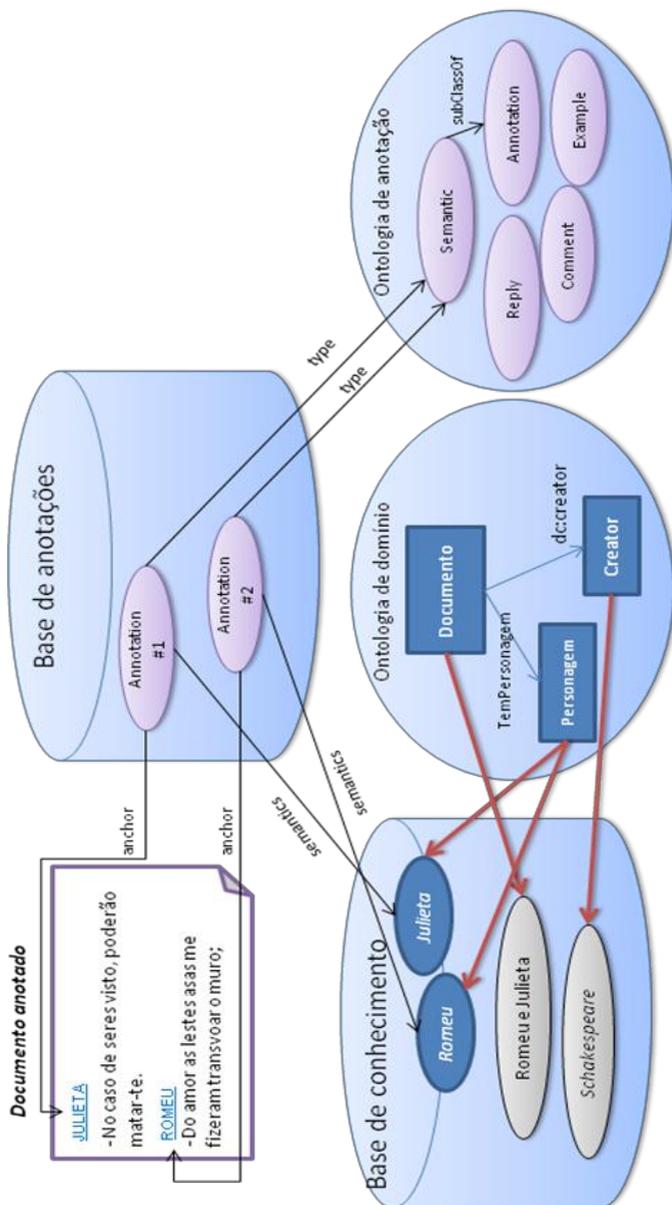


Figura 13 – Anotação semântica no sistema proposto

5.3 Avaliação das funcionalidades

O sistema DLNotes foi desenvolvido de maneira a atender aos requisitos identificados no capítulo anterior. A fim de atender o requisito (iii), o sistema DLNotes foi desenvolvido utilizando tecnologias Web, tais plataformas têm como principal característica o acesso através de navegadores. Estas plataformas têm como vantagens, entre outras: A flexibilidade e mobilidade, que permite que aplicativos fiquem disponíveis em qualquer lugar; Gerenciamento centralizado, permitindo que alterações sejam feitas em um único ponto e este seja replicado para todos os usuários.

Para os requisitos de não intrusão nos documentos da DL (vii), interoperabilidade das informações armazenadas pelo sistema (ix) e suporte para anotações livres e semânticas (iv) e (ii), todas as anotações e ontologias serão armazenadas em uma base de conhecimento, que é representada também, por uma ontologia. Como citado na seção Capítulo 2, o uso de ontologias traz vantagens como poder ser compartilhada por aplicações diferentes, e permitir o compartilhamento de conhecimento (GUIMARÃES, 2002). O uso de ontologias permite também, atingir o requisito (i), possibilitar que o sistema seja facilmente integrado a diferentes DL. Para isto, o sistema deverá permitir a importação de uma ontologia de domínio, que especifica os conceitos e relacionamentos tratados para um determinado domínio do conhecimento. Nesta proposta, a ontologia é representada utilizando RDFS (RDF *Schema*) e pode ser desenvolvida com editores de ontologias como Protégé ou OntoEdit.

Ainda para atender o requisito (ix), tornar o sistema interoperável e proporcionar que outros sistemas consigam utilizar a base de conhecimento mantida pelo sistema, o protocolo Annotea foi estendido. Este protocolo é utilizado por diversos sistemas de anotações Web, é um padrão mantido pela W3C para o enriquecimento de documentos Web baseado na colaboração e compartilhamento de anotações.

Para atender o requisito (x), será utilizado o padrão Dublin Core, este padrão tem como objetivo descrever objetos digitais, tais como textos, imagens, sons e vídeos na Web. É um padrão largamente utilizado pela comunidade de DLs (MARCHAND-MAILLET et al, 2006) para transferência de metadados entre diferentes DL, seu elemento de descrição mais simples, descreve itens como título, autor,

data de publicação, assunto, entre outros. Mais elementos podem ser incluídos (AALBERG, 2008).

O uso das tecnologias citadas acima permite o atendimento dos requisitos (v), (vi), (viii), (xi), (xii) e (xiii).

5.4 Considerações Finais

Este capítulo apresentou a proposta de um sistema de anotações para DL, chamado DLNotes. Foi apresentado a extensão do esquema de anotação Annotea, com o objetivo de dar suporte aos requisitos estabelecidos no capítulo anterior. Logo, foi exibida a arquitetura do sistema DLNotes e detalhado cada item que a compõe. Por fim, foi apresentada uma avaliação das funcionalidades disponíveis. O próximo capítulo exibe detalhes da implementação do DLNotes.

CAPÍTULO 6. PROTÓTIPO DESENVOLVIDO

Esta seção apresenta o protótipo do sistema de anotação, o protótipo é chamado de DLNotes (*Digital Libraries Notes*). Para realização de testes, o protótipo foi integrado à Biblioteca Digital de Literatura Brasileira (WILLRICH, 2006, NUPILL, 2009). A DL-LB tem como objetivo a disponibilização de uma grande coleção de obras literárias brasileiras e portuguesas de domínio público, informações sobre escritores, datas de publicação, editoras, gênero das obras, entre outras. A DL-LB atualmente possui 63645 obras e 16316 autores cadastrados, além de um total de 536 obras digitalizadas. Em sua implementação foi adotada a plataforma LAMP (Linux, Apache, MySQL, PHP) de código livre. O protótipo desenvolvido considerada somente obras armazenadas no formato HTML, que é o caso da DL-LB.

As seções que seguem, exibem a modelagem do protótipo através dos diagramas de modelagem UML, as tecnologias utilizadas no desenvolvimento do protótipo e exemplos da operacionalidade da aplicação.

UML (*Unified Modeling Language*) é uma linguagem de modelagem não proprietária, controlada pela OMG⁹ que contém uma família de notações gráficas com intuito de ajudar profissionais na descrição e no projeto de sistemas de *software* (FOWLER, 2007). O uso do UML possibilita uma melhor especificação, documentação e uma padronização das modelagens de *software* e será utilizado para a exibição do protótipo modelado.

6.1 Diagrama de casos de uso

Diagrama de casos de uso é um diagrama da UML, seu objetivo é “relacionar as funcionalidades do sistema modelado, através do elemento caso de uso, e os elementos externos que interagem com o sistema modelado, através do elemento sintático ator” (SILVA, 2007, p. 91). Atores e casos de uso são os principais elementos deste diagrama. Casos de uso são os “elementos responsáveis pela modelagem das funcionalidades em si” (SILVA, 2007, p. 101), são as operações de sistema. Atores, segundo XEXÉO (2007, p. 229), são entidades externas

⁹ OMG – *Object Management Group* – é um consórcio internacional de empresas que mantém padrões na área da orientação a objetos <http://www.omg.org>

ao sistema com comportamento próprio, eles interagem com o sistema para alcançar seus objetivos. Dentre os principais tipos de relacionamentos possíveis, pode-se citar a associação entre ator e caso de uso e a associação entre casos de uso. A primeira, associação entre ator e casos de uso, representa a participação da entidade externa (Ator) no caso de uso a ele associado.

A associação entre casos de uso pode aparecer em duas maneiras, inclusão ou extensão. Associação de inclusão, “estabelece que parte do comportamento inerente a um caso de uso está definida em outro caso de uso” (SILVA, 2007, p. 109), ou seja, toda vez que for executado o caso de uso, os casos de uso que são incluídos a ele, também serão executados. Associação por extensão, “estabelece uma relação em que um dos casos de uso tem seu comportamento estendido através do comportamento definido em outro caso de uso” (SILVA, 2007, p. 108), ou seja, o caso de uso que estende outro, explica melhor, mais detalhadamente o caso de uso estendido.

Este diagrama possui as seguintes notações (OESTEREICH, 2002): Atores podem ser representados por estereótipos textuais (normalmente representados por caixas com nome descrito no centro), estereótipos visuais (normalmente representados por uma figura de um boneco, relógio, cubos com seu nome abaixo ou acima da figura) ou uma mistura dos dois; Caso de uso é representado por uma elipse com o nome dentro. Podem adicionalmente conter um número para uma rápida identificação, notas textuais que o explicam mais detalhadamente. Relacionamentos são representados por linhas ou setas. Associação entre ator e caso de uso é representada por uma linha, ligando os dois elementos. Associação de extensão e inclusão são representadas por setas, que liga dois elementos e uma mensagem especificando o tipo de associação, “<<extend>>” (extensão) e, “<<include>>” (inclusão).

O sistema proposto na seção 5.2 define três atores: Usuário; Administrador Geral; Administrador Especialista. Usuário é o ator com objetivo de ler o documento da DL, criar, editar e visualizar as anotações do documento. Administradores são os atores que gerenciam o sistema de anotação na DL. Para representar as funcionalidades disponíveis para estes usuários, foi criado o diagrama de casos de uso, dividido em duas figuras no ANEXO 1, uma representando as interações do ator usuário e outra representando as interações dos atores administradores.

Na primeira figura do ANEXO 1, onde é apresentado os casos de uso referentes ao ator usuário, existem 5 funcionalidades principais disponíveis, que são representadas no diagrama, são elas: “Fazer *login*”, caso de uso onde o usuário entrará com seu nome de usuário e senha de acesso; “Manter Cadastro”, onde o usuário poderá criar, editar, visualizar e deletar seus dados e preferências; “Manter Anotações”, este caso de uso diz respeito a manutenção (criar/visualizar/editar/deletar) de anotações públicas e privadas, incluindo todos os tipos possíveis de anotações (*Semantic, Comment, Question*, etc); “Manter anotações resposta”, este caso de uso assim como o anterior, significa manter (criar/visualizar/editar/deletar) todas suas anotações respostas, incluindo todos os tipos possíveis de anotações respostas (*Comment, Agree*, etc); “Esconder/Mostrar anotações”, este caso de uso denota a exibição e ocultação das anotações no documento, a visualização de suas anotações privadas e anotações públicas e seus detalhes.

A segunda figura do ANEXO 1, apresenta os casos de uso referente a interação com os atores administradores, existem 8 funcionalidades principais disponíveis, descritas a seguir: “Fazer *login*”, caso de uso onde os administradores farão a validação de seu nome de usuário e senha para acesso ao sistema; “Importar ontologia”, caso de uso em que o administrador geral poderá importar a(s) ontologia(s) de domínio da DL; “Manter atividades de anotação”, Caso de uso em que o administrador geral poderá escolher diferentes tipos de anotações que serão possíveis em uma determinada atividade, este caso de uso é utilizado para o contexto de *e-learning*. Ainda neste caso de uso, os administradores poderão selecionar inclusive quais conceitos da ontologia de domínio farão parte da atividade; “Importar metadados da DL”, este caso de uso refere-se à importação dos metadados Dublin-Core da DL para a base de conhecimento; “Gerar base de conhecimento”, este caso de uso une a ontologia de domínio importada, os metadados importados e os contextos de anotações para a criação da base de conhecimento; “Gerenciar anotações”, este caso de uso é onde os administradores farão a verificação das anotações criadas, com o intuito de manter a consistência das anotações; “Gerenciar Usuários”, caso de uso onde o administrador geral pode especificar os tipos de usuários do DLNotes (especialista,normal), poderá bloquear e excluir usuários; “Manter cadastro”, caso de uso para os administradores manterem as informações cadastrais de seu perfil;

6.2 Diagrama de atividades

Diagrama de atividades segundo FOWLER (2005, p. 118) é “uma técnica para descrever lógica de procedimento, processo de negócio e fluxo de trabalho”, para (XEXÉO, 2007, p. 127), este diagrama foi proposto para “modelar aspectos dinâmicos de um sistema”, e têm como ênfase “detalhar o comportamento de um programa, isto é, descrevê-lo em execução”. (SILVA, 2007, p. 153), ou seja, “descrever a seqüência e a condição que esses comportamentos ocorrem” (XEXÉO, 2007, p. 127). Para ROQUES (2001, p. 58), este diagrama possui como uma de suas vantagens mostrar claramente o que está fora do sistema.

Este diagrama desempenha um papel semelhante aos fluxogramas, a principal diferença entre eles, destacadas por FOWLER (2005, p. 118), é que o diagrama de atividade suporta comportamento paralelo. O diagrama de atividades é composto principalmente por atividades e ações. Ação é a unidade de modelagem de comportamento, é atômica, não permitindo particionamento em termos de modelagem, já atividade é um elemento não atômico de modelagem e pode ser composta por ações ou atividades (SILVA, 2007, p. 157). No diagrama elas são representadas por um retângulo de cantos arredondados. O diagrama de atividades apresenta ainda, diversos fluxos de controle, descrito abaixo (SILVA, 2007, HAMILTON, 2006): Nodo inicial, representa o início da atividade, no diagrama ele é representado por um círculo totalmente preto; Nodo final, representado por um círculo parcialmente escuro, significa que a atividade chegou ao final de execução; Fluxo final, representado graficamente por um círculo com um “x” dentro, significa o final de um fluxo ou atividade, este nodo não significa necessariamente o fim da atividade, apenas de um fluxo dentro da atividade; Nodo de decisão, utilizados quando se quer executar diferentes seqüências de ações dependendo da condição, representado no diagrama por um losango; Nodo *Fork*, estabelece fluxos paralelos e; Nodo *Join*, estabelece a sincronização dos fluxos novamente. Estes dois últimos representados graficamente por uma barra. O diagrama de atividades pode apresentar o particionador de atividades, onde é possível especificar o responsável pela execução da ação e atividade modelada. Graficamente, cada partição é representada por um par de linhas paralelas contendo atividades e ações. A seqüência de realização das ações e atividades dentro de um diagrama é representada por setas.

O ANEXO 2 apresenta os diagramas de atividades para o sistema proposto. Cada caso de uso apresentado na seção anterior possui um diagrama de atividades que auxilia no detalhamento de cada caso de uso.

6.3 Diagrama de classes

O diagrama de classes, segundo SILVA (2007, P. 86), “é o modelo fundamental de uma especificação orientada a objetos”. É o diagrama que apresenta o conjunto de classes de um software, exibem seus atributos, métodos e os relacionamentos entre elas. Para FOWLER (2005), diagrama de classes “descreve os tipos de objetos presentes no sistema e os vários tipos de relacionamentos estáticos existentes entre eles”. Este diagrama possui a seguinte notação (OESTEREICH, 2002): As classes são representadas por retângulos, eles podem simplesmente ter o nome da classe ou podem ainda apresentar os atributos e operações da classe, respectivamente; Os atributos são listados pelo menos com seu nome, podem ainda apresentar especificações de tipo, valores iniciais e restrições; As operações ou métodos também podem ser simplesmente listadas pelos seus nomes e podem apresentar especificações como parâmetros, valores dos parâmetros, potencial valores e restrições; Os relacionamentos são elementos que ligam as classes umas nas outras, os relacionamentos possíveis no diagrama de classes da UML são herança, agregação e composição, mais informações sobre os relacionamentos possíveis em um diagrama de classes podem ser encontradas em SILVA (2007, p. 49). O ANEXO 3, apresenta o diagrama de classes da do servidor DLNotes. O diagrama de classe foi dividido em 4 imagens para proporcionar uma melhor visualização das classes e seus componentes.

A primeira imagem do ANEXO 3, apresenta as classes envolvidas no processo de manipulação de anotação, a classe principal chama-se “Anotação” e as outras classes apresentadas neste diagrama, são classes especializadas a partir da classe “Anotação”, criando o relacionamento denominado herança. As classes “*Example*” (Exemplo), “*Advice*” (Recomendação/Opinião), “*Question*” (Questionamento/Pergunta), “*SeeAlso*” (Veja mais), “*Change*” (Alteração), “*Comment*” (Comentário), e “*Explanation*” (Explicação) , são classes que foram definidas no padrão Annotea que segundo os requisitos do sistema proposto serão utilizadas no protótipo. A classe

“*Semantic*” estende o padrão Annotea a fim de dar suporte na manipulação de anotações semânticas.

A segunda imagem do ANEXO 3, apresenta as classes envolvidas no processo de manipulação de anotações resposta, estas classes suportam a comunicação e discussão entre os usuários da DL. A classe principal chama-se “*Reply*” (Resposta), e as classes que especializam esta classe chamam-se “*CommentReply*” (Comentário Resposta), “*AgreeReply*” (Concordar Resposta), “*DisagreeReply*” (Descordar Resposta), “*SeeAlsoReply*” (Veja Também Resposta), todas estipuladas também pelo padrão Annotea.

A terceira imagem do ANEXO 3, apresenta as classes envolvidas no processo de manipulação da ontologia, as classes que fazem parte desta manipulação chamam-se “Ontologia”, “*Namespace*”, “Classe”, “Propriedade”, e “Instancia”. Por fim as classes de conexão e configuração (quarta imagem do ANEXO 3). São elas: “ConexãoBanco”, classe para conexão com o banco de dados para controle de usuários; “ConexãoDBStore”, classe para conexão com o banco de dados para recuperação das bases de conhecimentos que são armazenadas de forma persistente em banco de dados; “*MemModel*” é a base de conhecimento que é recuperada pela classe “ConexãoDBStore”; “*KMPrivado*”, gerencia a manipulação das base de conhecimento privadas; “Usuário”, controla o acesso dos usuários, e gerenciamento de perfis e; “*TreadsController*” que é a classe que controla a seqüência das anotações e suas respostas. Neste diagrama, há uma ilustração que faz uma relação desta classe com a classe “Anotação” e “Reply” detalhadas nas primeiras imagens do ANEXO 3.

O protótipo trabalha com ontologias criadas em RDFS, para utilizar ontologias OWL, basta estender a classe “Propriedade” da terceira imagem do ANEXO 3 a fim de suportar os outros tipos de propriedades estipuladas na especificação da linguagem OWL.

6.4 Tecnologias Utilizadas

O servidor DLNotes, exibido na seção 5.2 foi desenvolvido utilizando a linguagem de programação PHP¹⁰ (versão 5.3.1). PHP é uma linguagem de programação interpretada, existente desde 1995,

¹⁰ PHP – *Hypertext preprocessor* – Mais informações podem ser encontradas em <http://php.net/>.

muito utilizada na geração de páginas dinâmicas na Web. O PHP pode ser usado para trabalhar com paradigmas de programação estruturados e orientado a objetos. Suas principais características são: robustez e velocidade; portabilidade; sintaxe similar a C/C++ e código-fonte aberto. Suporta diversas funcionalidades importantes como compatibilidade com diversas bases de dados, recebimento de arquivos, envio de *e-mails*, entre outros (PHP, 2009).

A linguagem de programação PHP é executada no servidor Web APACHE¹¹ (versão 2.0). O servidor APACHE é um software livre, o que significa que é possível estudá-lo, alterá-lo e utilizá-lo gratuitamente. É um software multiplataforma, podendo ser instalado e executado em diferentes sistemas operacionais. APACHE executa diversas linguagens de programação entre outras estão PHP, Pearl, Shell Script, Python, e ASP.

Para dar suporte a manipulação de ontologias e base de conhecimento foi utilizada a API RAP (versão 0.9.6). A API RAP¹² foi desenvolvida em PHP e tem o objetivo dar suporte a análise, consulta, serialização e manipulação de modelos RDF. A API RAP é mantida sob a licença LGPL¹³ e tem como principais características o uso de paradigmas centrado em afirmações, recursos, ontologias e cadeia; Possui linguagens de consulta RDQL e SPARQL e; Possui armazenamento em memória e armazenamento persistente em banco de dados, dando suporte a mais de trinta bancos de dados. A API RAP foi escolhida dentre uma gama de outras API's para manipulação de ontologias pela grande facilidade de manipulação e, código escrito na linguagem PHP, linguagem esta utilizada em diversas DLs.

RDF(S), como citado anteriormente é uma linguagem para representar informações sobre recursos na Web, é um padrão recomendado pela W3C, seu objetivo é criar um modelo simples de dados, com uma semântica formal. Os modelos RDF(S) e outras informações mantidas pelo protótipo (camada de dados) são

¹¹ APACHE HTTP *Server* – Mais informações podem ser encontradas em <http://www.apache.org/>.

¹² RAP – RDF API for PHP – mais informações podem ser encontradas em <http://sourceforge.net/projects/rdfapi-php/>.

¹³ LGPL - *Lesser General Public License* – é uma licença de software que permite de forma gratuita a cópia e distribuição dos softwares. A modificação dos softwares licenciados sob esta licença não é permitida. Mais informações podem ser encontradas em <http://www.gnu.org/copyleft/lesser.html>

armazenados de forma persistente em banco de dados. O sistema gerenciador de banco de dados (SGBD) utilizado é o MySQL¹⁴ (versão 5.4). Este SGBD é um dos mais utilizados no mundo (MYSQL, 2009), assim como o PHP, MySQL é um software multiplataforma que pode ser executado em diversos sistemas operacionais. Está sob a licença GPL¹⁵ e LGPL. Como principais características deste SGBD podem ser citadas o excelente desempenho e estabilidade; Pouco exigente em questão de *hardware*; Possui controle de transações e *triggers* e; interface gráfica (SUEHRING, 2002, p.43).

A camada de apresentação foi desenvolvida utilizando as linguagens de programação HTML, JavaScript, DOM e CSS.

HTML¹⁶ é uma linguagem de marcação utilizada para produção de páginas Web. Estas páginas são executadas nos navegadores Web. JavaScript¹⁷ é também uma linguagem de programação que é executada nos navegadores Web, como principais características desta linguagem pode-se citar a validação de formulários no lado do cliente, possibilitando uma verificação de possíveis erros entre a interação do usuário com o sistema; Interação com a página HTML através do uso de JavaScript e DOM¹⁸, dando um maior dinamismo pois com o uso destas tecnologias é possível alterar os estilos e os elementos de uma página HTML durante sua execução. O DOM permite que através da linguagem de programação JavaScript a estrutura das páginas Web possa ser alterada, como se fosse um conjunto de objetos. Permitindo com isso, que sejam feitas alterações instantâneas na página, por exemplo, em suas interfaces, redesenhando partes da página (CRANE et al., 2007, p. 22).

¹⁴ MySQL – Sistema gerenciador de banco de dados. Mais informações podem ser encontradas em <http://www.mysql.com/products/enterprise/server.html>.

¹⁵ GPL – *General Public License* – Os softwares regidos sob esta licença podem ser executados, modificados e redistribuídos. Este último possui o requisito de que a redistribuição deva ser liberada para toda comunidade a fim de que todos se beneficiem dele. Para isto, a liberação do código-fonte é imprescindível. Mais informações podem ser encontradas em <http://www.gnu.org/licenses/gpl.html>.

¹⁶ HTML - *HyperText Markup Language* (Linguagem de Marcação de Hipertexto). Mais informações podem ser encontradas em <http://www.w3.org/Markup/>.

¹⁷ JavaScript - Mais informações podem ser encontradas em <http://docs.sun.com/source/816-6408-10/>.

¹⁸ DOM - *Document Object Model* (Modelo de Objetos de Documentos). Mais informações podem ser encontradas em <http://www.w3.org/DOM/>.

CSS¹⁹ é uma linguagem de estilos criada e mantida pela W3C, usado para controlar o estilo de páginas Web. Os estilos são uma forma de controlar como o navegador processa os elementos HTML (LEMAY & COLBUM, 2006). O principal objetivo de seu uso é separar o conteúdo do formato do documento.

Para fazer a comunicação do servidor DLNotes com a interface foi utilizado AJAX. AJAX (*Asynchronous JavaScript and XML*) segundo NIEDERAUER (2007, p. 17) “é o uso sistemático de JavaScript e XML (entre outras tecnologias) para tornar o navegador mais interativo com o usuário, utilizando solicitações assíncronas de informações”. Com o uso de AJAX o protótipo pode interagir com o usuário sem a necessidade de recarregar toda página Web a cada interação. O uso de AJAX no protótipo se deu através do uso de uma API chamada SAJAX²⁰. Esta API foi desenvolvida com o intuito de facilitar o desenvolvimento de aplicações que utilizam AJAX, ela pode ser utilizada por sistemas feitos por diversas linguagens de programação, ela faz o gerenciamento da comunicação da camada de apresentação e a camada lógica.

A arquitetura das respectivas tecnologias utilizadas no desenvolvimento do protótipo é apresentada na figura abaixo em forma do diagrama de componentes de UML.

¹⁹ CSS - *Cascading Style Sheets* – Mais informações podem ser encontradas em <http://www.w3.org/Style/CSS/>.

²⁰ SAJAX - *Simple* AJAX toolkit – Mais informações podem ser encontradas em <http://www.modernmethod.com/sajax/>.

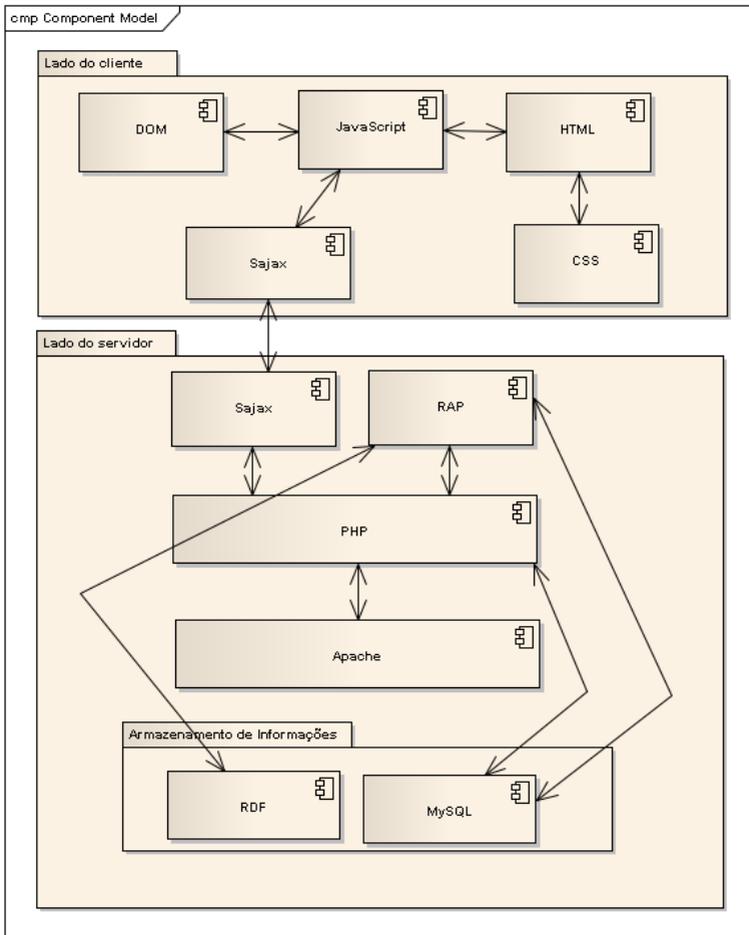


Figura 14 – Arquitetura das tecnologias utilizadas no protótipo

A figura acima além de exibir as tecnologias utilizadas no protótipo, mostra também a interação entre as tecnologias, através das linhas que fazem ligação entre elas. É possível notar que a comunicação entre o lado do servidor com o lado do cliente através do uso da API SAJAX. Ela que recebe/envia as requisições do lado do cliente para o lado do servidor, vice-versa. O armazenamento das informações pode se dar de duas formas. A primeira forma é através do acesso direto da linguagem de programação PHP ao SGBD MySQL, isto é feito no momento em que é necessária a manipulação de coleções de usuários

pelo módulo gerenciador de usuários e módulo de administração. Outra forma é o acesso direto da API RAP. Esta API faz acesso direto ao SGBD a fim de salvar e recuperar as ontologias e base de conhecimento e acesso aos modelos RDF(S) armazenados.

6.5 Ontologias utilizadas

Nesta seção será apresentada as ontologias que são usadas no protótipo.

6.5.1 Ontologia de domínio

Uma ontologia de domínio são utilizadas como forma de representação de conhecimento de um domínio específico, ou algumas partes deste (GUARINO, 1998). Segundo FREITAS (2003, p. 16), “procuram solucionar um problema específico de um domínio, como identificar doenças do coração, a partir de uma ontologia de domínio de cardiologia”.

O ANEXO 4 exhibe a ontologia do domínio de literatura brasileira, ontologia esta aplicada no contexto da DL-LB. A figura apresenta somente as classes e suas hierarquias, o relacionamento entre as classes, propriedades e instâncias não são apresentados por motivos de simplificação. A figura foi criada através do editor de ontologias *Protégé Ontology Editor and Knowledge Acquisition System*²¹.

A ontologia da DL-LB apresenta classes onde é possível criar instâncias semânticas de classes gerais como Localização (país, estado, cidade, bairro e até logradouro), Personagens (pessoa, animal ou objeto). Além das citadas, a ontologia apresenta classes restritas ao domínio de literatura brasileira como, tipos de figuras de linguagem, tipos de narradores, e identificação de foco narrativo. Outras classes que compõe a ontologia são as classes Obra, Autor e Evento, classe esta, que possibilita relacionar determinados eventos entre personagens em determinados lugares, por fim, a classe Anacronia, que é possível identificar a não coincidência da ordem dos acontecimentos com a ordem contada na história. Esta ontologia apresenta diversos tipos de relacionamentos, como por exemplo, relacionamentos de parentesco,

²¹ *Protégé Ontology Editor and Knowledge Acquisition System* – Mais informações em: <http://protege.stanford.edu/>

onde é possível indicar que um personagem é irmão, tio, pai de outro etc., ou que um país possui (tem) estados.

Para uma melhor compreensão, a Figura 15 exibe uma pequena parte da ontologia de domínio e sua relação com as instâncias que serão criadas através das anotações semânticas. Esta figura é dividida em duas partes, a parte superior, que representa o nível intencional e a parte inferior que representa o nível extensional. O protótipo utiliza os dois níveis para exibir e criar anotações semânticas para os usuários. Pode-se notar que a figura apresenta instâncias dos conceitos, a partir destas instâncias foram criados relacionamentos respeitando a estrutura conceitual.

Por exemplo, a *string* “Capitu” pode ser armazenada na base de conhecimento como instância do conceito “Mulher”, uma subclasse do conceito “Pessoa”, que por sua vez é subclasse de “Personagem”. O usuário pode ainda incluir outra instância do conceito “Pessoa”, chamada “Bento Santiago” (apelido Bentinho) e relacioná-la com “Capitu” através da propriedade “É casado”. Posteriormente, outro usuário pode associar ambas as personagens à instância do conceito “Cidade” através da propriedade “Mora em”.

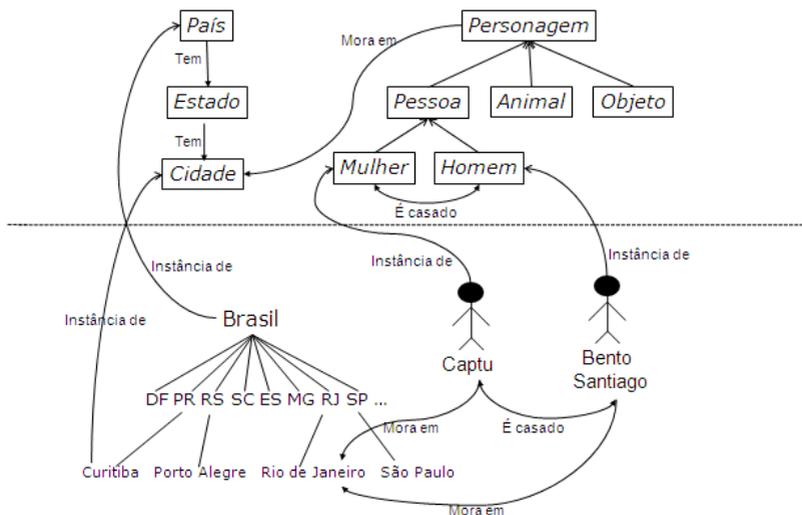


Figura 15 – Parte da ontologia da DL-LB

6.5.2 Ontologia de anotação

O sistema utiliza uma ontologia de anotações que segue o modelo proposto na seção 5.1.

A figura abaixo exhibe o esquema de anotação utilizado. Somente as classes e suas hierarquias são apresentadas nesta figura, o relacionamento entre as classes, propriedades e instâncias foram omitidos da figura a fim de proporcionar uma melhor visualização das classes utilizadas.

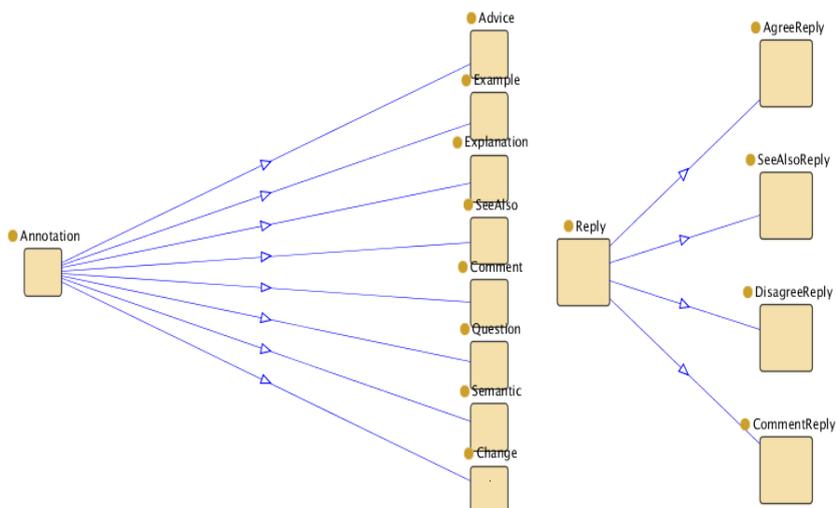


Figura 16 – Classes da ontologia usada no esquema de anotação do protótipo

6.6 Administração do protótipo

A administração do sistema contempla as funcionalidades dispostas no diagrama de casos de uso cujo diagrama é apresentado na figura 2 do ANEXO 1 e descrito na seção 6.1. Dentre os casos de uso já citados, o caso de uso “Gerenciar anotações” não foi implementado até o momento, todos os outros casos de uso estão concluídos.

6.7 Manipulação de Anotações

As funcionalidades da manipulação de anotações disponíveis neste protótipo foram apresentadas no diagrama de casos de uso

apresentado na primeira figura do ANEXO 1 e descrito na seção 6.1. Esta seção detalha alguns casos de uso e exibe sua execução.

6.7.1 Exibição do documento e inserção de anotações

A exibição do documento ao usuário se dá no momento que a DL faz uma chamada ao DLNotes informando a URL do documento que deve ser carregado. Após isto, o usuário deve se autenticar no DLNotes informando seu nome de usuário e senha. Caso os dados estiverem corretos, o protótipo verifica o perfil do usuário e verifica as preferências de visualização das anotações que podem ser as seguintes: Sem visualização das anotações; Visualização somente de suas anotações ou; Visualização de suas anotações e as anotações públicas de outros usuários.

Se o perfil indica a inclusão de anotação no documento, o sistema busca na base de conhecimento as anotações que satisfazem os requisitos e insere dinamicamente os links no documento. A figura abaixo exibe uma obra com *links* das anotações.

The screenshot shows a web browser window displaying a document titled "LITERATURA BRASILEIRA Textos literários em meio eletrônico". The document is an edition of "O Contrato, de Machado de Assis". The text is displayed with several annotations, each marked with a small icon. A sidebar on the right side of the browser window shows the DLNotes interface, which includes a search bar and a list of annotation types: Ajuda, Citar, Anotação Semântica, Citar, Anotação Livre, Exibir Todas, Anotações, Encerrar Todas, and Anotações. The document text includes a paragraph about a consórcio and a list of questions and answers related to a contract.

LITERATURA BRASILEIRA
Textos literários em meio eletrônico
O Contrato, de Machado de Assis

Edição de Referência:
http://www.uol.com.br/machadoceassis

Quem quiser celebrar um consórcio, examine primeiro as condições, depois as forças próprias, e, finalmente, faça um [contrato de consórcio](#). Foi o que não cumpriram estas duas meninas de colégio, cuja história vou contar em três folhas de almaço. [Eram filhas de funcionários públicos](#), e não se conheciam antes. Conheceram-se ali, simpatizaram uma com a outra, e travaram uma dessas [relações](#), que resistem aos anos, e são muita vez a melhor recordação do passado. [Uma delas](#) tinha mais um ano que [a outra](#), era a diferença. No mais as mesmas. Igual estatura, igual indole, iguais olhos e igual nascimento. Eram filhas de funcionários públicos, ambas dispoendo de um certo legado, que lhes deixara o [pai](#). Para que a semelhança seja completa, o [pai](#) era o mesmo, um certo [funcionário público](#), capitalista.

Com tal ajuste de condições e circunstâncias, não precisavam mais nada para serem amigas. [O colégio](#) [ligava-as](#) [dentro](#) [dos](#) [muros](#) [do](#) [colégio](#). No fim de poucos meses de freqüência, eram as mais unidas criaturas de todo ele, a ponto de causar inveja às outras, e até desconfiança, porque como cochichavam muita vez sozinhas, as outras imaginavam que diziam mal das [outras](#). Naturalmente, as relações continuaram cá fora, durante o colégio, e as famílias vieram a ligar-se, graças às meninas. Não digo nada das famílias, porque não é o principal do escrito, e eu prometi escrever isto em três folhas de almaço; basta saber que tinham ainda pai e mãe. Um dia, no colégio, contavam elas onze e doze anos, lembrou-se [uma delas](#) de propor à outra, adivinhem o quê? vamos ver se são capazes de adivinhar o que foi. Falavam: casamento de uma prima de [uma delas](#), e que há de lembrar a outra?

— Vamos fazer um contrato?
— Que é?
— Mas diga se você quer...
— Mas se eu não sei o que é?
— Vamos fazer um contrato: — casar no mesmo dia, na mesma igreja...
— Valeu! nem você casa primeiro nem eu; mas há de ser no mesmo dia.

Figura 17 – Obra com anotações

Como pode ser visto na Figura 17, cada tipo de anotação possui um ícone que é inserido após a *string* da anotação. Este ícone tem o intuito de diferenciar os diversos tipos de anotações. A figura abaixo exibe todos os tipos de ícones das anotações utilizados no sistema.

Imagem	Title	Título
	Comment	Comentário
	Change	Mudança
	Recommendation	Recomendação
	Example	Exemplo
	Agree	Concordar
	Disagree	Discordar
	Question	Questão
	Explantation	Explanação
	Semantic	Semântica
	See Also	Veja também

Figura 18 – Ícones para diferenciar os tipos de anotações

A figura 17 apresenta um menu de opções no canto inferior direito que possibilita ao usuário acessar algumas funcionalidades do DLNotes como ajuda, criar anotações semânticas, criar anotações livres, exibir ou esconder as anotações da obra. Este menu de opções fica sempre visível na tela e o usuário pode deslocá-lo para todos os cantos da tela (superior-esquerdo; inferior-esquerdo; superior-direito; inferior-direito). A figura 19 exibe o menu de opções do DLNotes.



Figura 19 – Menu de opções do DLNotes

6.7.2 Visualização de anotações

Quando o usuário desejar ver detalhes de uma anotação, basta dar um clique do botão esquerdo do *mouse* sobre o *link* da anotação, com isto, é aberta uma nova interface onde exibe dados referentes à anotação. A interface de exibição de uma anotação não-estruturada é mostrada na figura abaixo.



Figura 20 – Interface de exibição de anotação livre

A interface apresentada na Figura 20 mostra em sua parte superior dados como: Título; Criador da anotação; Data e hora em que a anotação foi criada e; Conteúdo da anotação. Na parte inferior da interface existem dois botões, um para fechar a interface de visualização dos detalhes da anotação e outro que possibilita o usuário criar uma anotação resposta. Anotação resposta é uma anotação referente a anotação principal. Uma anotação pode possuir diversas anotações respostas, caso uma anotação seja deletada, a primeira anotação resposta passará a se considerada como principal. A figura abaixo exibe a interface de exibição de anotação com uma anotação principal e uma anotação resposta.



Figura 21 – Interface de exibição de anotação livre com uma seqüência de uma anotação e uma anotação resposta

Quando o usuário desejar ver detalhes de uma anotação semântica, assim como as anotações anteriores, basta clicar sobre o *link* da anotação no documento. Será exibida uma interface onde são apresentados alguns dados da anotação como data de criação, data de atualização e lista de autores. Para este tipo de anotação, não é possível a criação de uma seqüência de anotações. As anotações semânticas que forem do tipo pública, vários autores podem editar. A figura abaixo exhibe a interface de exibição de anotação semântica.

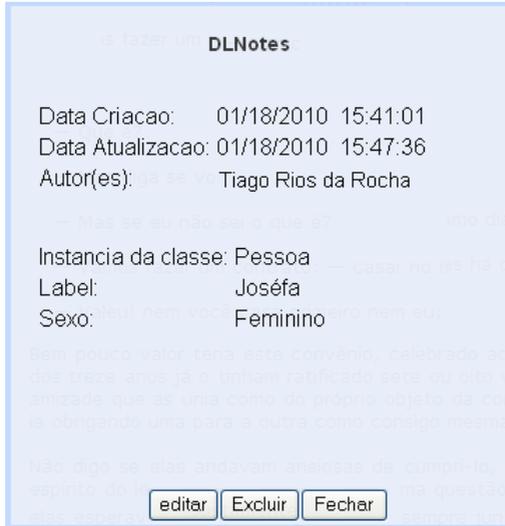
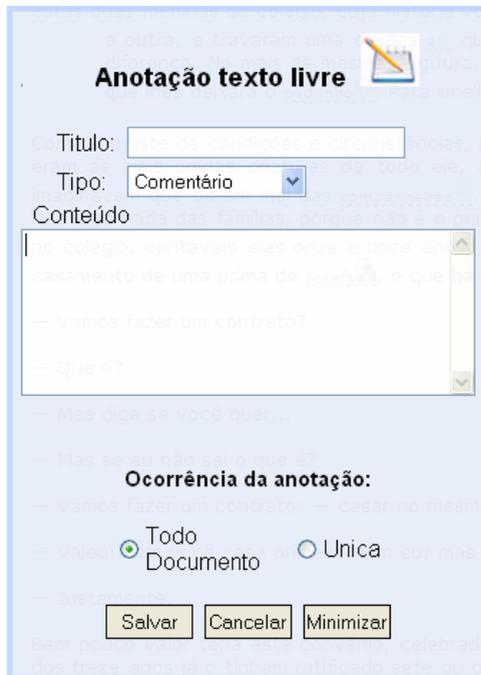


Figura 22 – Interface de exibição de anotação semântica

6.7.3 Criação de anotação

Para criar uma anotação (não-estruturada ou semântica), o usuário primeiramente deve selecionar a parte do texto (*string*) o qual deseja criar uma anotação. Após isto, há duas formas de criar a anotação, são as seguintes: O primeiro meio é selecionar a opção desejada no menu posicionado nos cantos da tela (“criar uma anotação livre” ou “criar uma anotação semântica”); A outra opção é o usuário clicar com o botão direito do mouse sobre o documento, é então aberto uma tela onde será possível o usuário selecionar que tipo de anotação deseja fazer. No caso de anotação não-estruturada, a seguinte interface é exibida na Figura 23.



Anotação texto livre

Título:

Tipo:

Conteúdo

No colégio, contavam elas onze e doze an...
casamento de uma prima de Josefa...
— Vamos fazer um contrato?
— Que é?
— Mas diga se você quer...
— Mas se...
— Vamos...
— Vale...
— Just...

Ocorrência da anotação:

Todo Documento Unica

Figura 23 – Interface de criação de anotação livre

Nesta interface (Fig. 23), o usuário poderá descrever um título para a anotação, escolher o tipo de anotação que ele deseja criar (comentário, questionamento, etc.) e descrever seu conteúdo. Por fim, poderá ainda definir se a ocorrência da anotação no documento ocorrerá em todo documento ou de forma única. A primeira significa que a anotação ocorrerá em todas as vezes que o texto selecionado aparecer no documento, por exemplo, toda vez que a *string* “Josefa” aparecer no texto, o protótipo fará uma relação com a anotação que está sendo criada; Ocorrência única, significa que a anotação que está sendo criada pertence somente aquela localização da obra, por exemplo, se o usuário definir que a anotação referente a *string* “Josefa” é somente válida em uma ocorrência da obra, todas as outras ocorrências desta *string* não serão relacionadas com a anotação. A interface de criação de anotação resposta é semelhante à interface exibida na figura anterior.

6.7.3.1 Criação de anotação semântica

Conforme detalhado anteriormente, criar uma anotação semântica é o ato de ligar instâncias encontradas no texto do documento, a seu referido conceito contido em uma ontologia. Para criar uma anotação semântica, primeiramente o usuário deve selecionar a *string* referente à anotação e acessar a opção “criar anotação semântica”, como explanado no item anterior. Após isto, o sistema buscará possíveis instâncias existentes na base de conhecimento, caso exista, é exibida para que o usuário possa escolher se a instância que ele deseja criar é de fato a encontrada pelo DLNotes. A interface que representa esta etapa é apresentada na Figura 24.

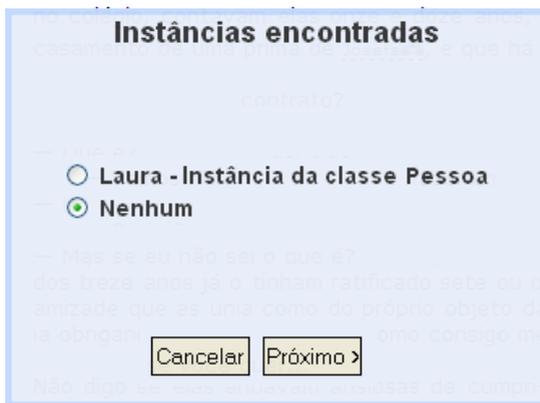


Figura 24 – Interface de seleção de possíveis instâncias da base de conhecimento

Caso o usuário identifique que a anotação que ele deseja criar é uma das recuperadas pela base de conhecimento ele pode acessá-la e alterá-la. Se o DLNotes não recuperou nenhuma instância ou as recuperadas não é de desejo do usuário, este pode criar uma nova instância onde a interface (Fig. 25) possibilita que o usuário selecione dentro de todos os conceitos da ontologia de domínio, qual conceito ele deseja instanciar a *string* selecionada.

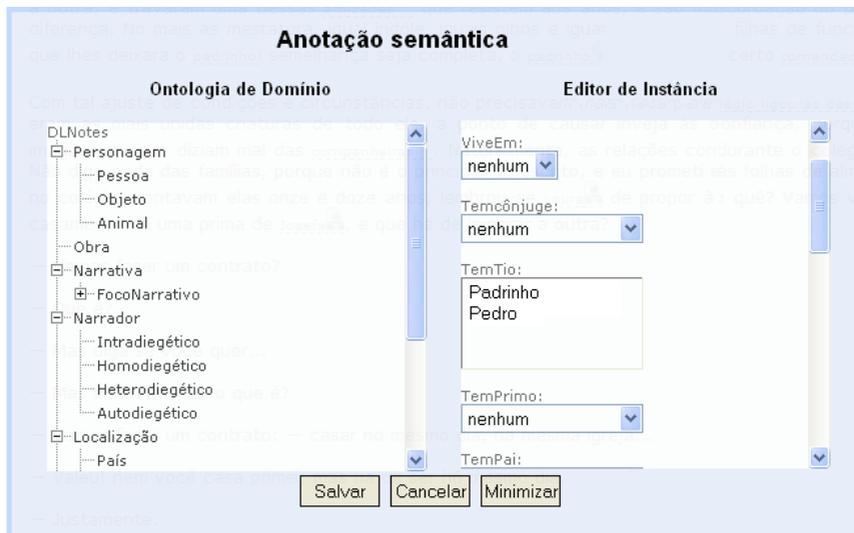


Figura 25 – Interface de criação de anotação semântica

Os termos exibidos no lado esquerdo da Figura 25, apresentados em forma de árvore, são os termos referentes à ontologia de domínio, estes termos são apresentados conforme a atividade de anotação que o usuário está anotando. Por exemplo, se o usuário está em uma atividade em um contexto em que somente 3 termos da ontologia lhe são disponibilizados, somente estes 3 termos serão exibidos.

O lado direito da figura é um formulário onde o usuário poderá preencher as propriedades do termo selecionado e fazer as relações com as instâncias da base de conhecimento. Por exemplo, em uma ontologia de parentesco é possível ao identificar uma instância do conceito “Pessoa”, relacionar outra instância de “Pessoa” da base de conhecimento através da propriedade “é pai”.

6.8 Possibilidades de uso do DLNotes no suporte ao ensino de literatura

Um dos objetivos do DLNotes é suportar atividades de *e-learning* através da colaboração e criação do conhecimento compartilhado entre os usuários da DL. Através de seu uso, usuários compartilham seus conhecimentos e trocam experiências a fim de construir colaborativamente um novo conhecimento comum. Esta sessão demonstra exemplos de atividades que podem ser desenvolvidas

através do uso do DLNotes em um domínio de literatura, domínio tratado na DL-LB.

Análise literária é o processo de compreensão de uma obra, pode ser feita através do uso do DLNotes com a DL-LB. A análise pode ser feita em vários estágios, e cada estágio pode ser representado através de uma atividade de anotação.

O primeiro exemplo do uso do DLNotes para análise literária é para a identificação de personagens. Nesta etapa, o estudante pode identificar os personagens com anotações livres (p.e., comentários) ou pela criação de anotações semânticas onde, por exemplo, pode criar uma instância da classe “Personagem” ou suas subclasses (p.e., “Pessoa”, “Animal”, “Objeto”). A Figura 26 apresenta um exemplo da identificação de personagens para o livro “Dom Casmurro”.

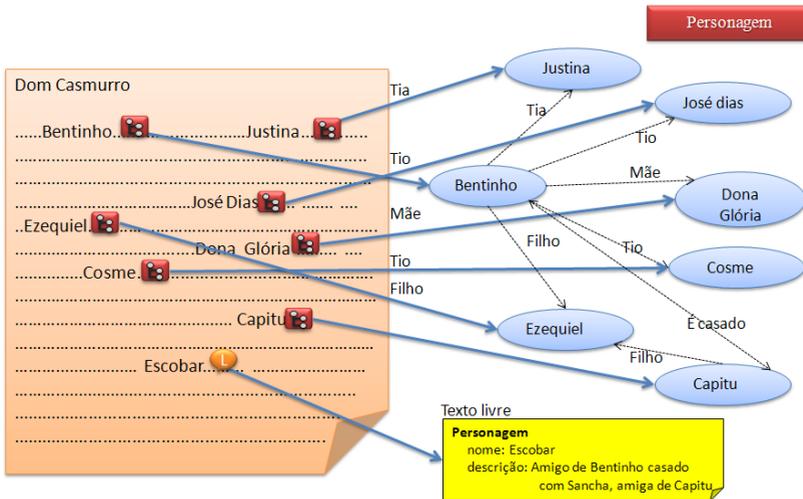


Figura 26 – Exemplo de atividade de identificação de personagens

Outro estágio da análise literária pode ser a identificação de figuras de linguagem. Neste estágio o estudante pode identificar as figuras de linguagem usadas no conteúdo e associar a respectiva passagem no texto com a correspondente subclasse de “figura de linguagem” na ontologia (p.e., antítese, eufemismo, pleonasmo). A figura 27 exibe um exemplo deste uso.

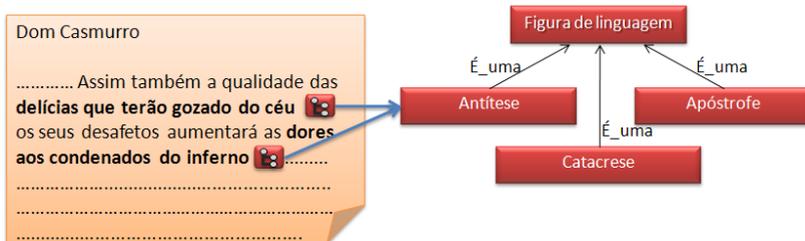


Figura 27 – Exemplo de atividade de identificação de figuras de linguagens

Ainda sobre a obra Dom Casmurro, a Figura 28 exibe outra possível etapa em uma análise literária, onde é possível descrever a seqüência de eventos e ações. Nesta etapa, as ações e eventos podem ser instanciados pelas classes "Evento" e "Ações", respectivamente. Logo, estas instâncias podem ser relacionadas com instancias de personagens a fim de identificar quem fez uma ação, por fim, ainda é possível relacionar outras instâncias de eventos e ações para indicar casualidades ou conflitos.

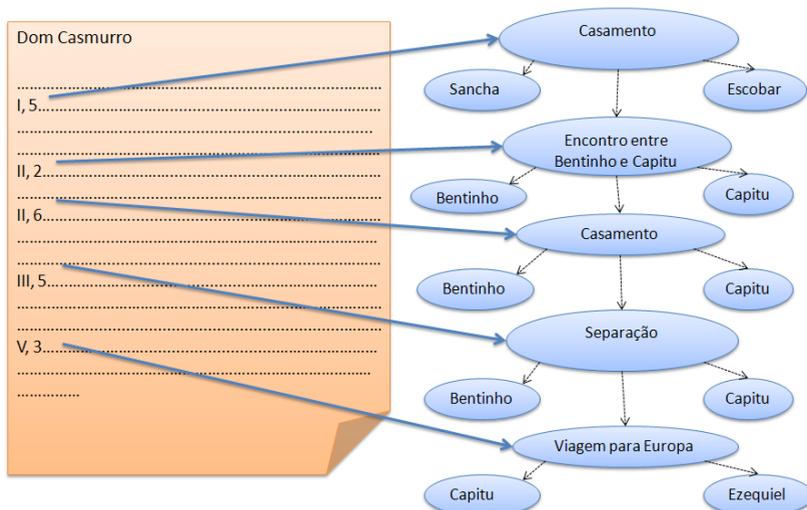


Figura 28 – Exemplo de atividade de identificação de ações e eventos

6.9 Considerações Finais

Este capítulo tratou do protótipo desenvolvido, detalhou sobre a modelagem do DLNotes apresentando os diagramas de casos de uso, diagramas de atividades e diagramas de classes de UML. Foi detalhada as tecnologias utilizadas no desenvolvimento do protótipo, e a comunicação entre as diversas tecnologias. Foi demonstrado detalhes da ontologia de domínio utilizada no contexto da DL-LB e a ontologia de anotações. Foram exibidas algumas interfaces do sistema e detalhado algumas funcionalidades. Por fim, foi demonstrado algumas possibilidades de uso do DLNotes para o ensino de literatura.

O próximo capítulo exibe as considerações finais desta dissertação e a proposta de trabalhos futuros.

CAPÍTULO 7. CONCLUSÕES

As Bibliotecas digitais (DLs) surgem em um cenário de expansão de informações como uma ferramenta capaz de suprir uma grande demanda por conhecimento, pois ela é uma coleção organizada de informações digitais dotadas de serviços integrados, devidamente organizados e descritos de modo a que essa informação se torne acessível mediante métodos de pesquisa.

DLs são divididas em duas categorias, a primeira diz respeito à infra-estrutura de informações, que define acesso a informação, o fluxo e o gerenciamento através de serviços de busca, catalogação, indexação e projetos de sistemas que organizam o conhecimento. A segunda categoria, diz respeito à infra-estrutura de informação, dando ênfase a infra-estruturas sociais em DLs, em termos de serviços para a criação e compartilhamento de conhecimento em espaços de conhecimento comuns e comunicação interativa entre os usuários (AGOSTI et al, 2005).

Anotações tem sido freqüentemente proposta como uma técnica para que os usuários adicionem conteúdos, e assim, compartilhem idéias dentro de sistemas de informação (NICHOLS et al., 2000). Anotações podem ser interpretadas em muitas maneiras: como um link, como uma construção de caminho, como comentário, como uma marca em torno de um texto, como uma descentralização de autoridade, como um registro de leitura e interpretação, ou como uma memória para comunidade (MARSHALL, 1998).

O processo de criar anotações em documentos digitais traz várias vantagens para os usuários de uma DL. Esta tarefa possibilita que os usuários identifiquem passagens importantes no texto, criem diferentes interpretações e façam relações do conteúdo do texto com conceitos do mundo real formalmente descritos em uma ontologia. Ontologias são definidas como "uma especificação explícita de uma conceitualização (GRUBER, 1993, p.2) e oferece uma forma de representação do conhecimento através de formalismos.

Com o conhecimento organizado através de uma ontologia é possível proporcionar o compartilhamento de conhecimento e uma troca de experiências entre os usuários proporcionando a construção do conhecimento através da discussão, reflexão e tomada de decisões de maneira colaborativa através de atividades de aprendizagem.

Este trabalho tem o intuito de promover uma mudança na interação dos usuários com as DLs e contribuir para o desenvolvimento de serviços de infra-estrutura social através da criação, e disseminação do conhecimento e, ao mesmo tempo, proporcionar uma possível evolução na infra-estrutura de informações com respeito a serviços de busca, catalogação e indexação.

Este trabalho mostra como recursos para criação e manipulação de anotações não-estruturadas e semânticas de maneira colaborativa podem ser incorporados em uma DL. Estes recursos permitem a análise colaborativa do conteúdo de documentos digitais e a organização da informação resultante em bases de conhecimento organizada segundo uma ontologia de domínio. Tais bases de conhecimento evolui com o uso de recursos de anotações e podem ser utilizadas para descrever partes do texto e suas relações semânticas. Isto promove a integração e o reuso dos resultados de reflexões de diversos usuários e contribui para a compreensão, organização e recuperação precisa de conteúdo.

Este trabalho apresentou as formas de representação de anotações propostas pelo W3C, as conceitualizações de anotações livres e anotações semânticas e um estudo das ferramentas encontradas atualmente na literatura. Baseado neste estudo, foram levantados os requisitos necessários para um sistema de anotações em DLs, logo, foi feito um comparativo dos sistemas de anotações para BDs encontrados na literatura com os requisitos definidos. Ainda, foi proposto um sistema para o suporte e manipulação de anotações em DL, chamado DLNotes.

DLNotes é um sistema com objetivo de proporcionar aos usuários, identificarem passagens importantes no texto, desde simples marcações até anotações referentes a uma conceitualização explícita e formal, proporcionar a colaboração entre as atividades dos usuários e o compartilhamento de conhecimento através da reflexão, discussão e tomada de decisão em grupo.

As anotações no DLNotes seguem uma extensão do esquema de anotação definido pelo Annotea. As anotações são armazenadas em uma base de conhecimento onde permite com que elas sejam processadas tanto manualmente por usuários, quanto automaticamente por *softwares*, a fim de revelar novos relacionamentos possíveis entre porções de dados, ou no suporte a outros tipos de serviços, como por exemplo, serviço de busca semântica.

O DLNotes emprega ontologias do domínio específico da DL. Com base na ontologia de domínio, este sistema permite gerar anotações colaborativas segundo uma conceitualização padronizada. Assim, o sistema permite a criação de uma base de conhecimento acerca do domínio específico tratado pela DL e pode ser facilmente incorporado em diferentes DLs.

7.1 Trabalhos Futuros

Como sugestões para trabalhos futuros destacam-se:

Validação do sistema com usuários a fim de avaliar a importância do sistema no processo de ensino/aprendizagem. Este trabalho tem como objetivo identificar a evolução da aprendizagem de alunos através do uso do sistema de anotação em um processo colaborativo. Identificar os principais objetivos dos usuários ao anotarem uma obra, possíveis dificuldades e sugestões de melhoria.

Para proporcionar a instrutores manterem e avaliarem atividades de ensino/aprendizagem, propõe-se a integração do DLNotes com um sistema de ensino (por exemplo, Moodle). A partir desta integração, instrutores poderão criar/manter/avaliar atividades de anotações diretamente do sistema de ensino.

Diversas DLs armazenam documentos no formato PDF, atualmente o sistema DLNotes trabalha com documentos no formato HTML. Uma proposta é a expansão do protótipo a fim de suportar anotações em documentos do tipo PDF ou um processo de conversão de documentos PDF para HTML.

A criação compartilhada de conhecimento e manutenção de diversas bases de conhecimento entre os usuários podem causar inconsistências no conhecimento adquirido, causado por informações ambíguas e informações incompletas. A fim de corrigir possíveis problemas propõe-se o desenvolvimento de um módulo para o tratamento de conhecimento inconsistente.

A busca semântica tem como objetivo facilitar e permitir uma maior relevância e precisão nos resultados apresentados, através da análise do significado das palavras nas pesquisas no seu mecanismo de busca. Outra proposta é a criação de serviços de buscas semânticas para os conteúdos da DL baseado no conhecimento adquirido.

Para a descoberta de novos relacionamentos sobre os conhecimentos armazenados, propõe-se o desenvolvimento de um módulo para criação de regras que permitam inferências sobre a base de conhecimento;

Por fim, a utilização de técnicas de mapeamento de ontologias para a integração de bases de conhecimento de diferentes DLs visando acesso unificado.

7.2 Publicações

ROCHA, T. R. ; WILLRICH, R. ; FILETO, R. ; TAZI, S. . Supporting Collaborative Learning Activities with a Digital Library and Annotations. In: WCCE - 9th IFIP World Conference on Computers in Education, 2009, Bento Gonçalves. WCCE Conference Book. Boston : Springer, 2009. p. 349-358.

ROCHA, T. R. ; WILLRICH, R. ; FILETO, R. ; SANTOS, A. L.; GIRARDELLO, R. . Anotações Semânticas em Bibliotecas Digitais Voltadas ao Ensino. In: SBIE - Simpósio Brasileiro de Informática na Educação – Resumo expandido, 2008, Fortaleza. XIX SBIE, 2008 - Fortaleza, CE, 2008.

CAPÍTULO 8. REFERÊNCIAS BIBLIOGRÁFICAS

AALBERG, T.; *Supporting relationships in digital libraries*. 2003. Norwegian University of Science and technology. Disponível em: <http://www.idi.ntnu.no/grupper/su/publ/phd/aalberg-thesis.pdf>. Acesso em: janeiro de 2010.

AGOSTI, M.; FERRO, N.; FROMMHOLZ, I.; THIEL, U.; *Annotations in Digital Libraries and Collaboratories - Facets, Models and usage*. ECDL: 2004. p. 244-255. 2004.

AGOSTI, M.; ALBRECHTSEN, H.; FERRO, N.; FROMMHOLZ, I.; HANSEN, P.; ORIO, N.; PANIZZI, E.; MARK PEJTERSEN, A. AND THIEL, U. *DiLAS: a Digital Library Annotation Service. In Proceedings of the Workshop on Annotation for Collaboration: Methods Tools and Practices*. Paris. 2005.

AGOSTI, M.; FERRO, N.; PANIZZI, E.; TRINCHESE, R.; *Annotation as a support to user interaction for content enhancement in digital libraries. Proceedings of the working conference on Advanced visual interfaces*, pp. 151-154. 2006.

ALENCAR, A. F.; *Bibliotecas Digitais: uma Nova Aproximação*. Em: Informação & Sociedade: Estudos, João Pessoa, v. 14, n. 1, p. 1-13, 2004. Disponível em <http://www.informacaoesociedade.ufpb.br/ojs2/index.php/ies/issue/view/10>. Acesso em: janeiro de 2010.

AMÂNCIO, M. J.; *Avaliação das bibliotecas digitais Perseus e Virgínia Tech: análise comparativa de sistemas de pesquisa e recuperação de informação*. Congresso Nacional De Bibliotecários, Arquivistas E Documentalistas, 9, 2007.

ARKO, R.A.; GINGER, K.M.; KASTENS, K.A.; WEATHERLEY, J.; *Using annotations to add value to a digital library for education*, D-Lib Magazine, Vol. 12 No.5, 2005.

ATKINS, D. et al.; *Report of the Santa Fe Planning Workshop on distributed Knowledge Work Environments: Digital Libraries*. 1997.

BORGMAN, C.; *What are digital libraries competing vision. Information Processing & Management*. Vol. 35 (3). pp. 227-243. ISSN: 0306-4573. 1999.

BRUSH, A.J.B. Annotating Digital Documents for Asynchronous Collaboration. Dept. of Computer Science and Engineering, Univ. of Washington. Technical Report. 2002.

CRANE, D.; PASCARELLO, E.; JAMES, D. Ajax em ação. Tradução Edson Furmankiewicz e Carlos Schafranski. São Paulo: Pearson Prentice Hall, 2007.

DAVIES, J., STUDER, R., AND WARREN, P.; Semantic Web Technologies: Trends and Research in Ontology-Based Systems. John Wiley & Sons. ISBN: 978-0-470-02596-3. 2006.

FOWLER, M.; UML Essencial: Um breve guia para a linguagem-padrão de modelagem de objetos / Fowler, M.; trad. João Tortello. 3ªed. Editora Bookman. Porto Alegre. 2005.

FREITAS, F. ; Ontologias e a Web Semântica. In: Renata Vieira; Fernando Osório. (Org.). Anais do XXIII Congresso da Sociedade Brasileira de Computação. Volume 8: Jornada de Mini-Cursos em Inteligência Artificial. Campinas: SBC, 2003, v. 8, p. 1-52.

GONÇALVES, P. R.; Utilização a API JENA 2 Ontology na especificação de um domínio. Trabalho de conclusão de curso. Universidade Luterana do Brasil. Palmas. 2004.

GRUBER, T.R.; Toward principles for the design of ontologies used for knowledge sharing. Em: Knowledge Systems Laboratory, Technical Report, Computer Science Department, Stanford University, California , Vol. KSL 92-04 (1993) , p. 907-928.

GUARINO, N.; Formal ontology and information systems. Em: Proc. Intl. Conf. on Formal Ontologies in Information Systems (FOIS), IOS Press. pp. 3–15. 1998.

GUARINO, N.; Semantic Matching: Formal Ontological Distinctions for Information Organization, Extraction, and Integration. Em: international Summer School on information Extraction: A Multidisciplinary Approach To An Emerging information Technology M. T. Pazienza, Ed. Lecture Notes. Em: Computer Science, vol. 1299. Springer-Verlag, London, 139-170. 1997.

GUIMARÃES, F. J. Z.; Utilização de ontologias no domínio B2C. Programa de Pós-graduação em Informática. Rio de Janeiro. 2002.

HAMILTON, K.; MILES, R.; *Learning UML 2.0*. ISBN: 0-596-00982-8. Ed. O'Reilly. 2006.

HAUBEN, J.; *Vannevar Bush and JCR Licklider: Libraries of the future 1945-1965*. Em: *Conference in the Institute for Library Science of Humboldt University in Berlin*. 2004. Disponível em: <http://www.ais.org/~jrj/acn/acn15-2.articles/jhauben.pdf>. Acesso em: janeiro de 2010.

HECK, R. M.; LUEBKE, S. M.; OBERMARK, C. H.; *A Survey of Web Annotation Systems*. disponível em: http://www.math.grin.edu/~rebelsky/Blazers/Annotations/Summer1999/Papers/survey_paper.html. Acesso em: janeiro de 2010.

HILTZ, S. R.; *Collaborative Learning in Asynchronous Learning networks: Building Learning Communities*. Em: *WebNet 98 World Conference of the WWW, Internet, and Intranet Proceedings*. Orlando, FL, 1998.

JAYAPRAKASH, A. VENKATRAMANA, R. *Role of Digital Libraries in E-learning. DRTC Conference on ICT for Digital Learning Environment*. 2006.

KHELIF, K.; DIENG-KUNTZ, R.; BARBRY, P.; *An Ontology-based Approach to Support Text Mining and Information Retrieval in the Biological Domain. Special Issue on Ontologies and their Applications of the Journal of Universal Computer Science (JUCS)*, p. 1881 - 1907. Vol. 13. Issue 12. 2007.

KIRYAKOV, A.; POPOV, B.; TERZIEV, I.; MANOV, D.; OGNJANOFF, D.; *Semantic annotation, indexing, and retrieval. Web Semantics: Science, Services and Agents on the World Wide Web*, 2(1), pp. 49-79. 2004.

KOIVUNEN, M.-R., SWICK, R.R.; *Collaboration through annotations in the semantic web*. Em: *Handschub S., Staab, S. (Eds.), Annotation for the Semantic Web. Frontiers in Artificial Intelligence and Applications*, vol. 96. IOS Press, Amsterdam, The Netherlands, pp. 46-62. 2003.

LASSILA, O.; SWICK, R.; *Resource Description Framework (RDF) model and syntax specification*.1.0. 1999. Disponível em: <http://www.w3.org/TR/1999/REC-rdf-syntax-19990222/>. Acesso em: janeiro de 2010.

LEMAY, L.; COLBUM, R.; *Sams Teach Yourself: Web Publishing With HTML and CSS in One Hour a Day*. Ed. Sams Publishing. ISBN: 0-672-32886-0. 2006.

LESK, M.; *Practical digital libraries: Books, bytes, and bucks*. São Francisco: Morgan Kaufman. 1997.

LESK, M.; *Understanding Digital Libraries*. 2ªed. Editora Morgan Kaufmann. ISBN: 1-55860-924-5. 2005.

LYNCH, C.; *Where Do We Go From Here? : The Next Decade for digital libraries*. *D-Lib Magazine*. Vol. 11, num. 7/8. ISSN: 1082-9873. 2005.

MARSHALL, C.; *Toward an ecology of hypertext annotation*. Em: HYPERTEXT '98: *Proceedings of the ninth ACM conference on Hypertext and hypermedia : links, objects, time and space---structure in hypermedia systems* (1998), pp. 40-49. 1998.

MARCHAND-MAILLET, S.; BRUNO, E.; MOËNNE-LOCCOZ, N.; *Structured Multimedia Description for Simplified Interaction and Enhanced Retrieval*. *ERCIM News 66. Special Theme:European Digital Library*. 2006.

MEIRINHOS, M.; OSÓRIO, A.; *Colaboração e comunidades de aprendizagem*. *International Symposium on Computers in Education*. 8. León. p. 270-277. 2006.

MIZOGUCHI, R.; VANWELKENHUYSEN, J.; IKEDA, M.; *Task Ontology for Reuse of Problem Solving Knowledge*. Em: *2nd International Conference on Very Large-Scale Knowledge Bases: Building and Knowledge Sharing*. 1995.

MYSQL. *Why MySQL?*. Disponível em: <http://www.mysql.com/why-mysql/>. Acesso em: janeiro de 2010.

NICHOLS, D. M.; PEMBERTON, D.; DALHOUMI, S.; LAROUK, O.; BELISLE, C.; TWIDALE, M. B.; *DEBORA: Developing an Interface to Support Collaboration in a Digital Library*. *Proceedings of the Fourth European Conference on Research and Advanced Technology for Digital Libraries* (ECDL2000), Lisbon, Portugal. Springer-Verlag. 239-248. 2000.

NIEDERAUER, J.; *Web Interativa com AJAX e PHP*. ISBN: 978-85-7522-126-6. Novatec Editora. São Paulo. 2007.

NUPILL. Biblioteca Digital de Literatura Brasileira. Disponível em: <http://www.literaturabrasileira.ufsc.br/>. Acesso em: janeiro de 2010.

OESTEREICH, B.; *Developing software with UML: Object-oriented analysis and design in practice*. 2ª ed. ed. Editora *Pearson Education* Ltd. ISBN 0-201-75603-X. 2002.

PENEDO, S.; *Uma técnica de recuperação adaptativa de obras em bibliotecas digitais baseada no perfil do usuário*. Dissertação de mestrado. Universidade Federal de Santa Catarina. Programa de Pós-graduação em Ciência da Computação. Florianópolis. 2005.

PHP. Características da linguagem de programação PHP. Disponível em: http://php.net/manual/pt_BR/features.php. Acesso em: janeiro de 2010.

PROCÓPIO, E.; *Construindo uma biblioteca digital*. Editora Edições inteligentes. ISBN: 85-7615-112-X. São Paulo, 2004.

ROCHA, T. R. ; WILLRICH, R. ; FILETO, R. ; TAZI, S. . *Supporting Collaborative Learning Activities with a Digital Library and Annotations*. In: WCCE - 9th IFIP World Conference on Computers in Education, 2009, Bento Gonçalves. WCCE Conference Book. Boston : Springer, 2009. p. 349-358.

ROQUES, P.; *UML in practice: The art of modeling software systems demonstrated through worked examples and solutions*. ISBN: 0-470-84831-6. Ed. Wiley. 2001.

SARACEVIC, T.; DALBELLO, M.; *A survey of digital library education*. Em: *Proceedings of the American Society for Information Science and Technology*. Vol. 38, p. 209-223. 2001.

SILVA, R. P.; UML: *Modelagem Orientada a Objetos*. Editora Books. ISBN: 85-7502-205-9. Florianópolis. 2007.

STUDER, R.; BENJAMINS, V.; FENSEL, D.; *Knowledge engineering: principles and methods*. *Studer 98 knowledge engineering*. 1998.

SUEHRING, S.; *MySQL Bible*. ISBN: 0-7645-4932-4. Ed. Wiley Publishing, Inc. 2002.

TEDD, L. A.; LARGE, A.; *Digital Libraries: principles and Practices in a Global Environment*. Editora K. G. SAUR. ISBN: 3-598-11627-6. 2005.

TENIER, S.; TOUSSAINT, Y.; NAPOLI, A.; POLANCO, X.; *Instantiation of Relations for Semantic Annotation*. IEEE/WIC/ACM International Conference on Web Intelligence - WI 2006. P. 463-472. 2006.

UREN, V.; CIMIANO, P.; IRIA, J.; HANDSCHUH, S.; VARGASVERA, M.; MOTTA, E. and CIRAVEGNA, F.; *Semantic annotation for knowledge management: Requirements and a survey of the state of the art*. Web Semantics: Science, Services and Agents on the World Wide Web, vol. 4, no. 1, pp. 14-28. 2006.

USCHOLD, M. et al.; *Ontologies: Principles, Methods and Applications*. *Knowledge Engineering Review*. Vo. 11. p. 93-136. 1996.

VAN HEIJST, G.; SCHREIBER, A. T.; WIELINGA, B. J. *Using explicit ontologies in KBS development*. *International Journal of Human-Computer Studies*, v. 46, n. 2-3, p. 183-192. 1997.

WALLER, R.; *Functionality in digital annotation: imitating and supporting real-world annotation*. Ariadne, Ed. 35, vol. 30 Abril de 2003. Disponível em: www.ariadne.ac.uk/issue35/waller/intro.htm. Acesso em: janeiro de 2010.

W3C-ANNOtea; *Annotea Protocols*. 2002. Disponível em: <http://www.w3.org/2002/12/AnnoteaProtocol-20021219>. Acesso em: janeiro de 2010.

W3C-OWL; *OWL Web Ontology Language Guide*. 2004. Disponível em: <http://www.w3.org/TR/2004/REC-owl-guide-20040210/>. Acesso em: janeiro de 2010.

W3C-RDF; *Resource Description Framework (RDF): Concepts and Abstract Syntax*. 2004. Disponível em: <http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/>. Acesso em: janeiro de 2010.

W3C-RDFS; *RDF Vocabulary Description language 1.0: RDF Schema*. 2004. Disponível em: <http://www.w3.org/TR/rdf-schema/>. Acesso em: janeiro de 2010.

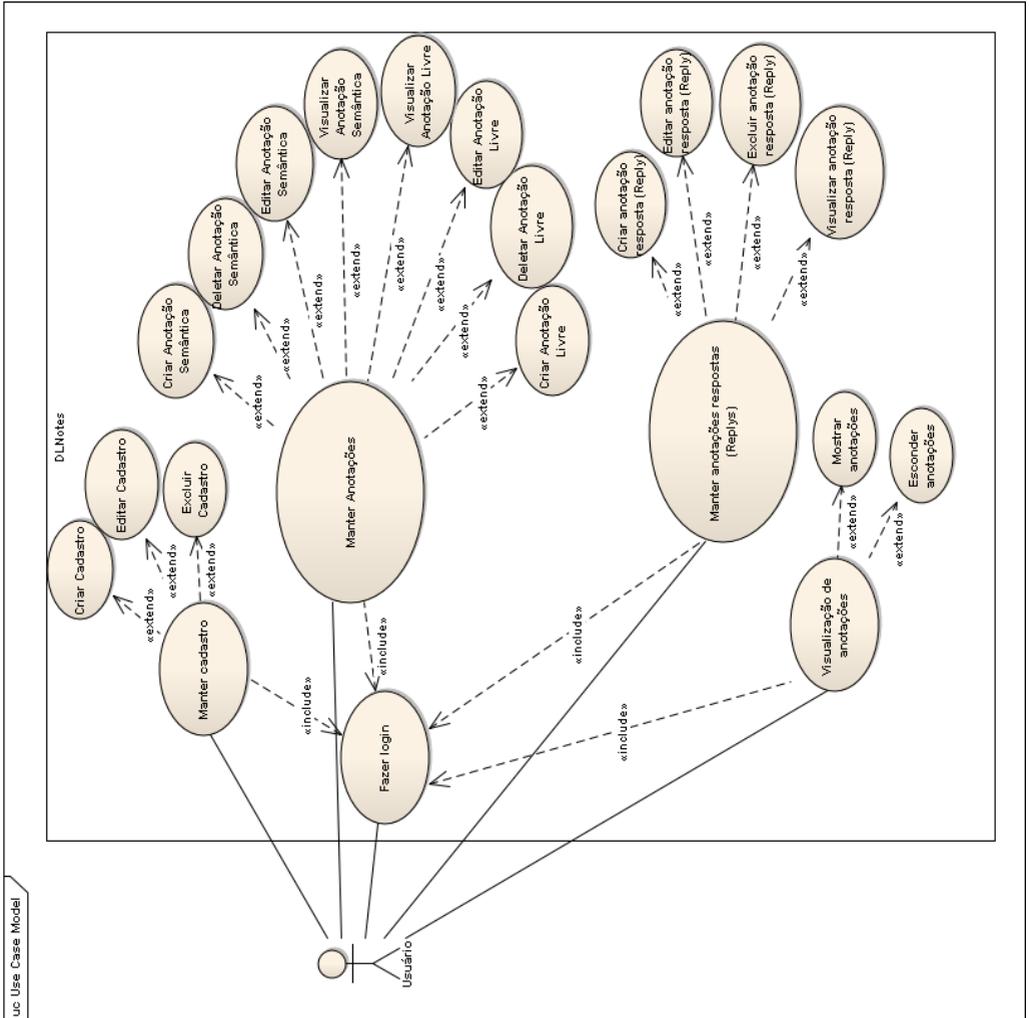
WILLRICH, R. et al.; *Sistema de Recuperação de Informações Adaptativo Aplicado a Biblioteca Digitais*. XII Simpósio Brasileiro de Sistemas Multimídia e Web, p. 165-173, 2006.

XEXÉO, G.; Modelagem de Sistemas de Informação: Da análise de requisitos ao modelo de interface. 2007. Disponível em http://wiki.xexeo.org/tiki-download_file.php?fileId=188. Acesso em: janeiro de 2010.

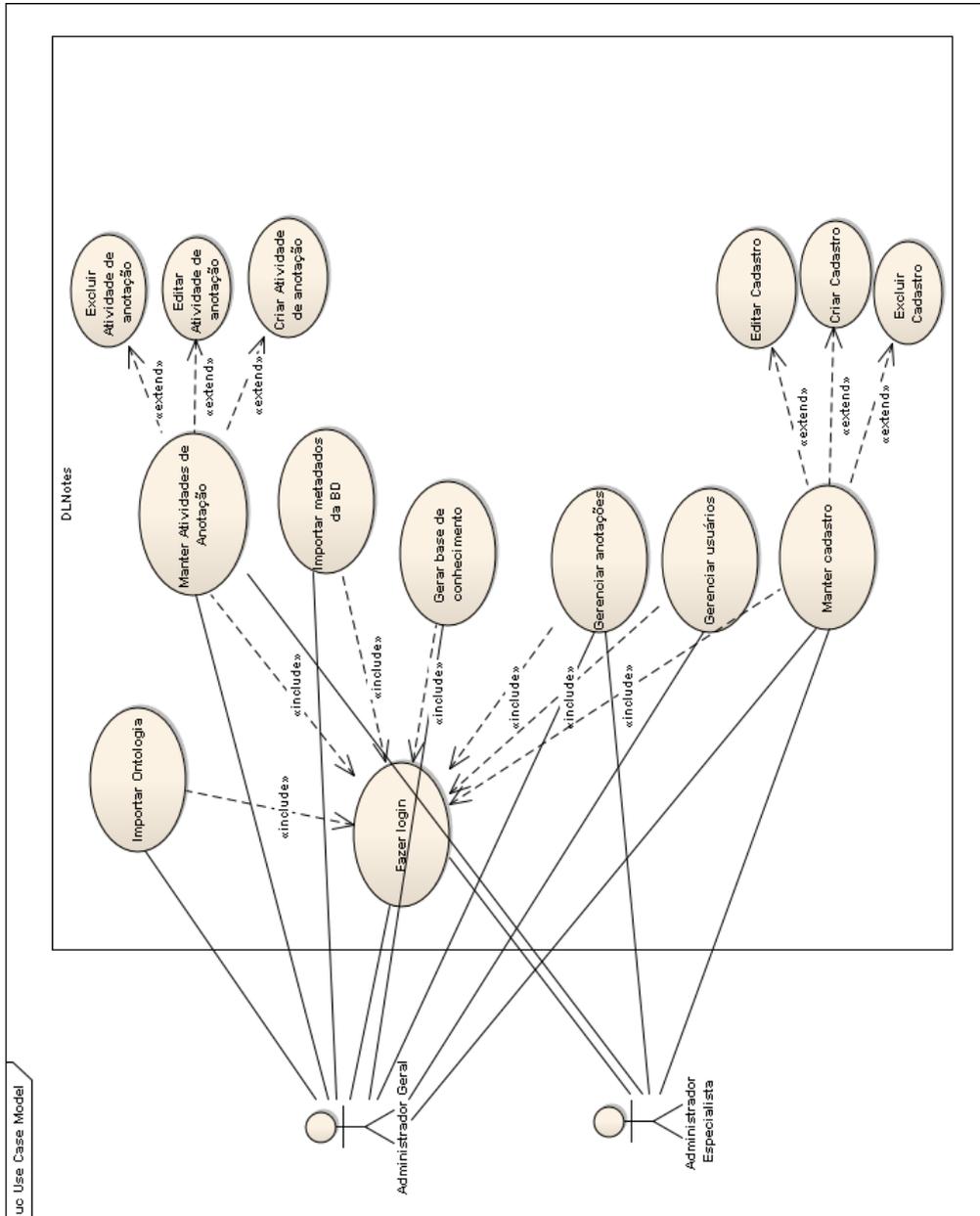
ANEXOS

ANEXO 1 – DIAGRAMA DE CASOS DE USO

1.1 Ator usuário.

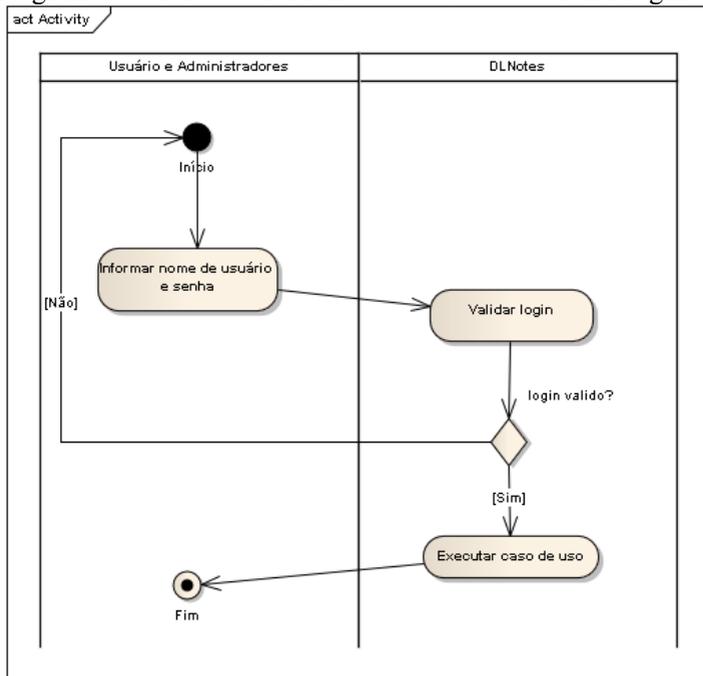


1.2 Atores Administradores

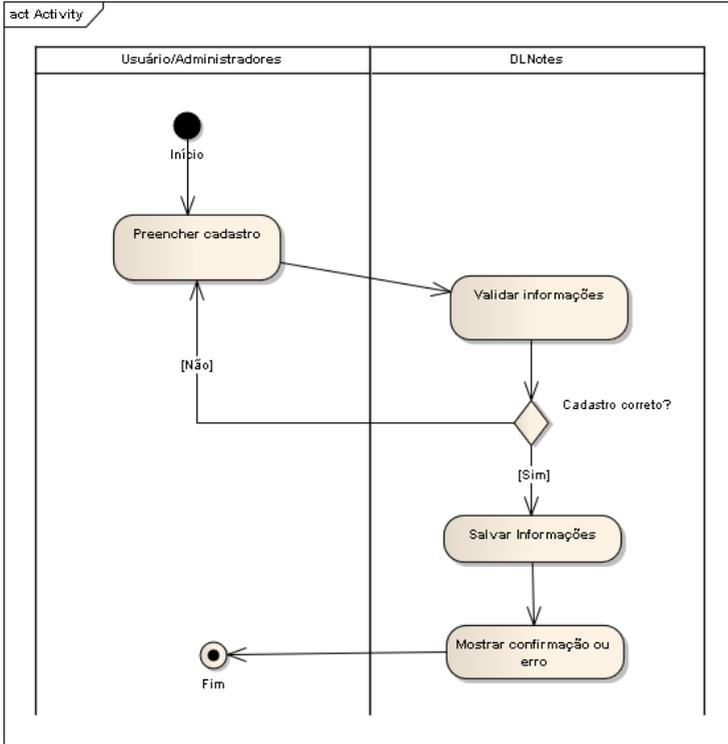


ANEXO 2 – DIAGRAMAS DE ATIVIDADES

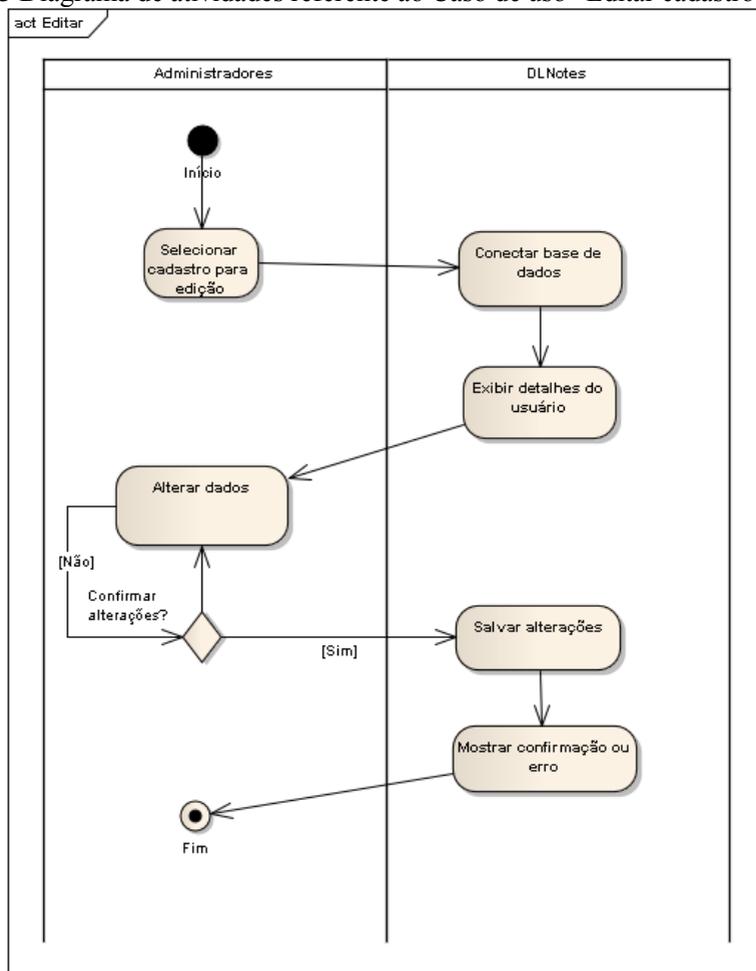
2.1 Diagrama de atividades referente ao Caso de uso “Fazer login”:



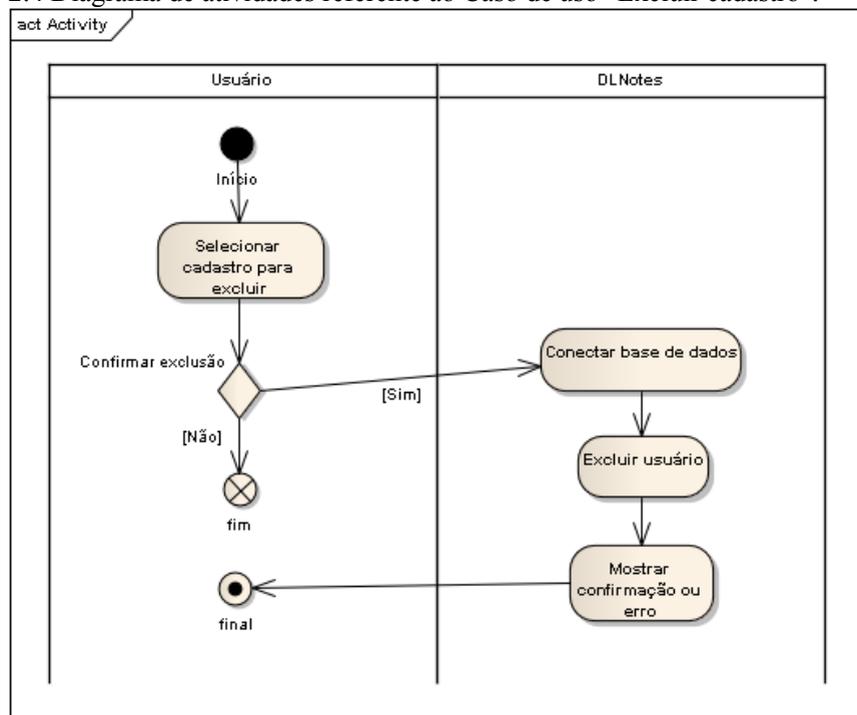
2.2 Diagrama de atividades referente ao Caso de uso “Criar cadastro”:



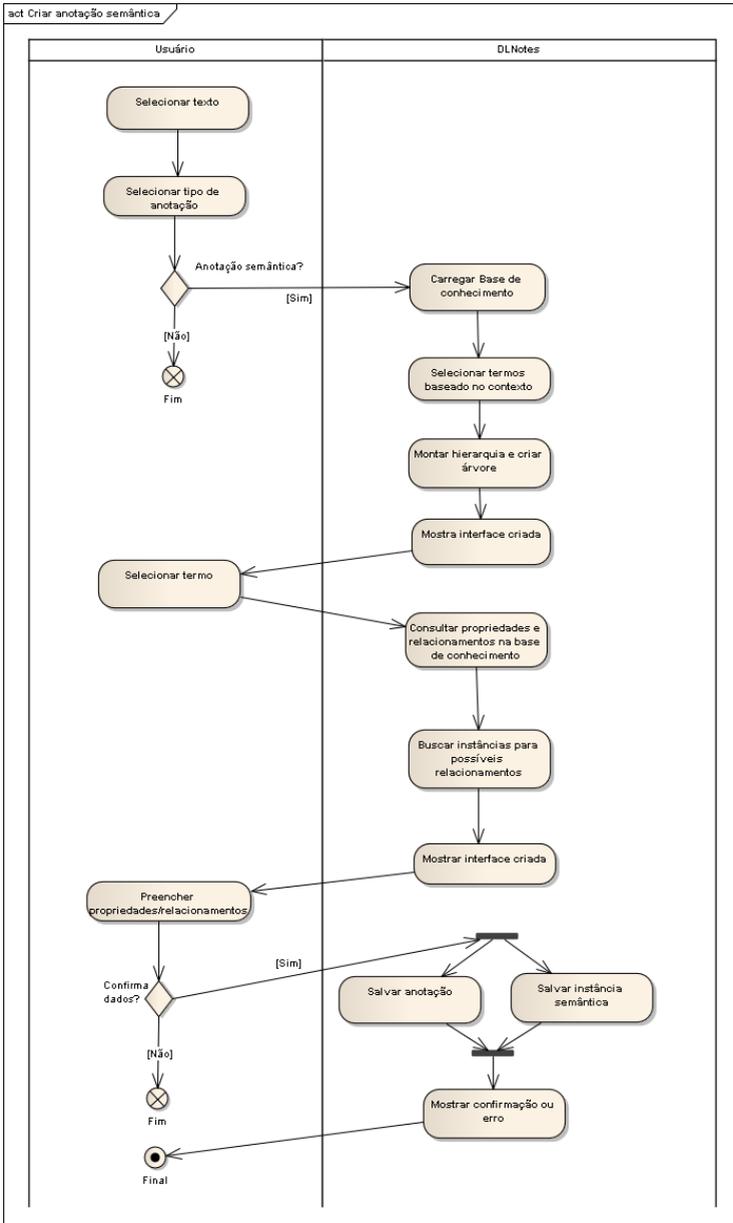
2.3 Diagrama de atividades referente ao Caso de uso “Editar cadastro”:



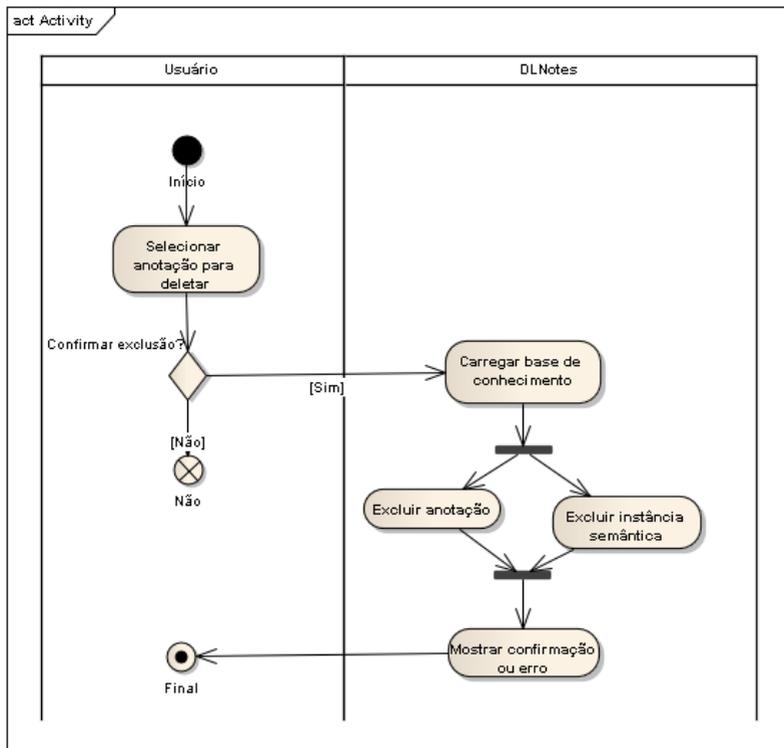
2.4 Diagrama de atividades referente ao Caso de uso “Excluir cadastro”:



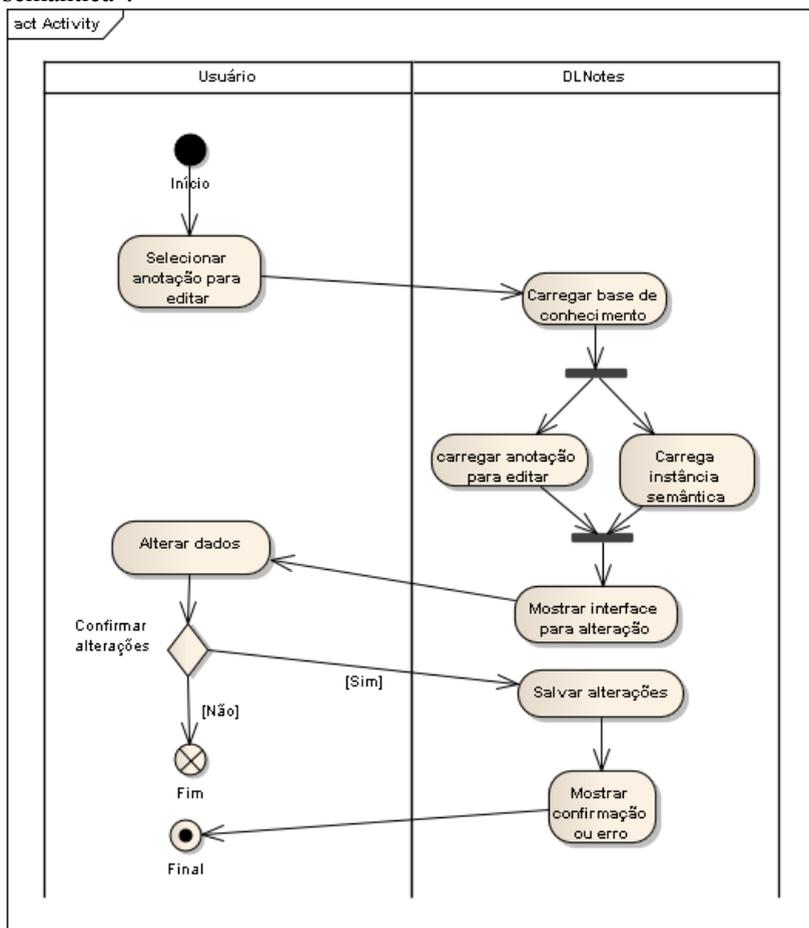
2.5 Diagrama de atividades referente ao Caso de uso “Criar anotação semântica”:



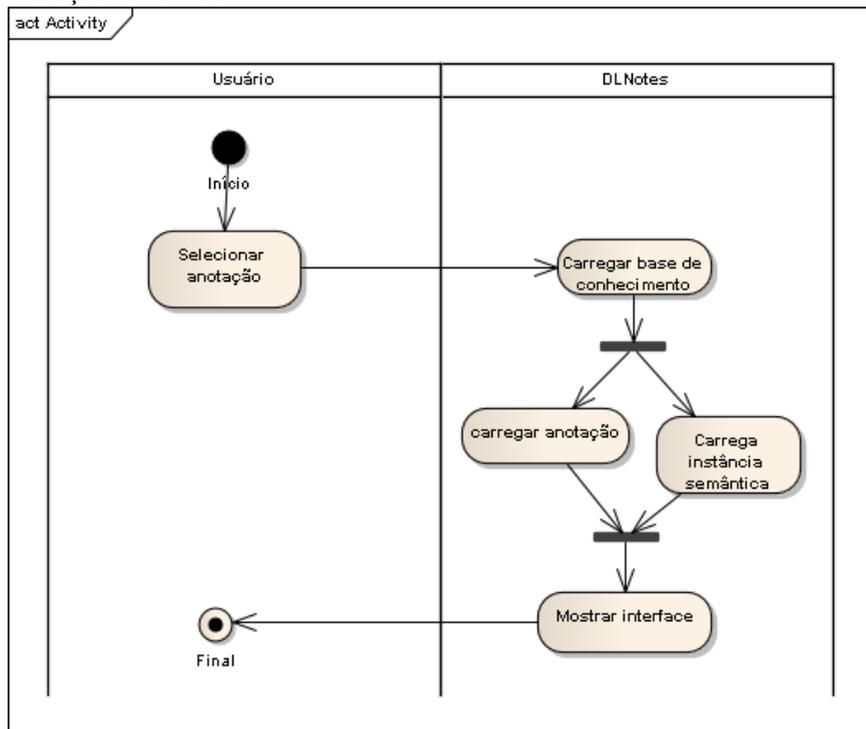
2.6 Diagrama de atividades referente ao Caso de uso “Deletar anotação semântica”:



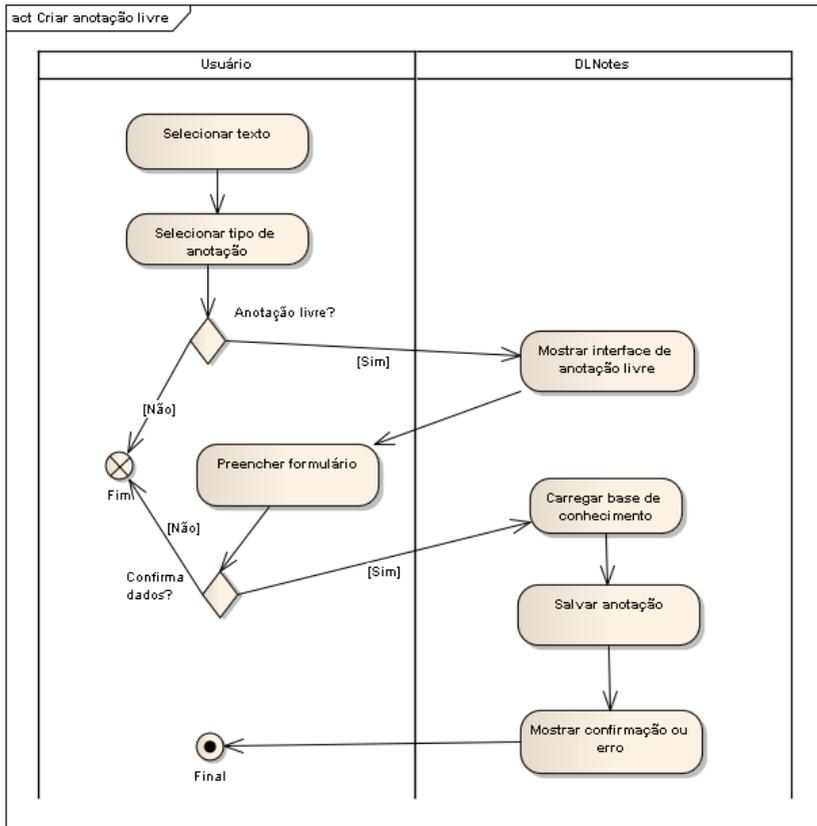
2.7 Diagrama de atividades referente ao Caso de uso “Editar anotação semântica”:



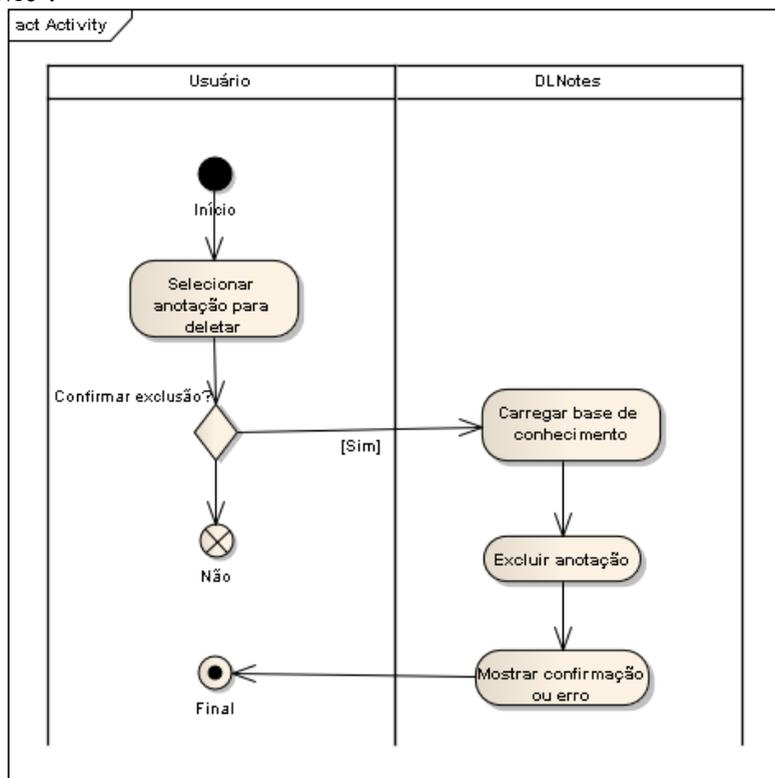
2.8 Diagrama de atividades referente ao Caso de uso “Visualizar anotação semântica”:



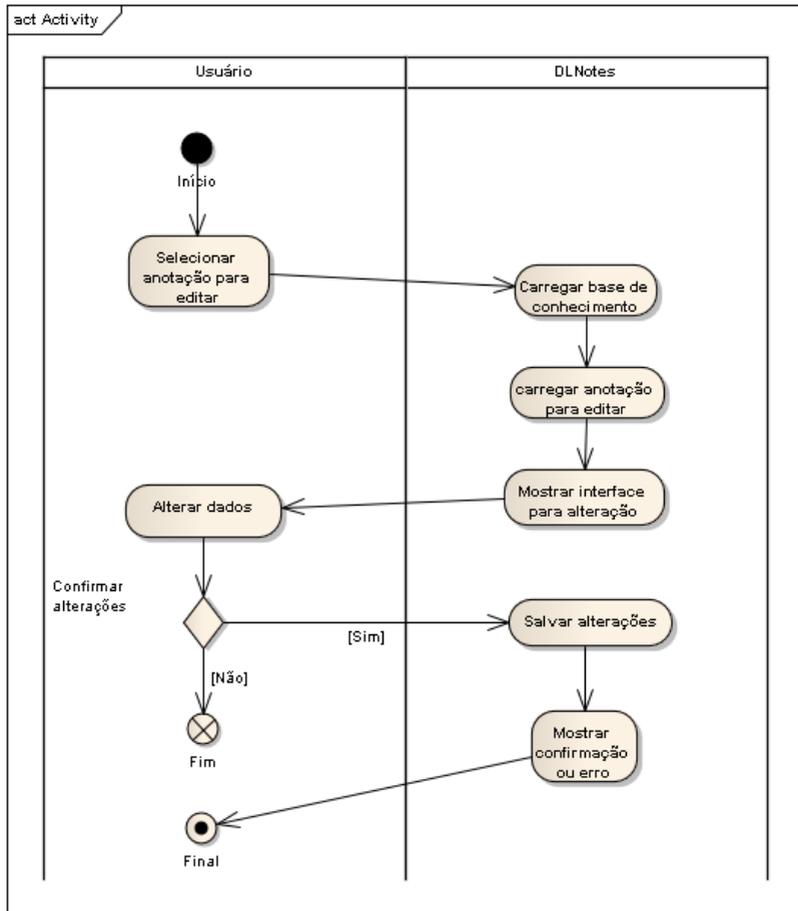
2.9 Diagrama de atividades referente ao Caso de uso “Criar anotação livre”:



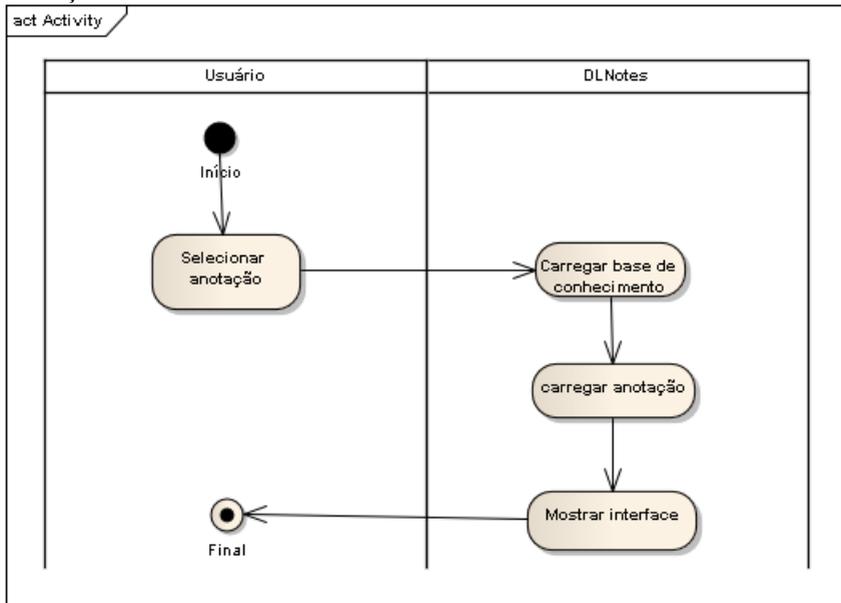
2.10 Diagrama de atividades referente ao Caso de uso “Deletar anotação livre”:



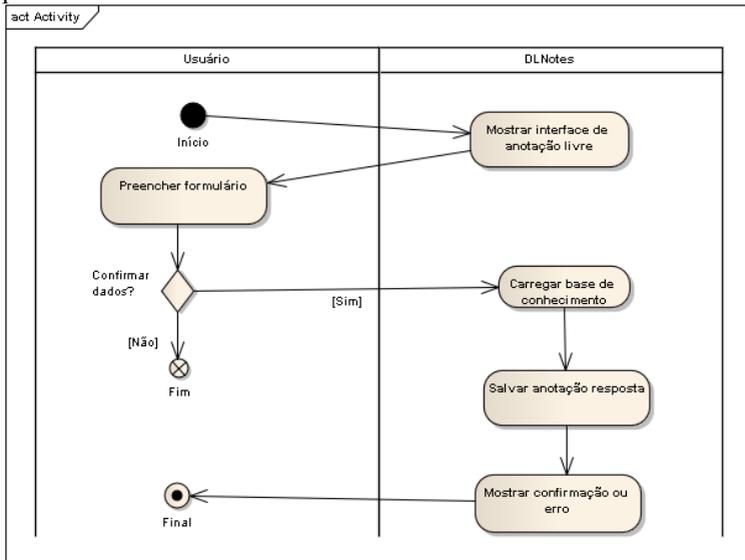
2.11 Diagrama de atividades referente ao Caso de uso “Editar anotação livre”:



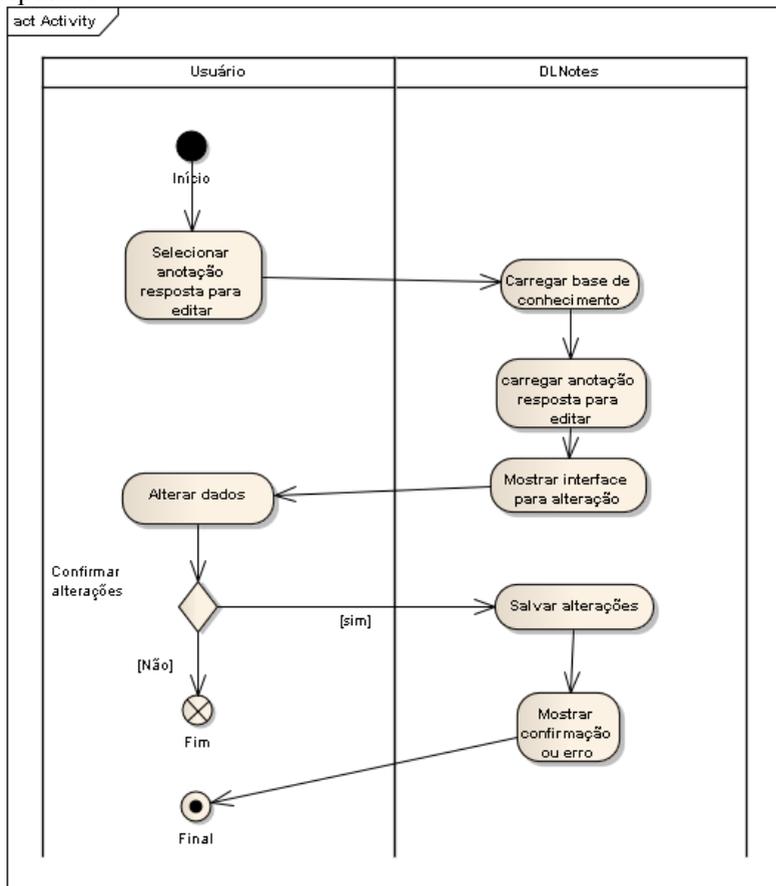
2.12 Diagrama de atividades referente ao Caso de uso “Visualizar anotação livre”:



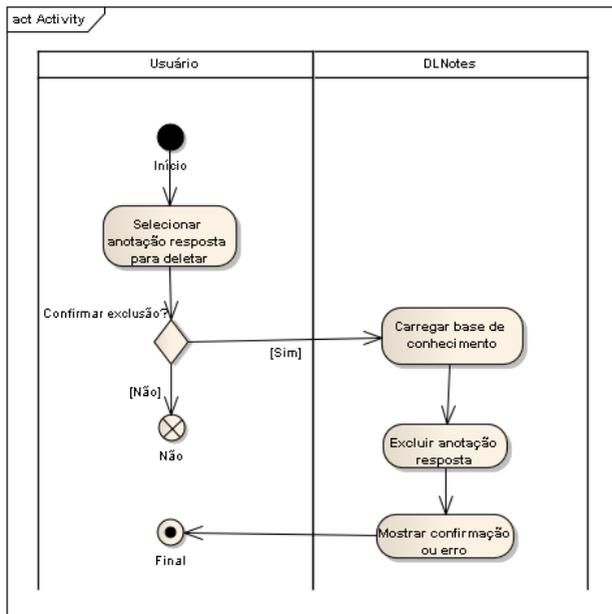
2.13 Diagrama de atividades referente ao Caso de uso “Criar anotação resposta”:



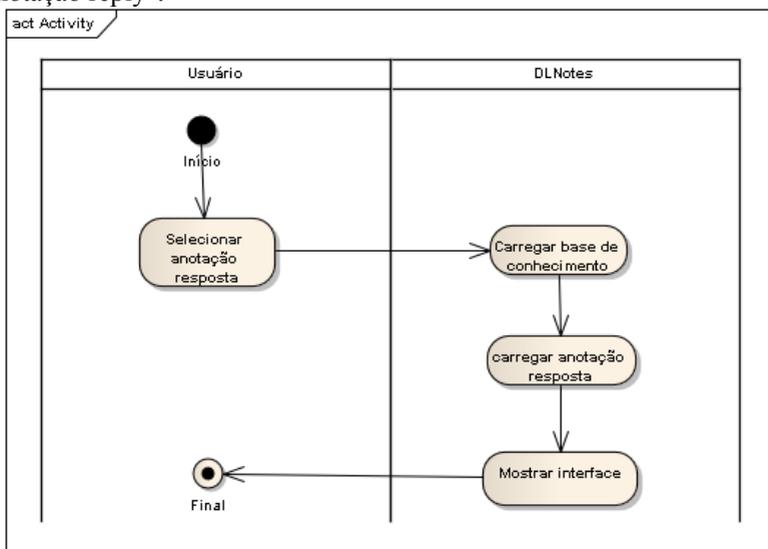
2.14 Diagrama de atividades referente ao Caso de uso “Editar anotação resposta”:



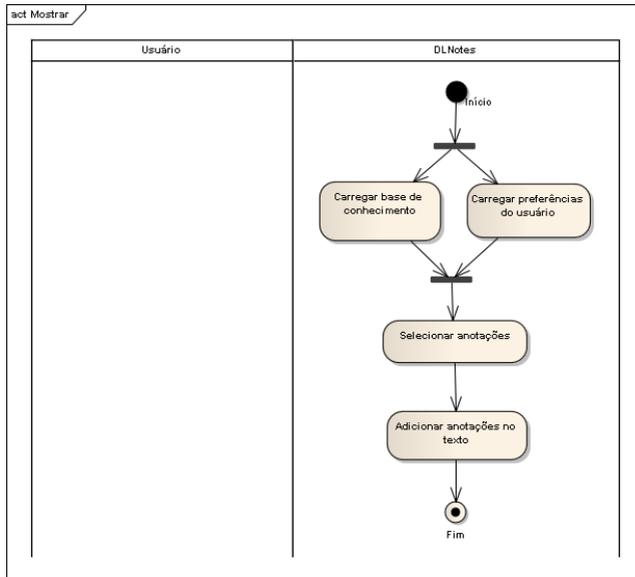
2.15 Diagrama de atividades referente ao Caso de uso “Excluir anotação reply”:



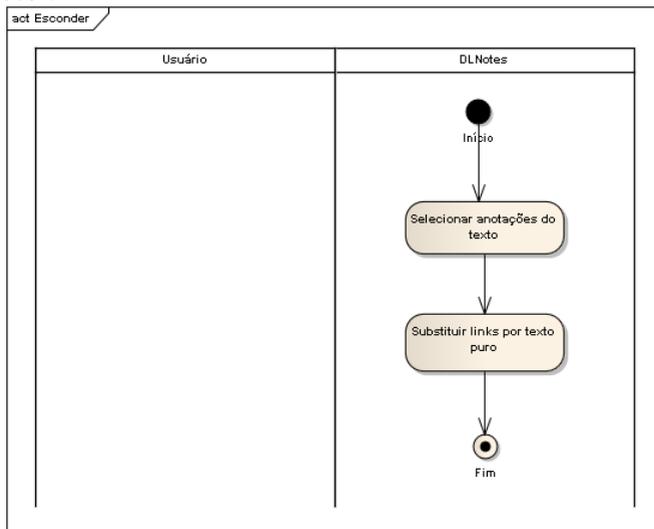
2.16 Diagrama de atividades referente ao Caso de uso “Visualizar anotação reply”:



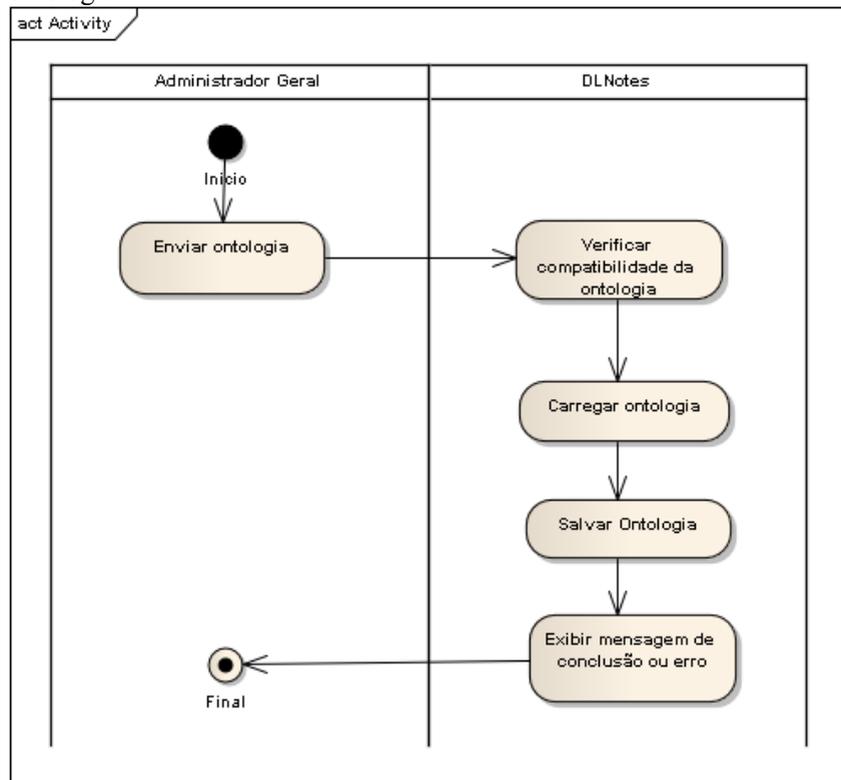
2.17 Diagrama de atividades referente ao Caso de uso “Mostrar anotações”:



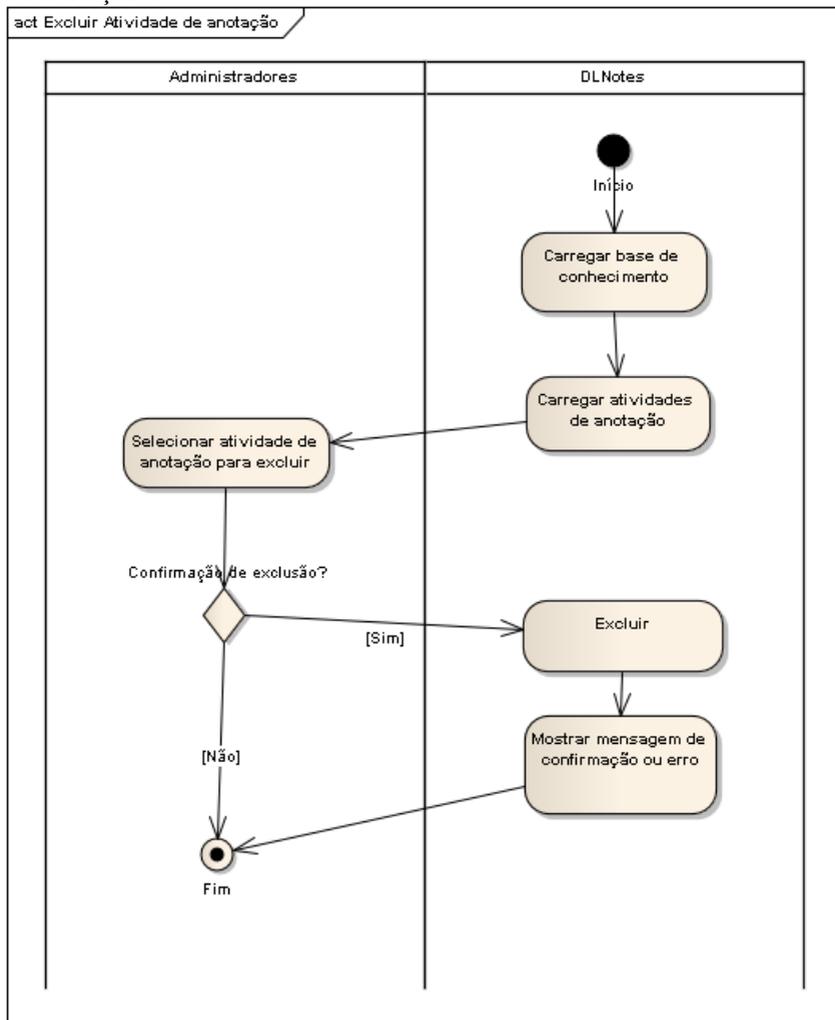
2.18 Diagrama de atividades referente ao Caso de uso “Esconder anotações”:



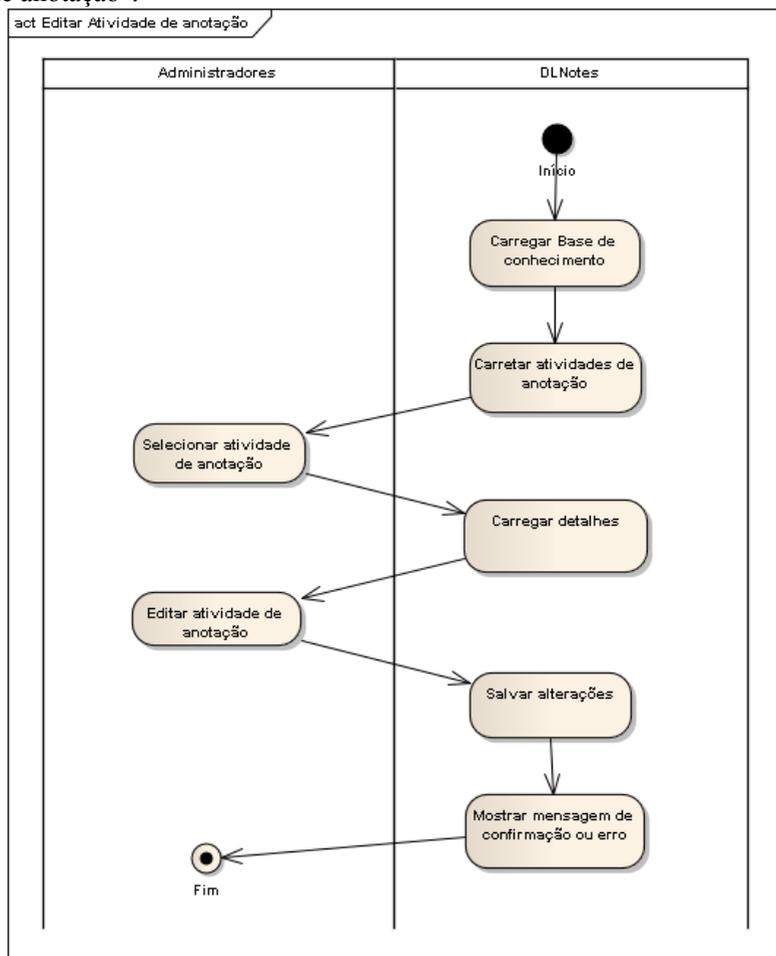
2.19 Diagrama de atividades referente ao Caso de uso “Importar ontologia”:



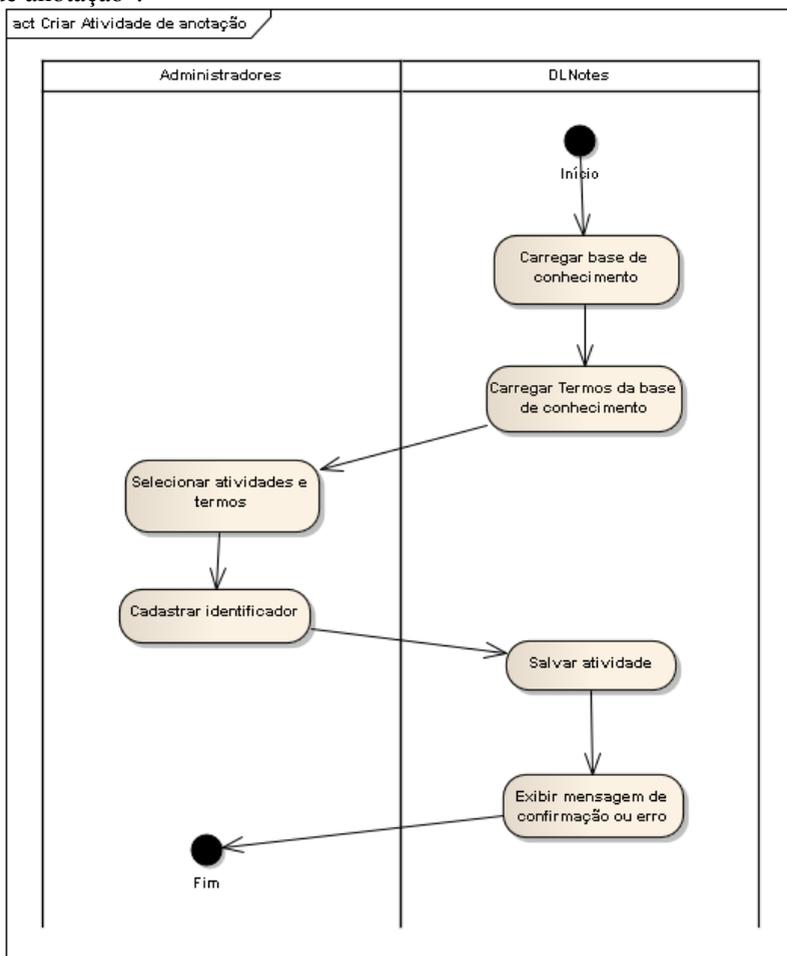
2.20 Diagrama de atividades referente ao Caso de uso “Excluir atividade de anotação”:



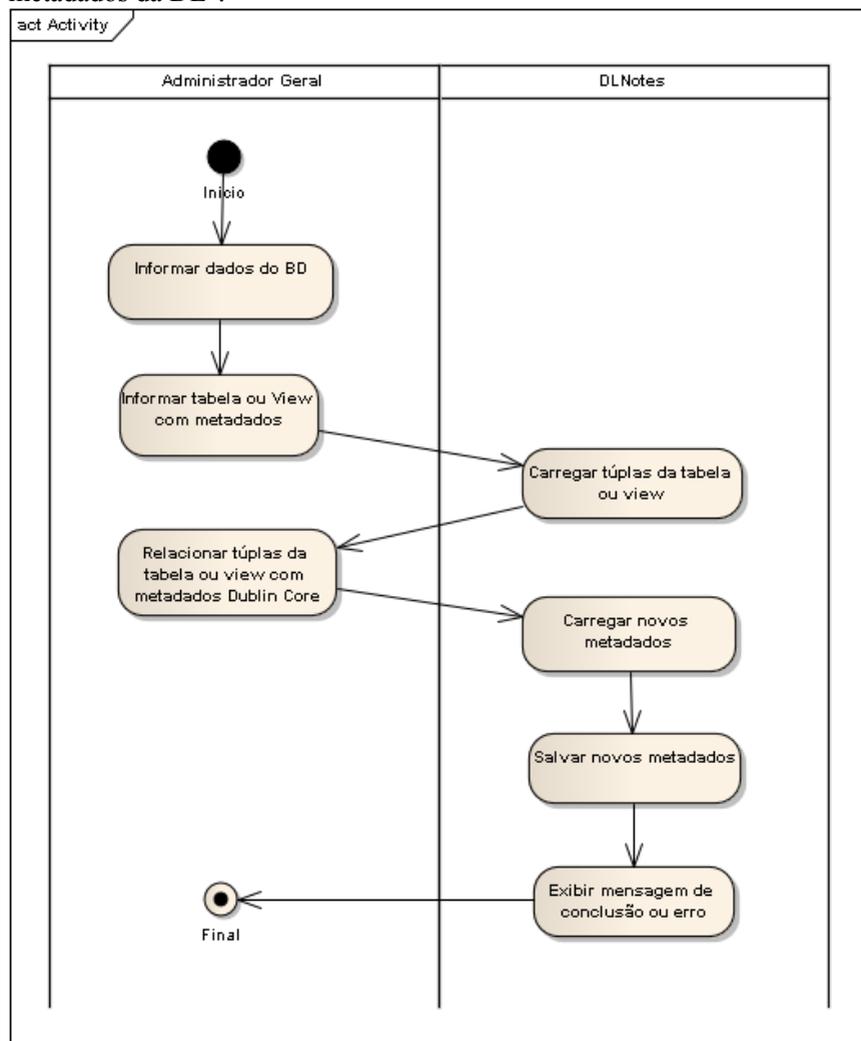
2.21 Diagrama de atividades referente ao Caso de uso “Editar atividade de anotação”:



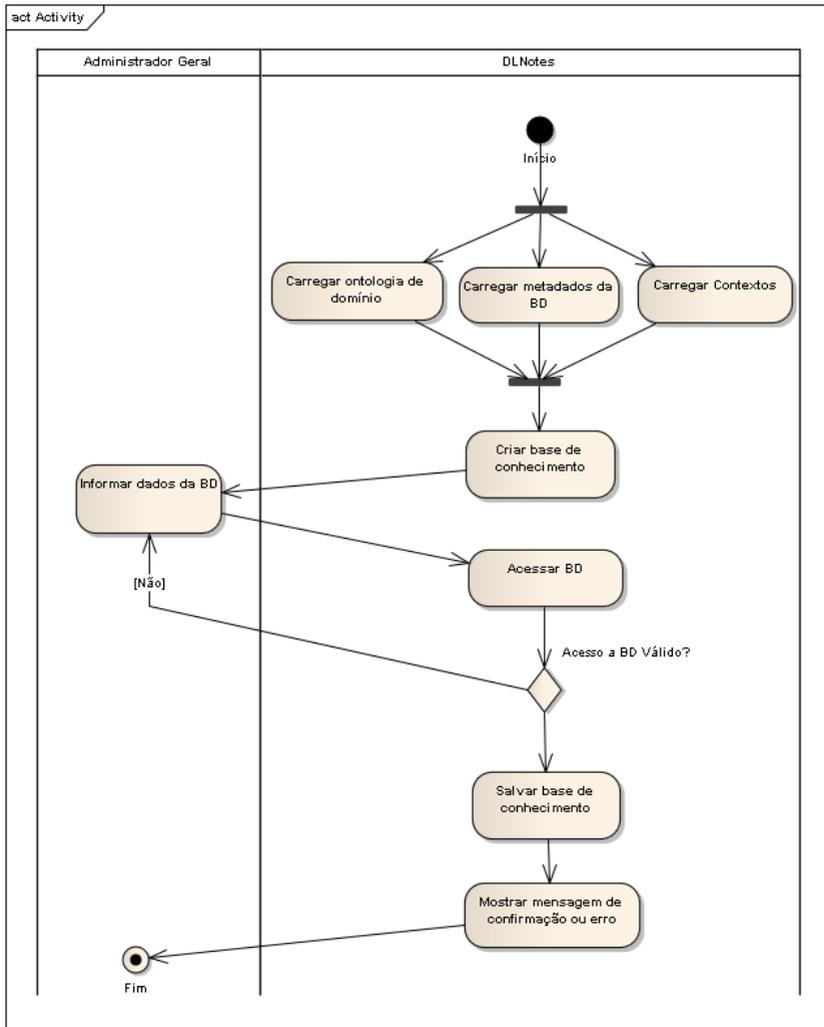
2.22 Diagrama de atividades referente ao Caso de uso “Criar atividades de anotação”:



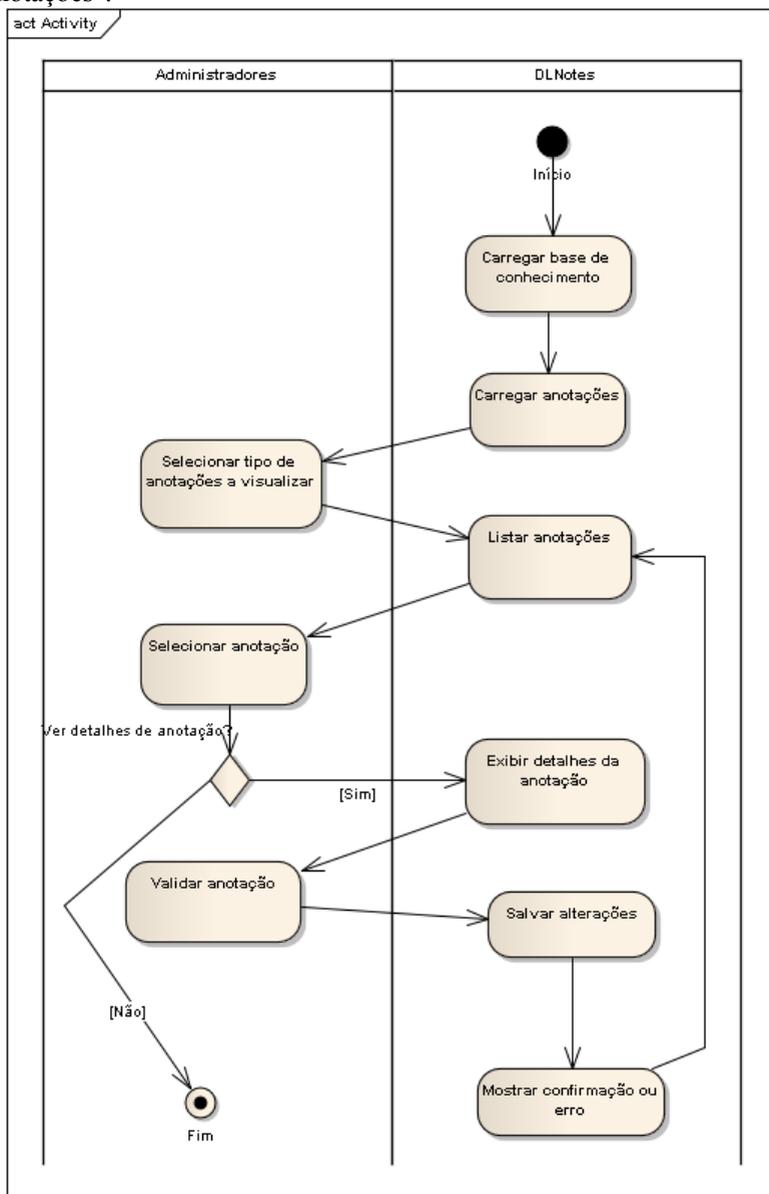
2.23 Diagrama de atividades referente ao Caso de uso “Importar metadados da DL”:



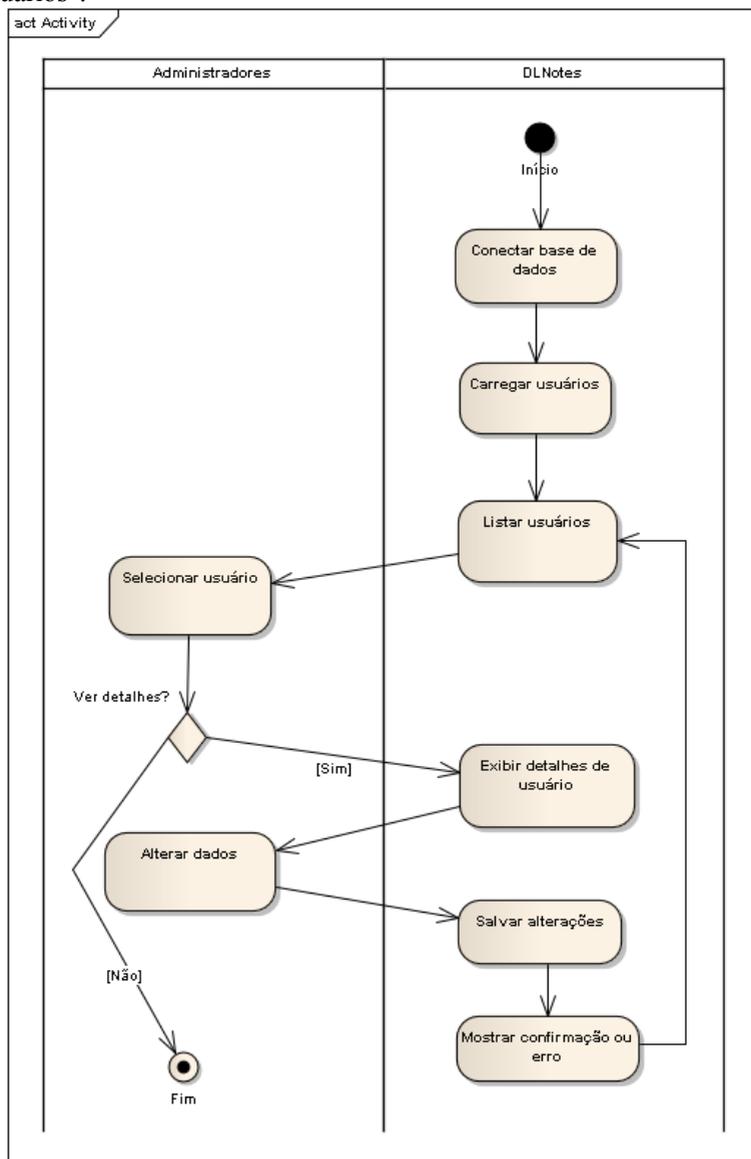
2.24 Diagrama de atividades referente ao Caso de uso “Gerar base de conhecimento”:



2.25 Diagrama de atividades referente ao Caso de uso “Gerenciar anotações”:

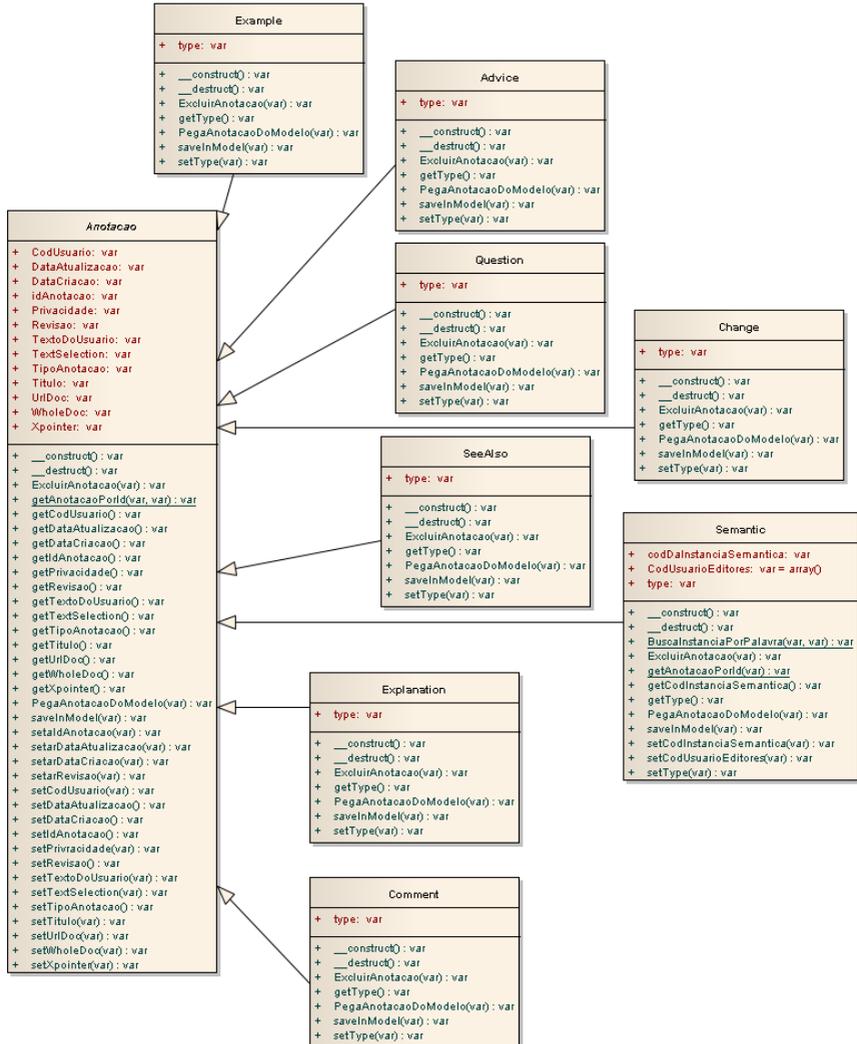


2.26 Diagrama de atividades referente ao Caso de uso “Gerenciar usuários”:

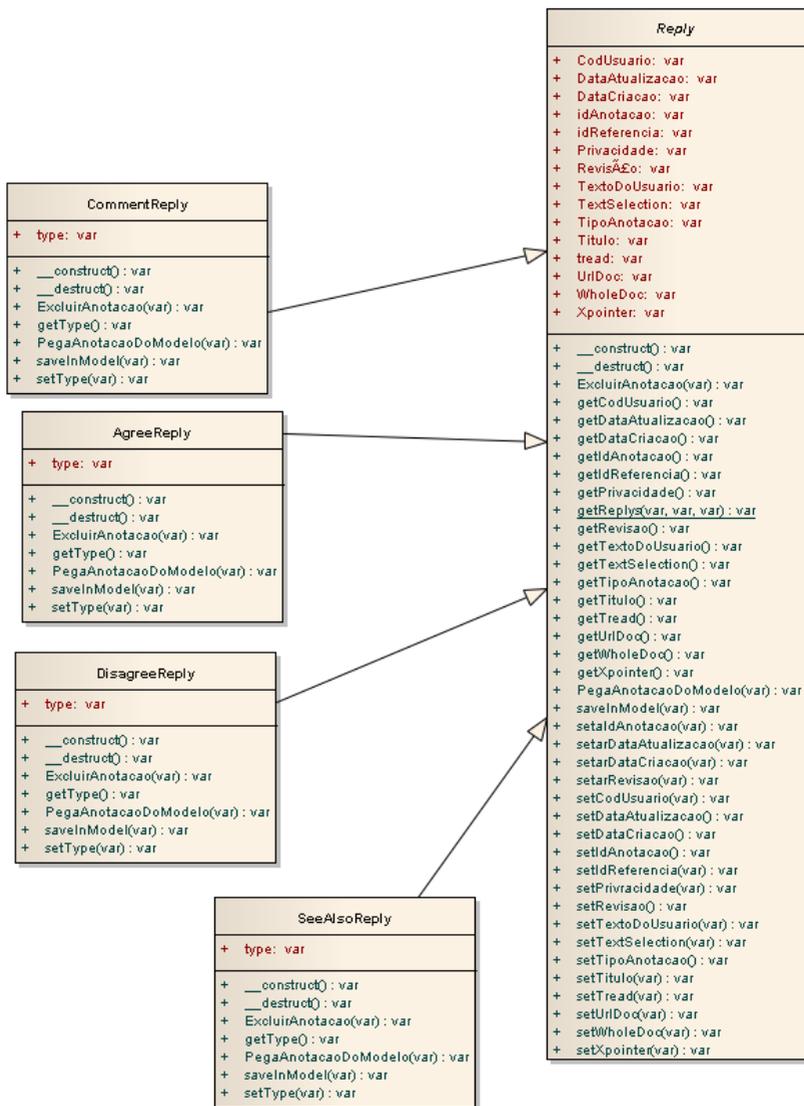


ANEXO 3 – DIAGRAMA DE CLASSES

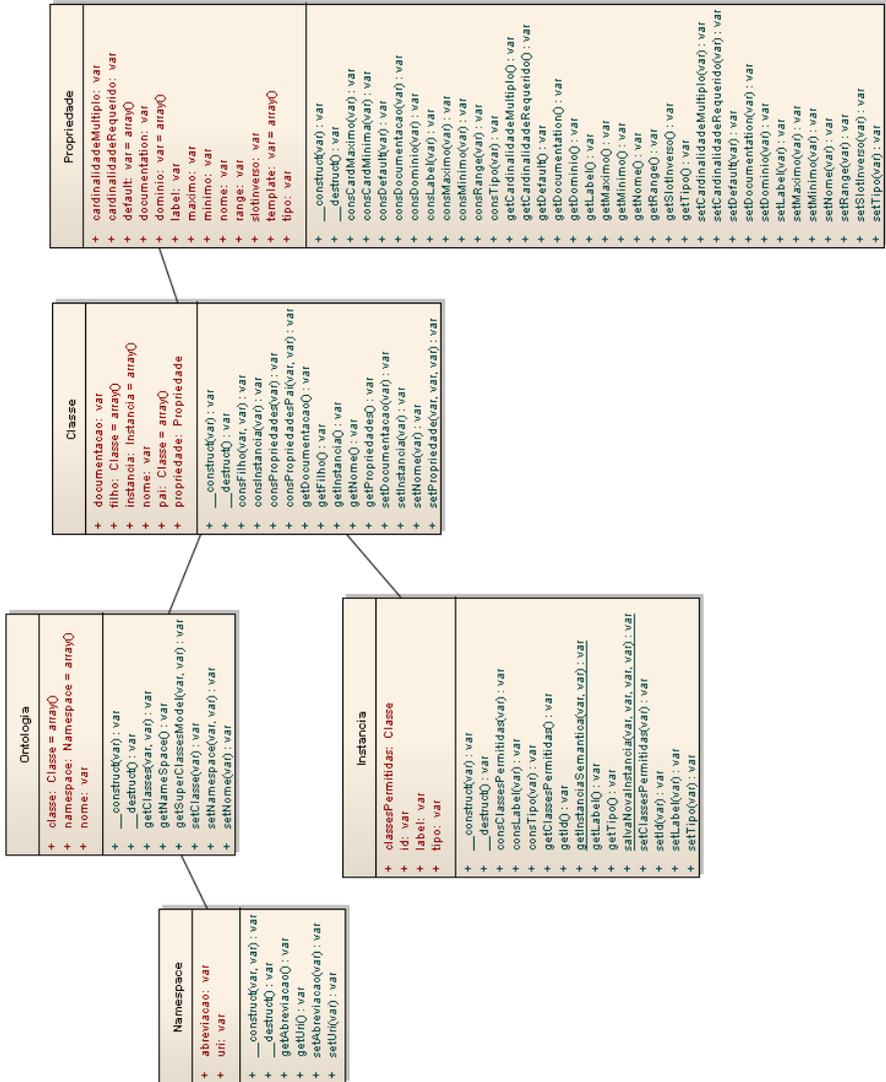
3.1 Classes anotação



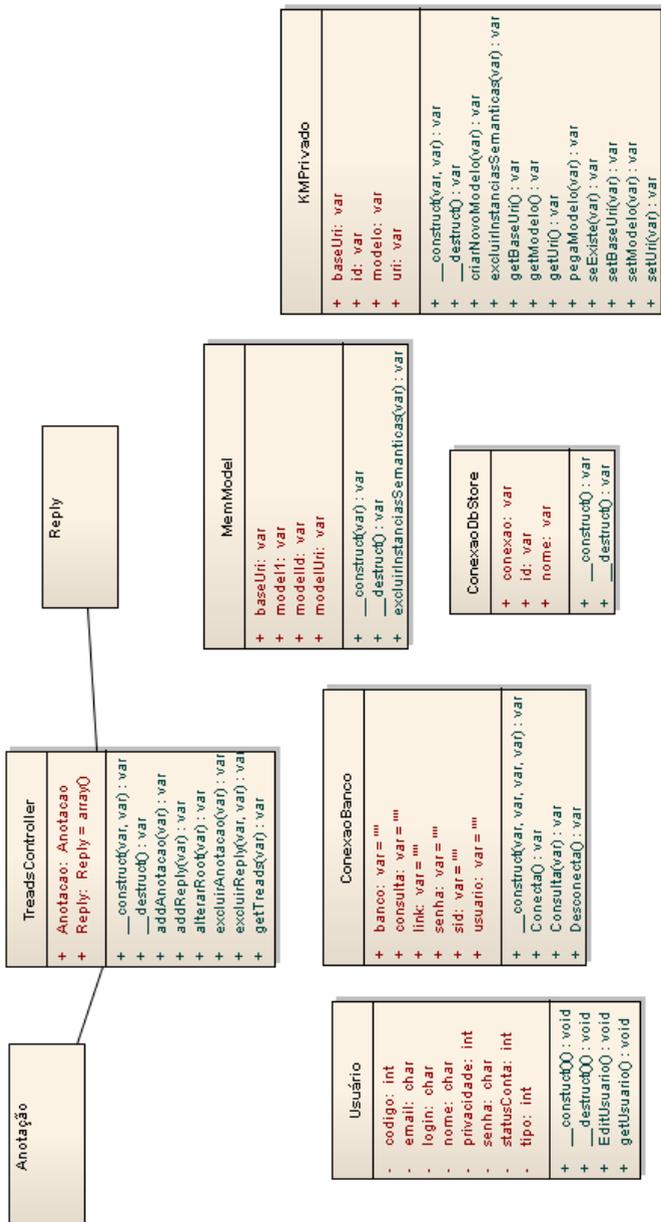
3.2 Classes anotação resposta



3.3 Classes de manipulação de ontologia



3.4 Classes de configuração e conexão



ANEXO 4 – ONTOLOGIA DE DOMÍNIO DA DL-LB

