

EP-RDF: SISTEMA PARA ARMAZENAMENTO E RECUPERAÇÃO DE IMAGENS BASEADO EM ONTOLOGIA

Edeilson Milhomem da Silva¹, Parcilene Fernandes de Brito¹

¹ Centro Universitário Luterano de Palmas (CEULP / ULBRA), Curso de Sistemas de Informação, Caixa Postal 160, CEP: 77.054-970, Palmas-TO, Brasil
{milhomem, pfb}@ulbra-to.br
Telefone para contato: (63) 218-8082 ou (63) 9994-5173
<http://www.ulbra-to.br>

Resumo

A web semântica representa a evolução da web atual. Ela fornece estruturas e possibilita a extração de significados dos conteúdos das páginas web, criando um ambiente onde agente de software e usuários possam trabalhar de forma cooperativa (BERNERS-LEE, 2001). Vários trabalhos vêm sendo desenvolvidos para criar um ambiente propício à existência de uma web com significado. Berners-Lee propôs uma arquitetura para a web semântica composta de três camadas: estrutural (permite a sistematização das informações), ontológica (fornece um conjunto de regras que define hierarquia, relacionamentos e restrições entre termos) e lógica (define um conjunto de regras de inferências). O objetivo do presente trabalho é demonstrar algumas das etapas de desenvolvimento da ferramenta de busca *EP RDF Images*, que permite a organização e a recuperação de imagens a partir da definição de uma ontologia para o domínio “imagens de instituições de ensino”, e um estudo comparativo com alguns mecanismos de buscas disponíveis na web (google.com e gettyimages.com).

Palavras-Chave: web semântica, ontologia, RDF e RDQL

II WORKSHOP DE INGENIERÍA DE SOFTWARE Y BASES DE DATOS (WISBD)

1. Introdução

A web atual, com suas características particulares, é considerada uma grande fonte de disseminação de informações. A disposição dessas informações não possui uma estrutura previamente definida, ou seja, uma estrutura única (homogênea). Isso, juntamente com a ausência de semântica dos dados, é um dos motivos que faz com que as ferramentas de busca funcionem através de indexações de palavras-chave, dificultando a coleta de informações. Assim, quando um usuário faz uma determinada busca é retornada uma quantidade elevada de informações que, às vezes, não atendem as suas necessidades. Por isso, o usuário deve realizar uma filtragem manual dessas informações. Em decorrência de tal deficiência, Berners-Lee (2001), responsável pelo projeto Web Semântica desenvolvido pela W3C, considerou que a solução seria adicionar semântica aos dados publicados na web e, a partir disso, a máquina poderia efetuar o processamento dessas informações, abstraindo um resultado semântico sobre elas.

Seguindo essa linha, a web semântica apresentará uma estrutura que além de interligar documentos, possibilitará também o reconhecimento do significado das informações contidas nesses documentos. Dessa forma, haverá uma modificação no meio atual, em que a informação é apenas apresentada para o meio, onde poderá ser, ainda, interpretada e processada. Isso será possível, segundo Berners-Lee (2001), através de agentes que funcionarão como programas coletores de conteúdos, advindos das mais variadas fontes, capazes de processar as informações e permutar resultados com outros agentes. Berners-Lee (2001) propôs uma arquitetura (Figura 1) para a web semântica composta por três camadas: uma camada de estrutura (possibilita a sistematização das informações), uma camada de ontologia (fornece um conjunto de regras que define hierarquia, relacionamentos e restrições entre termos) e uma camada lógica (definição de um conjunto de regras de inferências, além dos próprios mecanismos que as realizarão, atribuindo-as aos documentos que contêm dados).

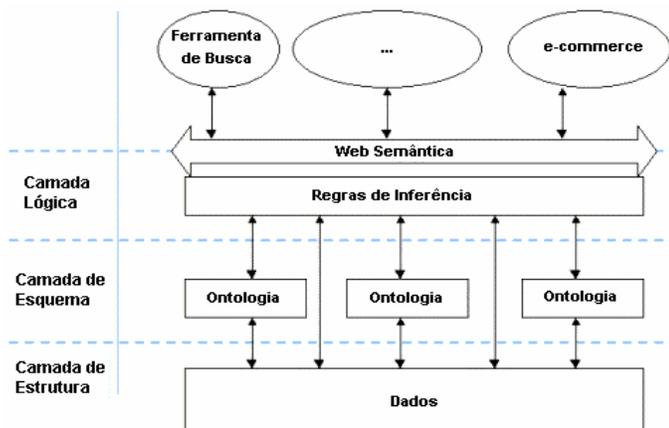


Figura 1: Arquitetura proposta para a Web Semântica (BERNERS-LEE, 2001)

O presente trabalho está vinculado à camada de estrutura e à camada de ontologia da web semântica, desta forma, a proposta visa a construção de um ambiente que utilize tecnologias capazes de propiciar uma sistematização e recuperação das informações. Para prover significados aos dados, será utilizado o padrão RDF, que, resumidamente, segundo Lassila (1999), pode ser entendido como um modelo que tem como funcionalidade prover metadados na web através da aplicação da linguagem XML.

O modelo RDF encontra-se definido em dois documentos: *Resource Description Framework (RDF) Model and Syntax Specification* (LASSILA, 1999), que descreve o modelo de dados RDF e *Resource Description Framework (RDF) Schema Specification* (BRICKLEY, 2000), que descreve as primitivas de modelagem utilizadas para a descrição de um domínio particular de interesse.

Para a recuperação das informações sistematizadas em RDF foi utilizada a linguagem de consulta RDQL, que, de acordo com Miller (2002) e Gutierrez (2003), é uma linguagem “orientada a dados”, pois ela não possibilita aplicar o mecanismo de inferência, permite apenas extrair dados de documentos RDF disponíveis na web ou armazenados em um meio físico qualquer. Para o desenvolvimento do ambiente foi utilizada a tecnologia JSP, que além de ser uma tecnologia voltada para a web, também trabalha com a linguagem Java. Isso foi um fator preponderante considerando que há uma API da linguagem Java (Jena) (JENA, 2004) desenvolvida pela HP, que dispõe de várias classes que permitem a interação com documentos RDF. Um outro ponto relevante, demonstrado neste artigo, foi a realização de uma análise qualitativa da ferramenta desenvolvida - *EP RDF Images* - em relação a alguns motores de buscas disponíveis na web.

2. A Estrutura do Sistema

Foi desenvolvida uma estrutura do funcionamento do sistema *EP RDF Images* (Figura 2), que tem como objetivo expôr as características individuais de cada módulo e das partes que os compõem. São dois os módulos: um que possibilita a descrição das imagens e o outro que possibilita a realização das consultas a essas imagens. Os funcionamentos de ambos os módulos são transparentes ao usuário, tendo em vista que o mesmo não conhece a estrutura interna de criação, armazenamento (módulo 1) e consulta às imagens (módulo 2) documentadas em RDF.

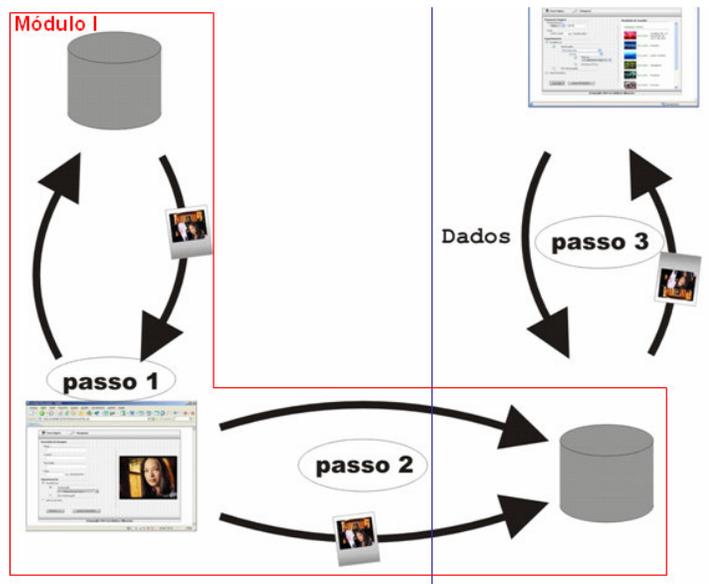


Figura 2: Partes que compõem a aplicação

Conforme representação na Figura 2 (Módulo I), havendo à disposição alguma imagem armazenada em um meio físico qualquer, o usuário, através do *browser*, pode acessar

o sistema que foi desenvolvido e informar os dados referentes a uma determinada imagem, objetivando, assim, descrevê-la (passo 1). A partir da confirmação do usuário para a respectiva descrição (passo 2), a imagem é, então, enviada e armazenada no servidor, juntamente com a descrição equivalente, ou seja, o documento RDF. É importante ressaltar que as descrições dessas imagens só poderão ser realizadas por usuários de nível administrador. Essas imagens armazenadas poderão ser, a qualquer instante, pesquisadas a partir do sistema demonstrado no Módulo II. Para a realização dessas consultas foi utilizada linguagem RDQL: uma implementação da linguagem de consulta *SquishQL*¹ para a API Jena (RDQL, 2004).

2.1 Ontologia para o domínio “Imagens de Instituição de Ensino”

Ontologia é um termo advindo da Filosofia, definido por Aristóteles como sendo a parte da filosofia “que trata do ser enquanto ser”. Na área computacional, ontologia é tratada por Uschold (1996) como “uma conceitualização explícita, formal e compartilhada, de uma área de conhecimento”. Para que uma ontologia possa ser composta e representada, são utilizadas algumas definições que podem ser visualizadas a seguir (Uschold, 1996):

- **Conceitos:** são as representações de qualquer coisa referente a um dado domínio.
- **Relacionamentos:** são as restrições, ou seja, interligações entre os conceitos de um referido domínio. Essas relações podem ser definidas formalmente como qualquer subconjunto de um produto de n conjuntos.
- **Funções:** as funções determinam os relacionamentos entre os elementos. Elas designam que os vários elementos envolvidos nesse tipo de relação são dependentes uns dos outros, embora sejam unívocos um em relação ao outro.
- **Axiomas:** representam as sentenças que restringirão a interpretação dos conceitos e relações. Essas sentenças são sempre verdadeiras. Por exemplo, um curso de graduação é composto pelo corpo docente e discente.
- **Instâncias:** representam os elementos de uma ontologia.

Não há uma regra explícita para a definição de ontologias, embora haja algumas técnicas definidas por alguns autores, a partir de experiências anteriores. Por exemplo, Uschold (1996) definiu uma metodologia a partir da definição, realizada por ele, de um conjunto de ontologias para a modelagem de empresas. Essa metodologia segue os seguintes passos: inicialmente (passo 1), deve ser definida a finalidade da ontologia (domínio a ser modelado) e o espaço da ontologia; (passo 2) deve ser construída a ontologia “capturando, codificando e integrando o conhecimento com ontologias já existentes”. Nos passos 3, 4 e 5, respectivamente, deve ser realizada a análise da ontologia criada, a sua documentação e o guia de referência para cada fase.

Constantemente novas linguagens para a definição de ontologias são desenvolvidas. Dentre essas, linguagens podem ser citadas: OIL (OIL, 2000), DAML+OIL (CONNOLLY, 2001), OWL (SMITH, 2003). Todas essas linguagens, apesar de suas particularidades, possuem uma característica em comum, a capacidade de representar ontologias em RDF e RDF *Schema*.

O RDF *Schema* dispõe de recursos que possibilita a representação de vocabulários diante da especificação de um determinado domínio. Esses recursos foram utilizados para a

¹ Uma linguagem de consulta a documentos RDF.

criação da camada ontológica, ou seja, a criação do conjunto de regras que determina a forma como serão criadas as instâncias RDF que descreverão as imagens. A Figura 3 demonstra o grafo dessa ontologia.

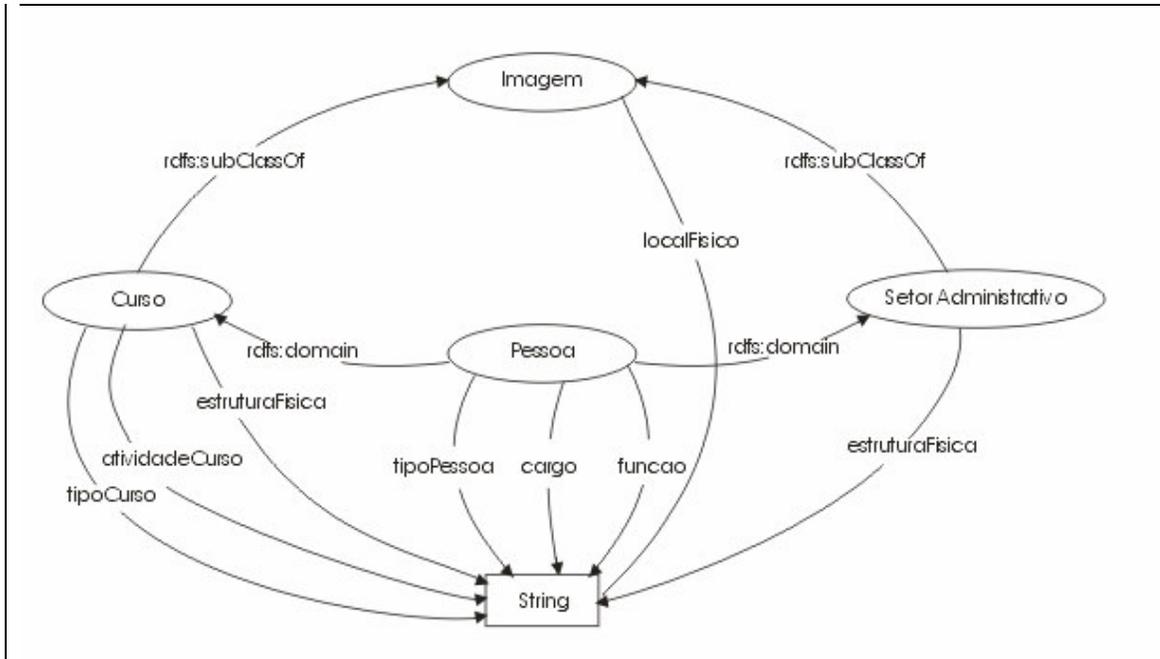


Figura 3: Grafo da ontologia criada.

As imagens podem pertencer a dois âmbitos na instituição: o acadêmico e o administrativo. O que há nos cursos, tanto de graduação quanto de pós-graduação, está ligado às atividades de ensino, pesquisa e extensão. Devido a isso, foi criada uma propriedade “atividadeCurso” que é utilizada para representá-las. Após a determinação da atividade, a imagem pode estar ligada a uma pessoa ou a uma estrutura física. Para os casos em que seja uma pessoa, foi criada uma classe “Pessoa” (que designa um dos três tipos: docente, discente e funcionário administrativo), assim como a propriedade “tipoPessoa”, que relaciona-se a esta classe. Se a pessoa for do tipo docente, estará implícito que é um professor, e que também poderá possuir um cargo. Da mesma forma, os discentes são compreendidos, com a diferença que estes não terão um cargo, porém (em alguns casos), exercerão uma determinada função na instituição (estagiário ou monitor, por exemplo). Há ainda a possibilidade da pessoa ser um funcionário administrativo e, assim, possuir somente a propriedade cargo. Caso contrário, ou seja, se a imagem for uma estrutura física, a mesma possuirá apenas um valor literal, que representará uma descrição.

2.2 Descrição das Imagens

Como dito anteriormente, dos módulos que compõem a aplicação, há um (Figura 4) que possibilita a descrição de imagens por usuários de nível administrador. Essa interface foi desenvolvida de acordo com a hierarquia estabelecida no esquema que representa o domínio.

Nova Figura | **Pesquisar**

Descrição da Imagem

Parte 1
Descrições Gerais

Título: Estagiários em um laboratório

Crédito: Coordenadora do Curso

Descrição: Estagiários do curso de Sist. Inf

Data: 15/06/2004 | ex: 02/05/2004

Departamento

Acadêmico

Graduação

Sistemas de Informação

Ensinos

Pessoa

Discente

Especificar

Estagiário

Pós-Graduação

Estrutura Física

Administrativo

Parte 2
Descrições Específicas

Parte 3
Imagem a ser Descrita

Próximo >> | Limpar Formulário

Figura 4: Tela que permite a descrição de uma imagem

A fotografia que está sendo descrita na Figura 4 é de um grupo de estagiários, alunos do curso de graduação de Sistemas de Informação e que estão vinculados à atividade ensino. O documento RDF que descreve a referida imagem pode ser visualizado na Listagem 1.

```

1. <rdf:RDF
2.   xmlns:rdf='http://www.w3.org/1999/02/22-rdf-syntax-ns#'
3.   xmlns:Imagem='http://www.ulbra-to.br/imagem/esquema#'
4.   xmlns:dc='http://purl.org/dc/elements/1.0/'>
5.   <rdf:Description rdf:nodeID='A1'>
6.     <rdf:type rdf:resource='http://www.ulbra-to.br/imagem/esquema#Imagem' />
7.     <dc:title>Estagiários em um Laboratório </dc:title>
8.     <dc:creator>Igor Flávio</dc:creator>
9.     <dc:date>15/06/2004</dc:date>
10.    <dc:description>Estagiários do curso de graduação de Sist. Inf. </dc:description>
11.    <Imagem:localFisico>documentos/jpeg/foto_625.jpg</Imagem:localFisico>
12.    <Imagem:curso rdf:resource='SistemasDeInformacao' />
13.  </rdf:Description>
14.  <rdf:Description rdf:about='SistemasDeInformacao'>
15.    <Imagem:atividadeNoCurso>Ensino</Imagem:atividadeNoCurso>
16.    <Imagem:tipoCurso>Graduacao</Imagem:tipoCurso>
17.    <Imagem:pessoa rdf:nodeID='A0' />
18.  </rdf:Description>
19.  <rdf:Description rdf:nodeID='A0'>
20.    <Imagem:tipoPessoa>Discente</Imagem:tipoPessoa>
21.  </rdf:Description>
22. </rdf:RDF>

```

Listagem 1: Documento RDF gerado que descreve a imagem representada na parte 2 da Figura 4.

Como forma de possibilitar um vínculo entre a imagem (.jpeg, .png, .bmp ...) e o documento RDF que a descreve, esse mesmo documento possui uma propriedade (linha 12) que recebe, como valor, o local físico onde a imagem está armazenada.

2.3 Processo de Armazenamento das Descrições das Imagens

Para cada imagem descrita pela aplicação é criada uma instância de um documento RDF. Para evitar durante o processo de busca a realização de consultas em todas essas instâncias, foi

criado um documento RDF intermediário. Esse documento tem a função de tornar as consultas (Módulo II) mais rápidas (otimizadas), pois funciona como um elemento norteador das consultas, ou seja, esse documento intermediário direciona para quais documentos serão realizadas as pesquisas, por exemplo, se está sendo consultada imagens referente ao curso de Sistemas de Informação, não há porque serem feitas consultas desnecessárias em outros cursos. Esse documento intermediário obedece a hierarquia estabelecida pela ontologia que rege a forma de descrição das imagens. Partindo-se do pré-suposto que foram descritas apenas 7 (sete) imagens, referentes a três cursos distintos, o documento intermediário estaria organizado da seguinte forma (Listagem 2).

```
1. <rdf:RDF
2.   xmlns:rdf='http://www.w3.org/1999/02/22-rdf-syntax-ns#'
3.   xmlns:Intermediario ='http://www.ulbra-to.br/imagem/Intermediario'>
4.   <rdf:Description rdf:about='curso/graduacao/SistemasDeInformacao/pessoa'>
5.     <Intermediario:arquivo>imagens/imagem_0.rdf</Intermediario:arquivo>
6.   </rdf:Description>
7.   <rdf:Description rdf:about='setorAdm/estruturaFisica'>
8.     <Intermediario:arquivo>imagem_1.rdf</NS0:arquivo>
9.     <Intermediario:arquivo>imagem_2.rdf</NS0:arquivo>
10.  </rdf:Description>
11.  <rdf:Description rdf:about='curso/graduacao/Administracao/estruturaFisica'>
12.    <Intermediario:arquivo>imagem_3.rdf</Intermediario:arquivo>
13.    <Intermediario:arquivo>imagem_4.rdf</Intermediario:arquivo>
14.    <Intermediario:arquivo>imagem_5.rdf</Intermediario:arquivo>
15.    <Intermediario:arquivo>imagem_6.rdf</Intermediario:arquivo>
16.  </rdf:Description>
17. </rdf:RDF>
```

Listagem 2: Documento RDF que mapeia as consultas

Para a explicação desse documento intermediário, será considerada a imagem descrita na Figura 4, que diz respeito a uma pessoa do curso de graduação de Sistemas de Informação. Para representação das imagens descritas que se encaixem nessa hierarquia da ontologia, foi criado um recurso (linha 4 a 6) com um URI de valor "curso/graduacao/SistemasDeInformacao/pessoa". Esse mesmo recurso possui uma propriedade arquivo (linha 5), que tem o objetivo de armazenar o nome do documento que descreve a imagem pertencente a esse nível da hierarquia.

2.4 Consultas às Imagens

Como forma de recuperar as imagens descritas a partir da utilização do sistema demonstrado no Módulo I, foi desenvolvido um segundo módulo que tem a responsabilidade de realizar pesquisas diante das informações fornecidas por um usuário. As pesquisas podem ser realizadas por quaisquer usuários, independente dele estar *logado* no sistema. No momento da pesquisa, se nenhuma informação é fornecida, são retornadas todas as imagens descritas, classificadas por: âmbito acadêmico e administrativo. Supondo que esse mesmo usuário queira fazer uma filtragem mais detalhada de uma pesquisa, é disponibilizada uma interface que possibilita tal operação (Figura 5).

Pesquisar imagens

Pesquisar por:
Nenhum

Data
ex: 02/05/2004

Departamento

Acadêmico

Graduação
---- Selecione um Curso ----

Pós-Graduação

Administrativo

Consultar Limpar Formulário

Figura 5: Tela de pesquisa

Como a linguagem de consulta RDQL, utilizada no trabalho para a realização das consultas, não disponibiliza a cláusula LIKE, que possibilita a realização de consultas de *string* com valores aproximados, a realização de pesquisas com valores aproximados só foi possível a partir da re-filtragem das informações, obtidas a partir da execução de uma determinada consulta RDQL. Dessa forma, para qualquer pesquisa que seja realizada, são feitas duas filtragens: a primeira é realizada através da utilização da linguagem RDQL, que utiliza para as cláusulas de restrições os valores advindos da parte do formulário que contém as informações específicas da imagem. A segunda filtragem é feita a partir do retorno obtido pela primeira. Pois, como a execução da primeira filtragem traz as informações gerais (título, crédito, descrição e data), torna-se possível realizar a manipulação desses valores, podendo assim, ser feita uma análise parcial das *strings*.

Uma outra limitação da linguagem RDQL é a falta de suporte a operadores lógicos, como, por exemplo, o operador lógico AND. Mesmo assim, a aplicação suportou a utilização de tal operador. Para isso, o esquema de implementação foi igual ao citado anteriormente, que possibilita a realização de consultas com *strings* de valores parciais (LIKE), em que é realizada uma seleção, tendo como base as restrições permitidas pela parte específica do formulário (parte 2 – Figura 4). A partir desse retorno, é feita uma re-filtragem dessas informações. O operador lógico AND é utilizado para parte das informações gerais (título, crédito e descrição) e o que determina sua utilização é um espaço em branco entre as palavras que compõem a *string* para pesquisa. Supondo que fosse realizada uma pesquisa, que dentre outras informações, contém o valor de descrição “encoinfo 2004”, seriam retornadas todas as imagens que contém em sua descrição as palavras “encoinfo” e “2004”, independente da ordem dos termos. O resultado da presente consulta pode ser visualizado na Figura 6.



Figura 6: Demonstração do resultado obtido pela busca

O resultado da pesquisa trouxe apenas 4 ocorrências, mas o suficiente para ressaltar que o retorno das pesquisas é classificado de acordo com o próximo nível, conforme a hierarquia estabelecida pelo esquema. Por exemplo, para as especificações realizadas, que trouxe como retorno as imagens representadas na Figura 6, referente à atividade extensão, o próximo nível seria a verificação das imagens relacionadas a discente, docente ou Funcionário Administrativo. Como foram encontradas imagens apenas de discentes, a classificação é feita apenas para eles.

3. Trabalho Correlato

Um dos primeiros passos para o desenvolvimento da *EP RDF Images* foi a realização de uma pesquisa que objetivou a coleta de informações a respeito de trabalhos correlatos. Do resultado dessa pesquisa, foi obtido o trabalho disponível em (GUUS, 2001) em que é utilizado o modelo de dados RDF para a descrição de imagens referente ao reino animal.

Assim como na *EP RDF Images*, em (GUUS, 2001) foi criada uma ontologia, que representa as regras de criação das instâncias RDF que descrevem as imagens. Para isso, foi utilizada a ferramenta *Protegê-2000 Ontology*, que gera um documento RDFS equivalente.

Na *EP RDF Images*, as imagens são descritas a partir de instâncias de documentos RDF e as consultas realizadas a partir da RDQL, diferentemente da ferramenta demonstrada em (GUUS, 2001), em que tanto a camada ontológica, as triplas RDF que descrevem as imagens e as consultas são representadas por distintos módulos de uma base de dados Prolog. Para a geração dessas triplas, que são utilizadas posteriormente pela ferramenta, é utilizada a ferramenta *SWI-Prolog RDF parser*². A ferramenta (GUUS, 2001) suporta o carregamento e a

² Versão demo disponível em <http://hcs.science.uva.nl/projects/SWI-Prolog/packages/sgml/online.html>

descrição dos arquivos de imagens (as fotografias) e, posteriormente, realiza o armazenamento dessas descrições (os documentos RDF). Uma arquitetura geral do sistema pode ser visualizada na Figura 7.

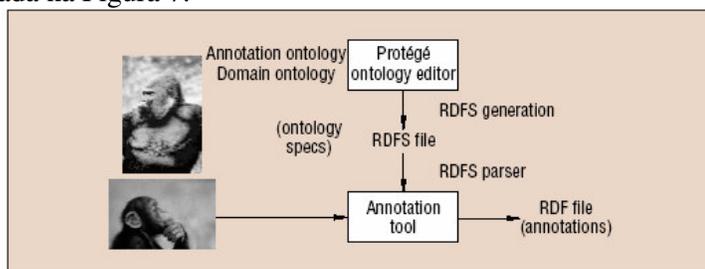


Figura 7: Arquitetura geral do sistema (GUUS, 2001)

Conforme supracitado, a ferramenta demonstrada em (GUUS, 2001) representa as triplas RDF em uma base de dados Prolog, principal diferencial entre ela e a *EP RDF Images*. Essa base é composta por três módulos: um que representa os dados ontológicos, ou seja, as regras de criação das instâncias RDF, outro que sistematiza as informações, ou seja, que representa as descrições das imagens e, por último, o módulo responsável pela realização das consultas.

4. Comparação Qualitativa com outras Ferramentas de Busca

Para quantificar o grau de funcionamento e eficiência das consultas realizadas na ferramenta *EP RDF Images*, foi realizado um estudo comparativo com alguns mecanismos de buscas existentes na web.

Na primeira parte da avaliação, foram considerados os resultados da pesquisa obtidos pelo motor de busca de imagens do *google.com*. Para isso, foram consideradas consultas simples, por exemplo, imagens de *teacher in classroom*. A submissão dessa consulta no referido mecanismo de busca trouxe como resultado 24.400 ocorrências, um número relativamente alto. Assim, é difícil estimar quais as ocorrências são realmente sobre *teacher in classroom*. Por isso, para a realização da análise, foram considerados os primeiros 200 resultados. Dessas 200 imagens apenas 69 eram referentes ao tema *teacher in classroom*, o que demonstra uma eficiência de apenas 34,5%. Das 131 imagens restantes 92 estavam relacionadas a imagens quaisquer, fora do contexto ensino, como por exemplo, imagens referentes a grupo de operários, paisagens ou até mesmo logomarcas. As demais 39 imagens estavam dentro do contexto ensino, com a diferença que 33 delas estavam relacionadas a imagens de, por exemplo, um aluno com a expressão de que está prestando atenção, um quadro negro, uma pessoa pegando um livro na biblioteca e a capa de um livro guia para o professor em sala de aula, e as 6 demais diziam respeito a imagens de professores fora da sala de aula.

Na segunda parte de análise de performance, utilizando a mesma consulta (*teacher in classroom*), foram considerados os resultados obtidos pelo motor de busca do *gettyimages.com*. Assim como na primeira análise, foram consideradas apenas as primeiras 200 ocorrências, de um total de 875 imagens. Desse total, 167 imagens estavam exatamente de acordo com consulta, o que estima uma eficiência de 83,5%. Das restantes, 21 estavam relacionadas ao contexto ensino, 10 não puderam ser exibidas (não encontradas) e apenas 2 estavam fora do contexto.

Ambos os mecanismos de busca (*gettyimages.com* e *google.com*) desconsideraram a palavra “in”, submetida inicialmente para a consulta, alegando ser uma palavra muito comum.

O *gettyimages.com* ainda permite que a consulta seja melhor especificada: "*teacher*" and "*Classroom (School)*" or "*teacher*" and "*Lecture Hall (University)*". Para esse estudo, foram consideradas as duas especificações. Foi verificado, a partir dessas pesquisas, que o motor de busca do *gettyimages.com* parece utilizar algum mapeamento parecido com ontologia para representação do conhecimento, o que justificaria o fato de uma melhor performance (grau de eficiência das consultas) em relação ao *google.com*. Um outro fato que justificaria a melhor performance do *gettyimages.com* em relação ao *google.com*, é o fato do último armazenar suas informações através de índices (faz constantes buscas em toda a web), diferentemente do *gettyimages.com*, em que as imagens estão disponíveis em um único meio (não faz buscas pela web).

Essa mesma consulta foi submetida a ferramenta *EP RDF Images* e trouxe como resultado uma eficiência de 100%. Mas, esse resultado era esperado, pois, a ferramenta *EP RDF Images* possui as imagens documentadas em RDF, o que permite que os dados sejam contextualizados a partir de determinados domínios, proporcionando assim a semântica necessária a determinados contextos. Neste caso, como o contexto é "instituição de ensino", a busca sobre algo relacionado a esse tema teria um alto grau de eficiência.

5. Conclusão

O presente trabalho objetivou, a partir do entendimento sobre os conceitos de ontologia, RDF, RDF Schema e RDQL, a demonstração de alguma das etapas de desenvolvimento de uma aplicação (*EP RDF Images*) que possibilitasse a sistematização e a recuperação de imagens descritas em RDF. Após o desenvolvimento da ferramenta, foi realizado um estudo comparativo entre a *EP RDF Images* e algumas das ferramentas de busca disponíveis na web: *google.com* e *gettyimages.com*. Nessa análise, foi considerada a mesma consulta (*teacher in classroom*) para os distintos mecanismos de busca, *google.com*, *gettyimages.com* e *EP RDF Images*, e os resultados advindos dessa busca solidificou a idéia de que a definição de contextos é fator preponderante na organização dos dados.

Como trabalhos futuros, podem ser acrescentadas novas funcionalidades à aplicação, como, por exemplo, o desenvolvimento de um módulo que possibilite a criação automática de esquemas, a partir de informações fornecidas pelo usuário, e a geração automática de interfaces adaptativas a partir do esquema criado. Desta forma, a partir da criação do esquema, o usuário poderá ter controle sobre o conteúdo e a disposição da informação na interface que será gerada. Para representação hierárquica dessas informações e o conjunto de dados que cada propriedade definida no esquema poderá assumir, pode ser utilizado um documento XML, pois um documento XML, por si só, possui um comportamento hierárquico. Para que o usuário não possa ficar "livre" para criar esquemas quaisquer, pode ser definida uma ontologia "geral", que determinará as regras de criação das demais. Para isso, o domínio poderia ser formalizado a partir da utilização de uma linguagem para ontologias, como, por exemplo, a linguagem OWL, que permite a definição do domínio de uma maneira mais complexa, dado o fato de ser possível através dela definir uma quantidade maior de axiomas.

6. Referências

Berners-Lee, T., Hendler, J. e Lassila, O. (2001) "The Semantic Web". Magazine Content. Disponível em: <http://www.scientificamerican.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21&catID=2>. Último acesso em: 15 de fevereiro de 2005.

- Brickey, D.; Guha R. V. (2000) (eds.) Resource Description Framework (RDF) Schema. Disponível em: <http://www.w3.org/TR/2000/CR-rdf-schema-20000327/>. Último acesso em: 22 de fevereiro de 2005.
- Connolly, D.; Harmelen, F.van; Horrocks, I.; McGuinness, D. L.; Patel-Schneider, P. F.; Stein, L. A. (2001) “DAML+OIL (MARCH 2001)” Recomendado pela W3C. Disponível em: <http://www.w3.org/TR/daml+oil-reference>. Último acesso em: 22 de fevereiro de 2005.
- “Description of OIL” (2004). Disponível em: <http://www.ontoknowledge.org/oil>. Último acesso em: 22 de fevereiro de 2005.
- Gutierrez, C., Hurtado, C. e Mendelzon, A.. (2003) “Formal aspects of querying RDF databases”, Semantic Web and Databases (SWDB), Alemanha.
- GUUS, Schreiber, DUBBELDAM, Barbara, WIELEMAKER, Jan, WIELINGA Bob, Ontology-Based Photo Annotation, Volume 16, Issue 3, Pages: 66 - 74 (May 2001) IEEE Intelligent Systems
- “Jena – A Semantic Web Framework for Java” (2004) Hewlett-Packard Development Company. Disponível em: <http://jena.sourceforge.net/license.html> Último acesso em: 14 de março de 2005.
- Lassila, O., Swick, Ralph R. (1999) “Resource Description Framework (RDF) Model and Syntax Specification”. Recomendado pela W3C. Disponível em: <http://www.w3.org/TR/REC-rdf-syntax/>. Último acesso em: 15 de março de 2005.
- Miller, L., Seaborne, A. e Reggiori, A..(2002) “Three Implementations of SquishQL, a Simple RDF Query Language”, International Semantic Web Conference on The Semantic Web (ISWC), vol. 2342 p. 423 – 435.
- “RDQL query” (2004): Disponível em: <http://www.hpl.hp.com/semweb/rdql.htm#RDQL> Usage Último acesso em: 25 de março de 2005.
- Smith, M. K., Welty, C.; McGuinness, D. L. (2003) “OWL Web Ontology Language Guide”, Recomendado pela W3C. Disponível em: <http://www.w3.org/TR/2003/CR-owl-guide-20030818/>. Último acesso em: 24 de março de 2005.
- Uschold M.; Gruninger M. (1996) “Ontologies: principles, methods, and applications”, Knowledge Engineering Review, vol. 11(2), 93--136, 1996.