



Marcelo Cohen de Azevedo

Gerador de aplicações para consultas a bases RDF/RDFS

Dissertação de Mestrado

Dissertação apresentada ao Programa de Pós-graduação em Informática da PUC-Rio como requisito parcial para obtenção do título de Mestre em Informática.

Orientador: Prof. Daniel Schwabe

Rio de Janeiro
Janeiro de 2010



Marcelo Cohen de Azevedo

Gerador de aplicações para consultas a bases RDF/RDFS

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Informática da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Daniel Schwabe

Orientador

Departamento de Informática - PUC-Rio

Prof.^a Simone Diniz Junqueira Barbosa

Departamento de Informática - PUC-Rio

Prof. Edward Hermann Haeusler

Departamento de Informática - PUC-Rio

Prof. José Eugênio Leal

Coordenador Setorial do Centro Técnico Científico - PUC-Rio

Rio de Janeiro, 07 de janeiro de 2010

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Marcelo Cohen de Azevedo

Graduou-se em Engenharia de Computação na PUC-Rio em 2007. Atuou como Analista de Sistemas e Programador em diversos projetos inovadores, sempre pesquisando e aplicando novas tecnologias. Possui interesse acadêmico e profissional nas áreas de Hípermidia, Engenharia de Software e Sistemas distribuídos.

Ficha Catalográfica

Azevedo, Marcelo Cohen de

Gerador de aplicações para consultas a bases RDF/RDFS / Marcelo Cohen de Azevedo ; orientador: Daniel Schwabe. – 2010.

134 f. : il. ; 30 cm

Dissertação (Mestrado em Informática) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2010.

Inclui bibliografia.

1. Informática – Teses. 2. Web semântica. 3. Geração de aplicações. 4. Ontologias. 5. RDF. 6. RDFS. I. Azevedo, Marcelo Cohen de. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

Dedico esse trabalho aos meus pais,
pelo apoio e amor incondicional.

Agradecimentos

Gostaria de agradecer ao departamento de informática da PUC-Rio e todo seu corpo docente por me proporcionar tamanha oportunidade de aprendizado.

Às agências de fomento à pesquisa CAPES e CNPq, pelo apoio financeiro que possibilitou este trabalho.

Aos membros da banca pela paciência e disponibilidade.

Ao meu orientador pelas idéias inspiradoras e entusiasmo contagiante.

Ao Cristiano Rocha, Daniel Orlean, Francisco Ferreira e José Leme pelo apoio e compreensão.

À Marena Cutnei pela ajuda em diversos momentos deste trabalho.

Aos meus amigos e familiares por tudo.

Resumo

Azevedo, Marcelo Cohen; Schwabe, Daniel. **Gerador de aplicações para consultas a bases RDF/RDFS**. Rio de Janeiro, 2010. 134p. Dissertação de Mestrado - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Com o crescimento da web semântica, cada vez mais bases de dados em RDF contendo todo tipo de informações, nos mais variados domínios, estão disponíveis para acesso na Internet. Para auxiliar o acesso e a integração dessas informações, esse trabalho apresenta uma ferramenta que permite a geração de aplicações para consultas a bases em RDF e RDFS através da programação por exemplo. Usuários podem criar casos de uso através de operações simples em cima do modelo RDFS da própria base. Esses casos de uso podem ser generalizados e compartilhados com outros usuários, que podem reutilizá-los. Com esse compartilhamento, cria-se a possibilidade desses casos de uso serem customizados e evoluídos colaborativamente no próprio ambiente em que foram desenvolvidos. Novas operações também podem ser criadas e compartilhadas, o que contribui para o aumento gradativo do poder da ferramenta. Finalmente, utilizando um conjunto desses casos de uso, é possível gerar uma aplicação web que abstraia o modelo RDF em que os dados estão representados, tornando possível o acesso a essas informações por usuários que não conheçam o modelo RDF.

Palavras-chave

Web Semântica; Geração de aplicações; Ontologias; RDF; RDFS.

Abstract

Azevedo, Marcelo Cohen; Schwabe, Daniel (Advisor). **An application builder for quering RDF/RDFS datasets.** Rio de Janeiro, 2010. 134p. MSc. Dissertation - Departamento de Informática, Pontificia Universidade Católica do Rio de Janeiro.

Due to increasing popularity of the semantic web, more data sets, containing information about varied domains, have become available for access in the Internet. This thesis proposes a tool to assist accessing and exploring this information. This tool allows the generation of applications for querying databases in RDF and RDFS through programming by example. Users are able to create use cases through simple operations using the RDFS model. These use cases can be generalized and shared with other users, who can reuse them. The shared use cases can be customized and extended collaboratively in the environment which they were developed. New operations can also be created and shared, making the tool increasingly more powerful. Finally, using a set of use cases, it's possible to generate a web application that abstracts the RDF model where the data is represented, making it possible for lay users to access this information without any knowledge of the RDF model.

Keywords

Semantic Web; Application Builder; RDF; RDFS.

Sumário

1	Introdução	15
2	Descrição do problema e motivação	20
3	Estado da arte	24
3.1	A linguagem de consultas SPARQL	24
3.2	O acesso a bases RDF através de aplicações pré-definidas	26
3.3	Ferramentas semânticas	27
4	Solução proposta	30
4.1	Excelplorator	31
4.2	O Explorator	39
4.3	Novos conceitos desenvolvidos	44
4.3.1	Operações customizadas	45
4.3.1.1	Transdutores e entrada de parâmetros via texto	55
4.3.2	Nova organização da área de trabalho	59
4.3.3	A parametrização de consultas	62
4.3.4	Compartilhamento aberto de consultas	75
4.3.5	O compartilhamento fechado de consultas	79
4.3.5.1	A construção da aplicação	80
4.3.5.2	A customização da camada de apresentação da aplicação gerada	98
5	Implementação da ferramenta proposta	109
5.1	A arquitetura utilizada	109
5.2	O novo modelo	111
5.2.1	A classe Application	113
5.2.2	A classe Workbench	114
5.2.3	A classe ResourceSet	115
5.2.4	A classe Operation	116

5.2.5	A classe Operand	116
5.2.6	A classe Parameter	117
5.2.7	A classe ParameterLink	117
5.2.8	A classe Transducer	117
5.2.9	A classe View	118
6	Avaliação	119
6.1	Dimensões Cognitivas	119
6.2	Avaliação com usuários	123
6.2.1	Resultados obtidos	125
7	Conclusões e trabalhos futuros	127
7.1	Contribuições	127
7.2	Trabalhos futuros	128
8	Referências bibliográficas	131

Lista de figuras

Figura 1 – Tela inicial da aplicação gerada.	33
Figura 2 – As âncoras de acesso aos casos de uso.	33
Figura 3 – Lista com os autores da base.	34
Figura 4 – Publicações do autor selecionado.	35
Figura 5 – Detalhes de uma publicação.	35
Figura 6 – Formulário para entrada de valores pelo usuário.	36
Figura 7 – Resultado de uma busca no Explorator	40
Figura 8 – Operação diferença no Explorator	40
Figura 9 – Resultado da operação diferença	41
Figura 10 – Visualização de um Recurso	42
Figura 11 – Propriedades existentes na base	43
Figura 12 – A operação SPO no Explorator	43
Figura 13 – Resultado da operação SPO	44
Figura 14 – Selecionando Classes na tela de desenvolvimento do Excelplorator	46
Figura 15 – Criando conjunto com todos os recursos do tipo “Product”	47
Figura 16 – Menu principal com a opção que gera o conjunto com as propriedades destacada.	48
Figura 17 – Operação SPO sendo executada	49
Figura 18 – Conjunto de triplas gerado pela operação SPO	49
Figura 19 – Opção “Operations” no menu principal do Excelplorator	50
Figura 20 – Lista das Operações disponíveis e formulário de criação.	51
Figura 21 – Formulário de criação de novas operações preenchido.	52
Figura 22 – Lista de operações disponíveis atualizada.	53
Figura 23 – Índice das operações disponíveis com o novo filtro habilitado.	53
Figura 24 – Drop-down com as operações habilitadas.	54
Figura 25 – Nova operação de filtro sendo utilizada para filtrar conjunto de triplas D.	54

Figura 26 – Criação do conjunto de triplas que possui a solução da tarefa do usuário.	55
Figura 27 – Opção do menu principal para criação de uma entrada de texto.	56
Figura 28 – Transdutor de texto criado para entrada de dados pelo usuário.	56
Figura 29 – Realizando a operação de filtro parametrizada com o valor 240.	57
Figura 30 – Novo conjunto de triplas criado (Conjunto H) através da operação de filtro parametrizada.	58
Figura 31 – Triplas do conjunto H são recalculadas quando o parâmetro “INPUT A” é atualizado.	59
Figura 32 – A barra inferior controla os workbenchs da aplicação.	60
Figura 33 – Basta um clique no nome para poder atualizá-lo de forma a manter a organização do ambiente.	61
Figura 34 – Botão para a criação de novo workbench.	61
Figura 35 – Seleção de um novo workbench ativo. O preenchimento cinza indica o workbench selecionado.	61
Figura 36 – Botão de remoção de workbench.	62
Figura 37 – Opção do menu que retorna um conjunto de triplas com as Classes da base.	63
Figura 38 – Através da classe Person, é possível criar um conjunto de triplas com todas as instâncias de pessoa da base.	64
Figura 39 – Realização de uma operação SPO, com uma pessoa na posição do Sujeito.	65
Figura 40 – Conjunto com as triplas que possuem a pessoa escolhida como sujeito. Entre as triplas, a propriedade “Made”	66
Figura 41 – Realização da operação SPO, S sendo o conjunto de triplas previamente criado e P sendo a propriedade “Made”	67
Figura 42 – Botão de edição do conjunto de triplas, localizado no topo do conjunto, logo acima do tooltip em verde da figura.	68
Figura 43 – Conjunto de triplas K no modo de edição.	69
Figura 44 – Opção de parametrização do conjunto de triplas.	70

Figura 45 – Opção para re-nomear o nome do parâmetro do conjunto. Tem como objetivo facilitar o entendimento da semântica do parâmetro.	71
Figura 46 – Parâmetro p1, que tinha “Steffen Staab” como valor, sendo substituído por “Alexander Kubias”.	72
Figura 47 – A mesma tela instantes após a substituição, o conjunto de triplas L, que depende do conjunto K, teve seus valores recalculados de forma instantânea e automática.	72
Figura 48 – O novo conjunto de triplas M, que possui três pessoas da base na posição do Sujeito.	73
Figura 49 – Conjunto M sendo arrastado para substituir o valor do parâmetro p1, do conjunto K.	74
Figura 50 – Logo após a substituição, o conjunto L já conta com seus novos valores, decorrentes da atualização do valor do parâmetro p1 do conjunto K.	75
Figura 51 – Caso de uso parametrizado em um workbench que será compartilhado. No detalhe, o workbench sendo re-nomeado.	77
Figura 52 – Opção do menu para compartilhamento e importação de workbenches.	77
Figura 53 – Tela com a lista de workbenches compartilhados pelo usuário e a lista de workbenches disponibilizados por outros usuários.	78
Figura 54 – Edição da descrição de um caso de uso que será compartilhado.	78
Figura 55 – A última coluna indica se o caso de uso está sendo compartilhado.	79
Figura 56 – Um dos casos de uso compartilhados por outros usuários.	79
Figura 57 – Após a importação, uma cópia é criada na sessão do usuário atual.	79
Figura 58 – Após a criação do conjunto de triplas K, que possui a propriedade “made” na posição do predicado.	82
Figura 59 – Momento da execução de uma operação SPO, que criará um novo conjunto.	82
Figura 60 – Conjuntos de triplas sendo re-nomeados.	83
Figura 61 – Workbench sendo re-nomeado.	83

Figura 62 – Autor da publicação sendo parametrizado para tornar possível a reutilização da consulta.	84
Figura 63 – Definição do conjunto de triplas como ponto de partida para o caso de uso.	85
Figura 64 – Adicionando o comportamento de passagem de valor para parâmetro ao conjunto de triplas.	86
Figura 65 – Após o comportamento ter sido adicionado, uma indicação informa que o valor do recurso na posição do sujeito que for selecionado, será usado como parâmetro do outro conjunto de triplas.	86
Figura 66 – Criação de um conjunto com as propriedades desejadas através de uma operação SPO.	88
Figura 67 – Parametrizando conjunto de triplas, depois do mesmo ter sido criado e re-nomeado.	89
Figura 68 – Resultado final: os três conjuntos de triplas (no modo de edição) que juntos implementam o caso de uso.	90
Figura 69 – Criação de um conjunto que possui as publicações na posição do sujeito e suas palavras-chaves na posição de Objeto.	92
Figura 70 – O transdutor de texto e os dois conjuntos de triplas gerados.	93
Figura 71 – Conjunto de triplas gerado pela operação do filtro.	94
Figura 72 – Link que leva o usuário ao ambiente de utilização da aplicação montada.	95
Figura 73 – Link que o usuário deve acessar para se autenticar.	96
Figura 74 – Usuário deve prover o seu provedor de identidade que suporte o protocolo open-id.	96
Figura 75 – Opção do menu para salvar a aplicação corrente ou carregar aplicações salvas.	97
Figura 76 – Índice de aplicações salvas e formulário para salvar a aplicação atual.	97
Figura 77 – Opção do menu que leva ao controle de customização de visualização.	99
Figura 78 – Tela de customização/criação de novas interfaces.	99
Figura 79 – A opção de descartar customização surge quando componente é modificado.	101

Figura 80 – Lista de novas interfaces criadas para customização da aparência da aplicação.	102
Figura 81 – Seleção do componente de interface que vai renderizar o conjunto de triplas.	103
Figura 82 – Índice de pessoas que é exibido após o acesso ao caso de uso.	105
Figura 83 – No comportamento padrão, após a seleção de uma pessoa, o índice de pessoas some da tela e um novo índice com as publicações é exibido.	106
Figura 84 – Já com o comportamento customizado, o índice de pessoas permanece na tela mesmo após a seleção de uma pessoa.	106
Figura 85 – No comportamento padrão, novas janelas surgem à esquerda das existentes.	107
Figura 86 – O comportamento pode ser customizado para que um conjunto seja exibido à direita dos conjuntos existentes.	108
Figura 87 – Modelo do Explorator.	111
Figura 88 – Modelo do Excelplorator	113