



Leonardo Magela Cunha

Um Framework de Aplicações para a Web Semântica

Tese de Doutorado

Tese apresentada como requisito parcial para obtenção do título de Doutor pelo Programa de Pós-Graduação em Informática da PUC-Rio.

Orientador: Carlos José Pereira de Lucena

Rio de Janeiro
Dezembro de 2006



Leonardo Magela Cunha

Um Framework de Aplicações para a Web Semântica

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

Carlos José Pereira de Lucena
Orientador
PUC-Rio

Marco Antonio Casanova
PUC-Rio

Simone Diniz Junqueira Barbosa
PUC-Rio

Renato Fileto
UFSC

Michel Klein
Vrije Universiteit

José Eugenio Leal
Coordenador(a) Setorial do Centro Técnico Científico - PUC-Rio

Rio de Janeiro, 18 de dezembro de 2006

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

Leonardo Magela Cunha

Graduou-se Bacharel em Informática pela Universidade Federal de Juiz de Fora (UFJF) em 1999. Obteve o título de Mestre em Informática pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) em 2002.

Ficha Catalográfica

Cunha, Leonardo Magela

Um framework de aplicações para a Web semântica / Leonardo Magela Cunha ; orientador: Carlos José Pereira de Lucena. – 2006.

131 f. : il. (col.) ; 30 cm

Tese (Doutorado em Informática)–Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2006.

Inclui bibliografia

1. Informática – Teses. 2. Web semântica. 3. Engenharia de software. 4. Aplicações para a Web semântica. 5. Frameworks. 6. Semantic Web challenge. I. Lucena, Carlos José Pereira. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

Para o Bruno, Dani Brauner e Anarosa.

Agradecimentos

Ao meu orientador, Carlos Lucena, pelo aprendizado, oportunidades, paciência e confiança no meu trabalho.

Aos membros da comissão avaliadora: Casanova, Fileto, Simone, Michel, Anarosa e Karin, por aceitarem o convite para a participação e por seus questionamentos e sugestões.

À Viviane, Anarosa e Dani Brauner, por conseguirem muitas vezes me fazer enxergar caminhos que eu não acreditava poder trilhar.

Aos colegas, professores e funcionários da PUC-Rio, do Departamento de Informática e do LES, por seu apoio e companheirismo. Em especial, à Vera, minha mãe da PUC, que mesmo quando não percebe, pode instaurar revoluções e mudanças com um olhar, um simples abraço ou uma palavra amiga. À Dani Küsel, por seus conselhos e companhia em vários momentos.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico, à PUC-Rio e à Fundação Padre Leonel Franca, pelo apoio financeiro.

Ao Bruno, por me acompanhar em novas jornadas sem saber o destino, por me apoiar e me dar a mão quando eu mais precisava, por ser um bom conselheiro, por acreditar em mim nos momentos em que nem mesmo eu podia mais. Não conseguiria listar aqui tudo que ele fez, mas foi muito e lhe sou muito grato por isso.

Aos meus amigos, companheiros e familiares, por sempre estarem disponíveis a me ouvir e a oferecer uma palavra de conforto. Em especial, ao Allyson, à minha mãe, meu pai e minha irmã, por seus ouvidos e corações imensuráveis.

A Deus, por tudo que consegui e por todos que estiveram em meu caminho.

Resumo

Cunha, Leonardo Magela; Lucena, Carlos José Pereira de. ***Um Framework de Aplicações para a Web Semântica***. Rio de Janeiro, 2006. 131p. Tese de Doutorado - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Até alguns anos atrás, a *Web* disseminava principalmente documentos. Com o advento das aplicações *Web*, as organizações puderam disponibilizar informações que estavam em seus bancos de dados e sistemas legados. Entretanto, a comunicação entre estas aplicações ou com aplicações de usuários finais, às vezes, não era possível devido a diferenças no formato de representação dos dados. O desenvolvimento de padrões (*standards*) e o uso da eXtensible Markup Language (XML) resolveram muitos destes problemas. Apesar das soluções desenvolvidas serem somente sintáticas elas funcionam em muitos casos, como por exemplo, na interoperabilidade de esquemas em sistemas *business to business* de *e-commerce*. Entretanto, a falta do aspecto semântico impossibilitou que as aplicações fizessem mais uso dos dados ou os utilizassem de forma mais “inteligente”. A idéia da *Web Semântica* é definir explicitamente o significado dos dados que se encontram na *Web*. Com isso, esperam-se aplicações capazes de “entender” o que significam os dados. E uma vez que estas aplicações entendam os dados, elas possibilitarão que os usuários utilizem essa nova *Web* “dirigida a dados” para facilitar as suas tarefas rotineiras. Esta tese propõe um *framework* para o desenvolvimento de aplicações para a *Web Semântica*. Considerando o que foi descrito no parágrafo anterior, o número de aplicações que podem ser construídas é quase infinito. Portanto, nós nos restringimos a observar as aplicações que tem por objetivo solucionar o problema apresentado pelo *Semantic Web Challenge*; e propor um *framework* que represente estas soluções. O *Challenge* tem como principal finalidade demonstrar como as aplicações podem atrair e beneficiar o usuário final através do uso das técnicas da *Web Semântica*. Conseqüentemente, nossa intenção é possibilitar que o desenvolvedor de aplicações possa atingir essa atração e benefícios, através do uso das técnicas de *Web Semântica* e de Engenharia de Software, utilizando um *framework* para o desenvolvimento das aplicações.

Palavras-chave

Web Semântica; Engenharia de Software; Aplicações para a *Web Semântica*; *Frameworks*; *Semantic Web Challenge*

Abstract

Cunha, Leonardo Magela; Lucena, Carlos José Pereira de. *Um Framework de Aplicações para a Web Semântica*. Rio de Janeiro, 2006. 131p. Tese de Doutorado - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Documents have been the main vehicle of the Web until some years ago. With the advent of Web applications, data stored in organizations' databases or legacy systems has been made available to users. However, very often, the exchange of data between those applications themselves or between them and "end-users applications" were not possible since they used different formats for the information representation. The development of standards and the use of the eXtensible Markup Language (XML) solved parts of the problem. That was a syntactic solution and it works for several cases, e.g., schema interoperability in Business-to-Business e-commerce scenarios. Nevertheless, the lack of semantics on these data prevented applications to take more advantage of them. The idea behind the Semantic Web is to define explicitly the semantics of data available on the Web. Therefore, we expect another step forward where applications, being them corporative or for end-users, will "understand" the meaning of the data available on the Web. Once those applications can understand it, they will be able to help users to take advantage of this "data driven" Web and to perform their daily tasks easily. This thesis proposes a framework for the development of Semantic Web applications. Considering the scenario described in the previous paragraph, the number of possible applications that can be developed is almost infinite. For this reason, we restricted ourselves to examine the solutions that aim to solve the problem presented at the Semantic Web Challenge; and to propose a framework that represent those solutions. The challenge is concerned in demonstrating how Semantic Web techniques can provide valuable or attractive applications to end users. Our main concern was then to demonstrate and help a developer to achieve that value addition or attractiveness, through Semantic Web techniques, in a Software Engineering approach using frameworks.

Keywords

Semantic Web; Software Engineering; Semantic Web Applications; Frameworks; Semantic Web Challenge

Sumário

1	Introdução	17
1.1.	O Problema	19
1.2.	Solução Proposta	20
1.3.	Objetivos	21
1.4.	Contribuições	21
1.5.	Trabalhos Relacionados	22
1.6.	Resumo	23
2	Web Semântica	24
2.1.	Ontologias	25
2.2.	O W3C e a Web Semântica	27
2.3.	A Controvérsia Sobre a <i>Semantic Web Stack</i>	29
2.4.	Resumo	31
3	O <i>Semantic Web Challenge</i> (SWC)	32
3.1.	Requisitos e Qualidades Desejáveis das Aplicações	33
3.2.	Classificação das Aplicações	34
3.3.	Descrevendo as Aplicações	35
3.3.1.	A Força Tarefa <i>Applications and Demos Task Force</i> no Grupo de Trabalho <i>Semantic Web Best Practices and Deployment</i> do W3C	35
3.3.2.	DOAP: <i>Description Of A Project</i>	36
3.3.3.	Uma Extensão do DOAP <i>Vocabulary</i>	40
3.4.	Resumo	46
4	Análise de Domínio das Aplicações do SWC	48
4.1.	<i>Browse Functionality</i>	50
4.1.1.	<i>Generation of Navigational Views Functionality</i>	51
4.1.2.	<i>Dynamic and Semantic Linking Hypertext Structures Functionality</i>	52
4.2.	<i>Search Functionality</i>	52
4.2.1.	<i>Semantic Search Functionality</i>	53
4.2.2.	<i>Semantic Query Expansion Functionality</i>	54

4.3. <i>Access through Diverse Devices Functionality</i>	54
4.4. <i>Support for Diverse Languages Functionality</i>	55
4.5. <i>Use of Multimedia Documents Functionality</i>	55
4.5.1. <i>Multimedia Handling Functionality</i>	55
4.5.2. <i>Multimedia Metadata Functionality</i>	56
4.5.3. <i>Multimedia Generation Functionality</i>	56
4.6. <i>Semantic Growth Functionality</i>	57
4.7. <i>Semantic Recommender Policy Functionality</i>	57
4.8. <i>Ontology Functionality</i>	58
4.8.1. <i>Ontology Schema Editor Functionality</i>	58
4.8.2. <i>Ontology Instances Editor Functionality</i>	58
4.8.3. <i>Ontology Repository Functionality</i>	59
4.9. <i>Tipos de Aplicações</i>	59
4.9.1. <i>Portal</i>	60
4.9.2. <i>Ontology Tool</i>	61
4.9.3. <i>Instance of a Framework</i>	61
4.9.4. <i>Semantic P2P Application</i>	62
4.9.5. <i>Semantic Collaborative Tool</i>	63
4.9.5.1. <i>Semantic Wiki</i>	63
4.10. <i>Tipos de Integração de Dados</i>	64
4.10.1. <i>Wrappers and Mediators Integration Functionality</i>	64
4.10.2. <i>Manual Integration Functionality</i>	65
4.11. <i>Resumo</i>	66
5 <i>Um Framework de Aplicações para a Web Semântica</i> (<i>SWAPpFW - Semantic Web Application Framework</i>)	67
5.1. <i>Requisitos</i>	67
5.1.1. <i>Os requisitos do SWC</i>	68
5.1.2. <i>O Processo de Manipulação de Metadados</i> (<i>Metadata Handling Process</i>)	70
5.2. <i>Arquitetura do SWAPpFW</i>	72
5.2.1. <i>O “4+1” View Model of Software Architecture</i>	72
5.2.2. <i>O Ajuste do “4+1” View Model</i>	74
5.3. <i>As Visões do SWAPpFW</i>	75
5.3.1. <i>Os Cenários (Scenarios)</i>	75
5.3.2. <i>A Visão de Análise (Analysis View)</i>	78

5.3.3. A Visão de Desenvolvimento (<i>Development View</i>)	78
5.3.4. A Visão de Processo (<i>Process View</i>)	83
5.3.5. A Visão Física (<i>Physical View</i>)	84
5.4. O <i>Design</i> do SWAPpFW	85
5.4.1. Possíveis Configurações de <i>Design</i> (“Complexidade”)	85
5.4.2. Uma Combinação Válida de Funcionalidades	92
5.5. Resumo	93
6 Conclusões, Contribuições e Trabalhos Futuros	94
6.1. Contribuições	95
6.2. Trabalhos Futuros	95
7 Referências Bibliográficas	97
Apêndice A - O SWDOAP <i>Vocabulary</i>	105
Anexo I – O DOAP <i>Vocabulary</i>	113
Anexo II – Glossário	130

Siglas e Abreviaturas

ACM.....	Association for Computing Machinery
ADTF	Applications and Demos Task Force
API.....	Application Programming Interface
BUSTER.....	Bremen University Semantic Translator for Enhanced Retrieval
CAS	CS AKTive Space
COHSE.....	Conceptual Open Hypermedia Service
CORESE.....	Conceptual Resource Search Engine
CS.....	Computer Science
CSS	Cascading Style Sheets
DAML+OIL.....	DARPA Agent Markup Language + Ontology Inference Layer
DARPA	USA's Defense Advanced Research Projects Agency
DDI.....	Data Documentation Initiative
DLS.....	Distributed Link Service
DOAP.....	Description Of A Project
DOPE.....	Drug Ontology Project for Elsevier
ECHO	Earth Observing System ClearingHouse
ELSST	European Language Social Science Thesaurus
FOAF	Friend Of A Friend RDF/XML vocabulary.
GAV	Global As View
GCMD.....	Global Change Master Directory
GO	Gene Ontology
GOHSE.....	GO + COHSE
GVV	Global Virtual View
HTML.....	Hypertext Markup Language
HTTP	HyperText Transfer Protocol
INRIA	Institut National de Recherche en Informatique et en Automatique
KmP.....	Knowledge Management Platform
KMS.....	Knowledge Management System
MADIERA	Multilingual Access to Data Infrastructures of the European Research Area
MDA.....	Model Driven Architecture
MLH	Multi-Layered Hypercube

MOMIS.....	<i>Mediator EnvirOnment for Multiple Information Sources</i>
MST	<i>A classification and code of disciplines GB/T 13745/92 by Ministry of Science and Technology, China</i>
NESSTAR.....	<i>Networked Social Science Tools and Resources</i>
OBA	<i>Ontology-Based Application</i>
ODL	<i>Object Definition Language</i>
ODL-I3.....	<i>A language that extends ODL with an underlying Description Logic</i>
ODMG.....	<i>Object Database Management Group</i>
OGC	<i>Open Geospatial Consortium</i>
OHS.....	<i>Open Hypermedia System</i>
OMG	<i>Object Group Management</i>
OMV.....	<i>Ontology Metadata Vocabulary</i>
OntoIQ	<i>Ontologent Interactive Query Tool</i>
OWL.....	<i>Web Ontology Language</i>
OWL DL.....	<i>A specie (DL) of Web Ontology Language (OWL)</i>
OWL Full.....	<i>A specie (Full) of Web Ontology Language (OWL)</i>
OWL Lite.....	<i>A specie (Lite) of Web Ontology Language (OWL)</i>
P2P	<i>Peer-to-peer</i>
PPR	<i>Personal Publication Reader</i>
PRF.....	<i>Personal Reader Framework</i>
RDF	<i>Resource Description Framework</i>
RDF Schema	<i>RDF Vocabulary Description Language 1.0</i>
RDQL.....	<i>RDF Data Query Language</i>
RSS	<i>RDF Site Summary (RSS 0.9 and 1.0)</i>
RSS	<i>Rich Site Summary (RSS 0.91, RSS 1.0)</i>
RSS	<i>Really Simple Syndication (RSS 2.0)</i>
RSS	<i>Real-time Simple Syndication (RSS 2.0)</i>
SEAL.....	<i>SEmantic portAL</i>
SECO.....	<i>SEmantic COllaboration</i>
SHOE.....	<i>Simple HTML Ontology Extensions</i>
SKOS.....	<i>Simple Knowledge Organisation System</i>
SPIA.....	<i>Semantic Portal of International Affairs</i>
SWAP	<i>Semantic Web Accessibility Platform</i>
SWApp.....	<i>Semantic Web Application</i>
SWAPSA	<i>Semantic Web and Peer-to-Peer Project – System Architecture.</i>

See Swapster.

Swapster..... *A generic platform to account for the general need of sharing semantic-based information in P2P fashion. See SWAPSA.*
SWBPD..... *Semantic Web Best Practices and Deployment Working group*
SWC *Semantic Web Challenge*
SWD *Semantic Web Documents*
SWEET *Semantic Web for Earth and Environmental Terminology*
SWWS'01 *Semantic Web Working Symposium (SWWS) 2001*
UML *Unified Modeling Language*
UNSO *UNSpecified Ontology*
URL..... *Uniform Resource Locator*
W3C..... *World Wide Web Consortium*
WCAG..... *Web Content Accessibility Guidelines*
WSML *Web Service Markup Language*
WSMO *Web Service Modeling Ontology*
WSMT..... *Web Service Modeling Toolkit*
WSMX..... *Web Service Execution Environment*
XML *eXtensible Markup Language*
XSLT..... *eXtensible Stylesheet Language Transformation*

Lista de Figuras

Figura 1 – Uma “interpretação” da lista de ferramentas ou tecnologias da <i>Web Semântica</i> de Fensel <i>et alli</i> (Fensel <i>et al.</i> , 2003)	18
Figura 2 – <i>An ontology spectrum</i> (McGuinness, 2003)	26
Figura 3 - <i>Semantic Continuum ...</i> (Uchold, 2003).....	26
Figura 4 – Decisões pragmáticas baseadas na <i>Semantic Web stack</i> para projetar e implementar uma SWAPp (Tummarello & Morbidoni, 2005).....	30
Figura 5 - Uma versão atualizada da <i>Semantic Web stack</i> (Berners-Lee, 2005).....	30
Figura 6 - O Processo de Manipulação de Metadados (<i>Metadata Handling Process</i>)	71
Figura 7 - <i>The "4+1" View Model</i> (Kruchten, 1995)	73
Figura 8 - O <i>"4+1" View Model</i> ajustado.....	74
Figura 9 - O Cenário Geral do SWAPpFW	75
Figura 10 - Cenário 0: Os Requisitos do SWC	76
Figura 11 - Cenário 1: Requisitos do SWC + Processo de Manipulação de Metadados.....	77
Figura 12 – O modelo de visão de análise do SWAPpFW	78
Figura 13 - Pacote <i>Application</i>	79
Figura 14 - Pacote <i>Functionality</i>	80
Figura 15 - Pacote <i>Integration</i>	81
Figura 16 - Pacote <i>Ontology</i>	82
Figura 17 - Dependências entre pacotes.....	83
Figura 18 - Aplicação cliente-servidor na <i>Web</i>	84
Figura 19 - Rede de aplicações P2P	84

Lista de Tabelas

Tabela 1 – Resumo do <i>Semantic Web Challenge</i>	34
Tabela 2 - Propriedades do DOAP	37
Tabela 3 - Propriedades do SWDOAP (DOAP Estendido).....	41
Tabela 4 – Resumo das Funcionalidades	49
Tabela 5 – Resumo dos Tipos de Aplicação	50
Tabela 6 – Resumo dos Tipos de Integração.....	50
Tabela 7 - Aplicações <i>Portal</i>	86
Tabela 8 - Aplicações <i>Ontology Tool</i>	89
Tabela 9 - Aplicações <i>Instance of a Framework</i>	91

O meu olhar é nítido como um girassol.
Tenho o costume de andar pelas estradas
Olhando para a direita e para a esquerda,
E de vez em quando olhando para trás...
E o que vejo a cada momento
É aquilo que nunca antes eu tinha visto,
E eu sei dar por isso muito bem...
Sei ter o pasmo comigo
Que tem uma criança se, ao nascer,
Reparasse que nascera deveras...
Sinto-me nascido a cada momento
Para a eterna novidade do mundo...

Creio no mundo como num malmequer,
Porque o vejo. Mas não penso nele
Porque pensar é não compreender...
O mundo não se fez para pensarmos nele
(Pensar é estar doente dos olhos)
Mas para olharmos para ele e estarmos de acordo.

Eu não tenho filosofia: tenho sentidos...
Se falo na Natureza não é porque saiba o que ela é,
Mas porque a amo, e amo-a por isso
Porque quem ama nunca sabe o que ama
Nem sabe por que ama, nem o que é amar...

Amar é a eterna inocência,
E a única inocência é não pensar...