

Daniela Francisco Brauner

**Uma Arquitetura para Catálogos de Objetos
baseados em Ontologias**

Dissertação de Mestrado

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

Orientadores: Prof. Carlos José Pereira de Lucena
Prof. Marco Antonio Casanova

Rio de Janeiro
Abril de 2005



Daniela Francisco Brauner

**Uma Arquitetura para Catálogos de Objetos
baseados em Ontologias**

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

Prof. Carlos José Pereira de Lucena

Orientador

Departamento de Informática - PUC-Rio

Prof. Marco Antonio Casanova

Co-orientador

Departamento de Informática - PUC-Rio

Prof^a. Simone Diniz Junqueira Barbosa

Departamento de Informática - PUC-Rio

Prof^a. Viviane Torres da Silva

Departamento de Informática - PUC-Rio

Prof. José Eugenio Leal

Coordenador Setorial do Centro

Técnico Científico - PUC-Rio

Rio de Janeiro, 14 de abril de 2005

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

Daniela Francisco Brauner

Graduou-se em Ciência da Computação pela UFPel (Universidade Federal de Pelotas) em março de 2003. Atualmente é doutoranda em Informática pela PUC-Rio.

Ficha Catalográfica

Brauner, Daniela Francisco

Uma arquitetura para catálogos de objetos baseados em ontologias / Daniela Francisco Brauner ; orientadores: Carlos José Pereira de Lucena, Marco Antonio Casanova. – Rio de Janeiro : PUC, Departamento de Informática, 2005.

141 f. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática.

Inclui referências bibliográficas.

1. Informática – Teses. 2. Catálogo de objetos. 3. Ontologia. 4. Interoperabilidade. 5. Federação. I. Lucena, Carlos José Pereira de. II. Casanova, Marco Antonio. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. IV. Título.

CDD: 004

Agradecimentos

Em primeiro lugar, quero agradecer aos meus orientadores pela confiança que depositaram em mim, pelas numerosas críticas, correções e sugestões que fizeram ao longo do trabalho e principalmente pelo apoio constante ao longo de meu aprendizado.

Um agradecimento especial ao Prof. Casanova pela orientação. Além de longas e produtivas conversas, merecem minha sincera gratidão sua capacidade de ensinar, seu rigor, exigência e competência, bem como seus conselhos amigos.

Obrigada a todos os colegas do Laboratorio de Engenharia de Software, onde encontrei um espaço privilegiado de criatividade, dinamismo e amizade.

Agradeço à CAPES e ao CNPq pelo financiamento desta pesquisa.

Agradeço ao Departamento de Informática e à PUC-Rio pelo apoio.

Aos meus pais, por me darem motivação para vencer os desafios.

Ao meu amor, Alexandre, por todo carinho e principalmente, pela companhia imprescindível e o apoio incondicional.

Finalmente, agradeço à Deus!

Resumo

Brauner, Daniela Francisco; Lucena, Carlos José Pereira de; Casanova, Marco Antonio. **Uma arquitetura para catálogos de objetos baseados em ontologias**. Rio de Janeiro, 2005. 141p. Dissertação de Mestrado - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

O compartilhamento de dados é um fator crítico para o sucesso das organizações devido à emergente necessidade de comunicação com seus parceiros de negócios. Porém, a heterogeneidade das fontes de dados é uma das barreiras para a efetiva interoperabilidade entre sistemas de informação. Esta dissertação inicialmente introduz o conceito de Catálogo de Objetos baseado em Ontologia (OnOC) como uma estratégia para endereçar o problema de interoperabilidade entre diferentes fontes de dados. Um OnOC oferece recursos para definir e manter um vocabulário comum, chamado ontologia de referência, que facilita o intercâmbio de dados entre diferentes fontes. Em seguida, propõe uma arquitetura para OnOCs que combina conceitos tais como catálogos de metadados, servidores de ontologias, mediadores e federações de bancos de dados.

Palavras-chave

Catálogo de objetos, ontologia, federação, fontes de objetos, heterogeneidade, interoperabilidade.

Abstract

Brauner, Daniela Francisco; Lucena, Carlos José Pereira de; Casanova, Marco Antonio. **An architecture to ontology based object catalog**. Rio de Janeiro, 2005. 141p. MSc. Dissertation - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Data sharing is a critical factor for the success of organizations due to the increasing need to facilitate communication with business partners. However, data source heterogeneity proved to be a barrier for the effective interoperability between information systems. The concept of an Ontology-based Object Catalog (OnOC) is first introduced as a strategy to address the interoperability problem between heterogeneous data sources. An OnOC offers support for the definition and maintenance of a common vocabulary, called a reference ontology, that facilitates data interchange among different sources. Then, an architecture for OnOCs is proposed that combines concepts such as metadata catalogs, ontology servers, mediators and database federations.

Keywords

Federation, heterogeneity, interoperability, objects catalog, object sources, ontology.

Sumário

1 Introdução	15
1.1 Motivação	15
1.2 Solução proposta	16
1.3 Organização da dissertação	16
2 Preliminares	18
2.1 Motivação	18
2.2 Dimensões para classificação de Sistemas de Informação	19
2.2.1 Distribuição	20
2.2.2 Heterogeneidade	20
2.2.3 Autonomia	22
2.2.4 Flexibilidade	23
2.3 Abordagens tradicionais para integração	23
2.3.1 Arquitetura para múltiplos bancos de dados	24
2.3.2 Bancos de dados federados	25
2.3.3 Arquitetura de mediadores	27
2.3.4 Armazém de dados	30
2.4 Abordagens inovadoras para integração	31
2.4.1 Arquitetura baseada em serviços	31
2.4.2 Abordagens baseadas no uso de ontologias	32
2.5 Servidores de ontologias e catálogos de metadados	35
2.5.1 Servidores e serviços de ontologias	35
2.5.2 Serviço de catálogo de metadados no Grid	38
2.5.3 Catálogo de feições do OpenGIS	38
2.6 <i>Dataweb</i> : infra-estrutura para compartilhamento de dados na Web	40
3 Viabilizando interoperabilidade através de OnOCs	42
3.1 Motivação	42

3.2 Metodologia	43
3.2.1 Descrição	43
3.2.2 Criação das ontologias e mapeamentos	50
3.2.2.1 Composição básica das ontologias	50
3.2.2.2 Criação da ontologia de referência	54
3.2.2.3 Criação da ontologia local	57
3.2.2.4 Definição dos mapeamentos	65
3.2.3 Criação da federação	72
3.3 Catálogo de Objetos baseado em Ontologias (OnOC)	73
3.3.1 Conceito	73
3.3.2 Situações de uso	76
3.3.2.1 Catálogo de metadados	76
3.3.2.2 Catálogo de instâncias de referência	77
3.3.2.3 Mediador para acesso a objetos distribuídos	77
3.3.3 Requisitos	78
4 Visão geral da arquitetura para OnOCs	81
4.1 Motivação	81
4.2 Componentes da arquitetura	82
4.3 Inter-relacionamentos da arquitetura	85
4.4 Padrões de software utilizados	88
4.4.1 Padrão arquitetural	89
4.4.2 Padrões de projeto	90
4.4.2.1 Padrão <i>Facade</i>	92
4.4.2.2 Padrão <i>Singleton</i>	92
4.4.2.3 Padrão <i>Adapter</i>	93
4.5 Serviços da arquitetura	94
5 Detalhamento da arquitetura para OnOCs	95
5.1 Motivação	95
5.2 Casos de uso	95
5.3 Diagramas de pacotes	101

5.4 Diagramas de classes	102
5.5 Diagramas de seqüência	108
6 OntCatalog: Um Protótipo de um OnOC	118
6.1 Motivação	118
6.2 Descrição	118
6.3 Testes	120
6.3.1 Critérios de teste utilizados	120
6.3.2 Descrição dos casos teste	121
6.3.2.1 Testes unitários	121
6.3.2.2 Testes funcionais	122
7 Conclusão e trabalhos futuros	127
8 Referências bibliográficas	128
Anexo A – Dicionário de dados do OntCatalog	134
Anexo B – Ontologias utilizadas para teste do OntCatalog	135
Anexo C – Relatório de testes do OntCatalog	140

Lista de figuras

Figura 1 – Caracterização das arquiteturas de bancos de dados.	19
Figura 2 – Arquitetura de cinco níveis de esquemas para FDBS.	27
Figura 3 – Arquitetura mediador-adaptador.	28
Figura 4 – Hierarquia de mediadores.	30
Figura 5 – Arquitetura do modelo de referência.	39
Figura 6 – Exemplo de heterogeneidade dos metadados.	44
Figura 7 – Mapeamento dos metadados.	45
Figura 8 – Exemplo de heterogeneidade na identificação de objetos.	46
Figura 9 – Armazenamento explícito do mapeamento entre objetos equivalentes.	46
Figura 10 – Representação abstrata das ontologias e mapeamentos entre classes.	48
Figura 11 – Representação abstrata das instâncias e mapeamentos.	49
Figura 12 – Repositório de ontologias.	49
Figura 13 – Tag delimitadora de uma ontologia em OWL..	51
Figura 14 – Exemplo de declaração dos namespaces de uma ontologia em OWL.	51
Figura 15 – Exemplo de definição dos metadados do documento OWL.	54
Figura 16 – Exemplo de definição dos metadados da ontologia de referência.	56
Figura 17 – Exemplo de banco de dados relacional.	58
Figura 18 – Exemplo de descrição de uma classe em OWL.	58
Figura 19 – Exemplo de representação em OWL de uma tabela e seus atributos.	61
Figura 20 – Exemplo de identificação usando o identificador original do objeto.	62
Figura 21 – Exemplo de identificação concatenando o nome da tabela ao identificador original do objeto.	63
Figura 22 – Exemplo de identificação utilizando uma estrutura de diretórios.	64

Figura 23 – Exemplo de descrição de igualdade entre instâncias em OWL.	66
Figura 24 – Exemplo de descrição de equivalência entre classes em OWL.	67
Figura 25 – Exemplo de descrição de uma subclasse em OWL.	67
Figura 26 – Exemplo de regra em SWRL para restrição.	71
Figura 27 – Exemplo de regra em SWRL para mapeamento.	72
Figura 28 – Diagrama de interação entre o OnOC, o cliente e as fontes de dados.	74
Figura 29 – Arquitetura em módulos para OnOCs.	82
Figura 30 – Subdivisão do módulo <i>CatManager</i> .	82
Figura 31 – Organização do módulo <i>CatRegistry</i> .	83
Figura 32 – Subdivisão do módulo <i>OntManager</i> .	83
Figura 33 – Subdivisão do módulo <i>OntRepository</i> .	84
Figura 34 – Arquitetura detalhada para OnOCs.	85
Figura 35 – Primeiro nível de inter-relacionamentos entre os módulos da arquitetura.	85
Figura 36 – Detalhamento dos inter-relacionamentos da arquitetura – Descreve termo.	86
Figura 37 – Detalhamento dos inter-relacionamentos da arquitetura - Consulta.	87
Figura 38 – Arquitetura usando o padrão em camadas.	90
Figura 39 – Uso do padrão de projeto <i>Facade</i> .	92
Figura 40 – Estrutura do padrão <i>Adapter</i> (Gamma et al., 1995).	93
Figura 41 – Atores do OnOC.	95
Figura 42 – Casos de uso do ator “Aplicação Cliente”.	96
Figura 43 – Detalhamento dos casos de uso 1.3.	96
Figura 44 – Casos de uso do ator “Aplicação Participante”.	97
Figura 45 – Casos de uso do ator “Usuário Administrador”.	97
Figura 46 – Pacote <i>OnOC</i> .	101
Figura 47 – Diagrama de pacotes do protótipo implementado.	102
Figura 48 – Diagrama de classes do pacote <i>OnOCManager</i> .	103
Figura 49 – Diagrama de classes do pacote <i>CatManager</i> .	106

Figura 50 – Diagrama de classes do pacote <i>CatRegistry</i> .	107
Figura 51 – Diagrama de classes do pacote <i>OntManager</i> .	107
Figura 52 – Diagrama de classes do pacote <i>OntRepository</i> .	108
Figura 53 – Diagrama de seqüência da operação <i>getOntology</i> (ontologia de referência).	109
Figura 54 – Diagrama de seqüência da operação <i>getOntology</i> (ontologias locais).	110
Figura 55 – Diagrama de seqüência da operação <i>getTermDescription</i> .	111
Figura 56 – Diagrama de seqüência da operação <i>registerNewApp</i> .	112
Figura 57 – Diagrama de seqüência da operação <i>registerObjectSrc</i> .	113
Figura 58 – Diagrama de seqüência da operação <i>setAppPermission</i> .	114
Figura 59 – Diagrama de seqüência da operação <i>setReferenceOntology</i> .	115
Figura 60 – Diagrama de seqüência da operação <i>setAdmin</i> .	116
Figura 61 – Diagrama de seqüência da operação <i>query</i> .	117
Figura 62 – Pacote <i>OntCatalog</i> .	119
Figura 63 – Diagrama de classes do pacote <i>OntRepository</i> .	120

Lista de tabelas

Tabela 1 – Exemplo de heterogeneidade esquemática.	22
Tabela 2 – Web versus <i>Dataweb</i> .	41
Tabela 3 – Namespaces freqüentemente utilizados em ontologias.	53
Tabela 4 – Mapeamento entre os tipos de dados do SQL para o XML <i>Schema</i> .	59
Tabela 5 – Implicações das propriedades <i>rdfs:subClassOf</i> e <i>owl:equivalentClass</i> .	68
Tabela 6 – Classificação dos padrões de projeto (Gamma et al., 1995).	91
Tabela 7 – Serviços oferecidos pela camada de interface.	94
Tabela 8 – Operações básicas de um OnOC.	103
Tabela 9 – Parâmetros da operação <i>getOntology</i> .	104
Tabela 10 – Parâmetros da operação <i>getTermDescription</i> .	104
Tabela 11 – Parâmetros da operação <i>query</i> .	104
Tabela 12 – Parâmetros da operação <i>registerObjectSrc</i> .	105
Tabela 13 – Parâmetros da operação <i>registerNewApp</i> .	105
Tabela 14 – Parâmetros da operação <i>setAppPermission</i> .	105
Tabela 15 – Parâmetros da operação <i>setReferenceOntology</i> .	106
Tabela 16 – Parâmetros da operação <i>setAdmin</i> .	106

Lista de siglas

ACL	Agent Communication Language
API	Application Program Interface
CORBA	Common Object Request Broker Architecture
DAML+OIL	DARPA Agent Markup Language with Ontology Inference Layer
DW	Data Warehouse
FDBS	Federated DataBase System
FIPA	Foundation for Intelligent Physical Agents
HTML	HyperText Markup Language
JDBC	Java Database Connectivity
MCS	Metadata Catalog Service
ODBC	Open DataBase Connectivity
OGC	OpenGIS Consortium
OnOC	Ontology based Object Catalog
OWL	Web Ontology Language
REST	Representational State Transfer
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
RDQL	RDF Data Query Language
SGBD	Sistema Gerenciador de Bancos de Dados
SOA	Service Oriented Architecture
SOAP	Simple Object Access Protocol
SWRL	Semantic Web Rule Language
UDDI	Universal Description Discovery and Integration
URI	Unified Resource Identifier
URN	Unified Resource Name
XML	eXtensible Markup Language
WSDL	Web Services Description Language
W3C	World Wide Web Consortium