

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO



Simone Leal de Moura

**Uma arquitetura para Integração de Repositórios de
Objetos de Aprendizagem baseada em Mediadores e
Serviços *Web***

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. Rubens Nascimento Melo
Prof. Sean Wolfgang Matsui Siqueira

Rio de Janeiro
Agosto de 2005



Simone Leal de Moura

**Uma arquitetura para Integração de Repositórios de
Objetos de Aprendizagem baseada em Mediadores e
Serviços Web**

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. Rubens Nascimento Melo

Orientador
PUC-Rio

Prof. Sean Wolfgang Matsui Siqueira

Co-Orientador
PUC-Rio

Prof. Antônio Luz Furtado

PUC-Rio

Prof. Álvaro César Pereira Barbosa

UFES

Prof^a. Maria Helena Lima Baptista Braz

IST/UTL

Prof. José Eugenio Leal

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

Rio de Janeiro, 19 de agosto de 2005

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

Simone Leal de Moura

Graduou-se em Tecnologia em Processamento de Dados nas Faculdades Integradas Simonsen em 1993. Especializou-se em Análise, Projeto e Gerência de Sistemas no curso de Pós-Graduação Lato-Sensu da PUC-Rio em 1997. Tornou-se monitora e posteriormente professora de disciplinas de Banco de Dados em cursos de extensão e especialização da PUC-Rio. Administradora do ambiente de Banco de Dados do TecBD/DI/PUC-Rio. Desde 2000 vem trabalhando no projeto, construção e manutenção dos Bancos de Dados que constituem o programa Delegacia Legal.

Ficha Catalográfica

Moura, Simone Leal de

Uma arquitetura para integração de repositórios de objetos de aprendizagem baseada em mediadores e serviços Web / Simone Leal de Moura ; orientadores: Rubens Nascimento Melo, Sean Wolfgang Matsui Siqueira. – Rio de Janeiro : PUC-Rio, Departamento de Informática, 2005.

158 f. ; 30 cm

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

Inclui bibliografia.

1. Informática – Teses. 2. Integração de dados. 3. Metadados. 4. Mediador. 5. Serviços Web. 6. Ontologia. 7. Objetos de aprendizagem. 8. PGL. I. Melo, Rubens Nascimento. II. Siqueira, Sean Wolfgang Matsui. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. IV. Título.

CDD: 004

Aos meus pais, Moura e Mariete
Ao Paulinho

Agradecimentos

Aproveito esse momento para expressar meus sinceros agradecimentos a tantos que me acompanharam e fizeram parte de um dos capítulos da minha vida nessa jornada. Agradeço imensamente.

À Deus, pela força, coragem e perseverança por ter chegado até aqui;

À PUC, instituição que me favoreceu com uma bolsa de isenção garantindo a realização desta dissertação de mestrado;

Aos meus pais, pelo carinho e compreensão nas ausências familiares, principalmente nos fins de semana, onde minha dedicação era exclusivamente à PUC, e pelas palavras tranquilizadoras quando me encontravam em total angústia;

Ao meu orientador, Prof.Dr.Rubens Nascimento Melo, por me direcionar com seu tema de pesquisa, pelo estímulo e tranquilidade com suas orientações e, principalmente, por sua amizade e pela alegria de trabalharmos juntos;

Ao Sean, um grande amigo e co-orientador que me despertou desde o início a forma organizada e responsável de se trabalhar no mestrado, desenvolvendo a dissertação, assim como as publicações de artigos (principalmente nas traduções para o inglês) e por todas as sugestões e críticas que foram de total importância para a conclusão deste trabalho;

Ao Fábio Coutinho, que também foi um grande amigo e co-orientador e deu uma enorme ajuda no processo de produção da dissertação principalmente na fase final. Valorizei bastante os comentários e críticas que recebi sobre o texto, mesmo quando divergíamos. Aproveito para me desculpar por todos os incômodos que lhe causei tirando-o de seu trabalho nos finais de semana para discutir os capítulos da dissertação. Nunca comi tanta pizza Domino's no TecBD, nesses últimos seis meses;

A grande amiga Diva (mãezona), pelos tantos “puxões de orelha”, conselhos e dicas de como levar bem o mestrado, de nossos bate papos e cafezinhos no TecBD antes de iniciarmos nossos estudos;

A Geórgia, uma amiga de fé, que mesmo estando somente às segunda-feira no TecBD, causava um enorme movimento. Apesar de toda a nossa bagunça, tínhamos momentos de seriedade em nossas pesquisas, mas no final da tarde não podia faltar o nosso hamburger quentíssimo no bar das meninas, do lado de fora da PUC;

A amiga Julita, nossa mascote do TecBD, que sempre com um sorriso no rosto nos ajudou com tarefas administrativas quando um artigo era aceito e também, pelas correrias nas traduções de alguns trabalhos realizados, além da sua amizade e companheirismo do dia a dia no TecBD, sempre, é claro, com muita animação;

Ao grande Luis Pereira, pela colaboração e amizade, acompanhados de suas piadas codimentadas que nos extraíam gargalhadas, mesmo em momentos de muita tensão;

Ao Portelinha, um amigo que mesmo recém chegado ao TecBD, estava sempre disposto a ajudar.

Finalmente, agradeço ao Paulinho, por sua enorme paciência, companheirismo e compreensão nos momentos mais difíceis que passei durante todo o mestrado. Sou grata por todas as liberações no trabalho, o que me favoreceu em termos de tempo, para uma maior dedicação ao trabalho na PUC.

Obrigada à todos vocês, meus inesquecíveis amigos TecBDianos.

Resumo

Moura, Simone. **Uma Arquitetura para Integração de Repositórios de Objetos de Aprendizagem baseada em Mediadores e Serviços Web**. PUC-Rio, 2005. 158p. Dissertação de Mestrado - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Na educação baseada na *Web* há uma grande preocupação em relação ao compartilhamento de materiais instrucionais devido à complexidade do processo de desenvolvimento de materiais com boa qualidade. Isto leva a uma tendência em adotar a abordagem de orientação a objetos no desenvolvimento destes materiais, originando os chamados "Objetos de Aprendizagem". Além disto, instituições e empresas interessadas na educação baseada na *Web* vêm formando parcerias no sentido de promover o compartilhamento de tais objetos. De modo a contribuir para estes esforços desenvolvemos uma arquitetura de mediadores e tradutores enriquecidos pelo uso de ontologias, que é implementada como serviços *Web* de modo a permitir a autonomia local com uma visão integrada. O uso de mediadores possibilita que uma consulta seja redefinida em sub-consultas que são distribuídas às fontes de dados e o resultado integrado. Os tradutores permitem que as sub-consultas sejam entendidas pelas fontes de dados e as respectivas respostas sejam entendidas pelo mediador. A implementação dos componentes da arquitetura como serviços *Web* possibilita uma maior flexibilidade e interoperabilidade entre os participantes da rede. O tratamento da heterogeneidade semântica faz uso do formalismo de ontologias para descrever os conceitos existentes nos metadados de cada repositório participante na rede e encontrar as possíveis equivalências entre eles. Desta forma, o desenvolvimento desta arquitetura resultou em LORIS, um sistema de integração de repositórios de objetos de aprendizagem. O LORIS está sendo aplicado no PGL, um projeto de cooperação internacional entre instituições de ensino e pesquisa para promover a educação baseada na *Web*.

Palavras-chave

Integração de Dados; Metadados; Mediador; Serviços *Web*; Ontologia; Objetos de Aprendizagem; PGL

Abstract

Moura, Simone. **An Architecture based on Mediators and Web Services for Integrating Learning Objects Repositories**. PUC-Rio, 2005. 158p. Master Dissertation - Computer Science Department, Pontifical Catholic University of Rio de Janeiro.

In web-based education there is an emphasis on reusing and sharing instructional content due to the complexity of the development process of high-quality learning materials. It leads to the learning objects orientation as well as to partnerships among institutions to promote sharing of these objects. In order to contribute to these efforts, we proposed an architecture based on the mediators and wrappers for integrating learning objects repositories. The components of this architecture were implemented by the use of web-services and the integration processes were enriched by ontologies. The use of mediators allows a query to be redefined as sub-queries that are distributed to the data sources and the results to be integrated. The wrappers allow the data sources to understand the sub-queries and the mediator to understand the respective answers. The implementation of the architecture components as web services allows more flexibility and interoperability among the participants of the community. The formalism of ontologies is used to deal with the semantic heterogeneity as the metadata concepts of each repository are described and the equivalences are established. The development of this architecture is called LORIS, an integration system of learning objects' repositories. LORIS is being adopted by PGL, an international partnership project for promoting web-based education.

Keywords

Data Integration; Metadata; Mediator; Web Services; Ontology; Learning Objects; PGL

Sumário

Abreviaturas e Siglas	16
1 Introdução	17
1.1. Descrição do Contexto	17
1.2. Motivação	18
1.3. Objetivos da Dissertação	19
1.4. Organização da Dissertação	20
2 Fundamentação	22
2.1. Objetos de Aprendizagem	22
2.2. Modelos de Representação de Metadados	24
2.2.1. Dublin Core	24
2.2.2. Ariadne	26
2.2.3. IEEE-LOM	27
2.3. Integração de Dados	30
2.3.1. Sistema de Informação Federado	31
2.3.1.1. Propriedades das Fontes de Dados em um Processo de Integração de Dados	31
2.3.1.1.1. Autonomia	31
2.3.1.1.2. Heterogeneidade	32
2.3.1.1.3. Distribuição	34
2.3.1.2. Tipos de SIF	34
2.3.1.2.1. Banco de Dados Distribuído	34
2.3.1.2.2. Sistema de Informação Fracamente Acoplado	35
2.3.1.2.3. Sistema de Informação Fortemente Acoplado	36
2.3.1.2.4. Sistema de Banco de Dados Federado	36
2.3.1.2.5. Sistema de Informação Baseado em Mediadores	38
2.3.2. Estratégias de Desenvolvimento	42
2.4. Serviços Web	44
2.4.1. Arquitetura Básica de Serviços Web	45

2.4.1.1. Componentes	45
2.4.1.2. Entidades	46
2.4.1.3. Operações	46
2.4.2. Tecnologias de Serviços Web	47
3 A Abordagem Adotada e Trabalhos Relacionados	49
3.1. Alternativas para Integração de LOs	49
3.1.1. Centralizada	49
3.1.2. Distribuída com Replicação	50
3.1.3. Par-a-Par	51
3.1.4. Distribuída Baseada em Mediadores e Tradutores	52
3.2. Trabalhos de Integração	53
3.2.1. Projeto Integra	53
3.2.2. Projeto Edutella	54
3.2.3. Projeto ONES	55
3.2.4. Projeto MIWEB	56
3.2.5. Projeto IBHIS	57
4 Tratamento de Heterogeneidade dos Metadados	59
4.1. Meta-esquemas e Padrões de Metadados para LOs	59
4.2. Mapeamento entre Modelos	60
4.2.1. Mapeamento Semântico	61
4.2.2. Mapeamento Estrutural	62
4.2.2.1. Mapeamento de Elementos	62
4.2.2.2. Mapeamento de Valores de Elementos	63
4.3. Mapeamento no LORIS	63
4.3.1. Representação Estrutural	64
4.3.2. Representação Semântica	65
4.3.2.1. Ontologias dos Padrões	68
4.3.2.2. Ontologia Comum	70
4.3.3. Processo de Mapeamento	72
4.4. Considerações Finais	75
5 Arquitetura Proposta	76

5.1. Visão Geral da Arquitetura	76
5.2. As Camadas da Arquitetura	78
5.2.1. Camada de Aplicação	78
5.2.2. Camada de Mediação	78
5.2.2.1. Serviço de Registro das Fontes de Dados	82
5.2.2.2. Serviço de Gerência de Informações	82
5.2.2.3. Serviço de Mediação	82
5.2.2.4. Serviço Mapeador	83
5.2.2.5. Serviço de Refinamento	83
5.2.3. Camada de Acesso aos Dados	84
5.2.3.1. Serviço de Tradução	84
5.3. Diagrama de Casos de Uso	85
5.4. Considerações Finais	88
6 Desenvolvimento do Protótipo	89
6.1. Estudo de Caso	89
6.1.1. Partnership in Global Learning	89
6.1.2. Ambiente PGL	90
6.2. Análise do Sistema LORIS	91
6.2.1. Subsistemas do LORIS	91
6.2.1.1. Subsistema de Mediação	91
6.2.1.2. Subsistema de Tradução	94
6.2.1.3. Subsistema GUI	95
6.2.1.4. Módulo Utilities	96
6.3. Implementação do Protótipo	97
6.3.1. Arquitetura do Protótipo	97
6.3.2. Linguagem de Consulta	98
6.3.3. Implementação do Subsistema de Mediação	101
6.3.4. Implementação do Subsistema de Tradução	102
6.3.4.1. Sítio A	102
6.3.4.2. Sítio B	102
6.3.4.3. Sítio C	102
6.3.5. Implementação do Subsistema GUI	103
6.3.5.1. Interface de Busca Simples	103

6.3.5.2. Interface de Busca Avançada	104
6.4. Considerações Finais	106
7 Conclusão	107
7.1. Contribuições Resultados	108
7.2. Considerações envolvendo os Trabalhos Relacionados	109
7.3. Trabalhos Futuros	110
8 Referência Bibliográfica	112
Anexo A – Tecnologias de Serviços <i>Web</i>	119
A.1 WSDL - Web Service Description Language	119
A.2 SOAP – Simple Object Access Protocol	123
A.3 UDDI - Universal Description, Discovery and Integration	127
Anexo B –Trabalhos de Integração de Dados baseado em Mediadores	130
B.1 TSIMMIS	130
B.2 SIMS	131
B.3 GARLIC	132
B.4 MIX	132
B.5 HERMES	133
B.6 IDIMS	134
B.7 MBM	134
Apêndice A – Ontologia Comum em OWL	135
Apêndice B – Esquema do documento XML que representa a consulta	158

Lista de Figuras

Figura 1.1: Ambiente de uma comunidade de <i>e-learning</i> .	19
Figura 2.1: Arquiteturas diferentes para integração de dados	34
Figura 2.2: Arquitetura em cinco camadas do SBDF	37
Figura 2.3: Arquitetura de um SIBM	39
Figura 2.4: Abordagem GaV (a) versus LaV (b).	41
Figura 2.5: Projeto de um SIF (a) <i>Bottom-up</i> (b) <i>Top-down</i>	43
Figura 2.6: Arquitetura básica orientada a serviço	45
Figura 4.1: Representação estrutural dos padrões	65
Figura 4.2: Representação semântica e estrutural dos padrões	67
Figura 4.3: Visão completa das ontologias Ariadne, LOM e DC	69
Figura 4.4: Visão semântica de <i>autor</i> e <i>requisito</i> na ontologia comum	71
Figura 4.5: Visão completa da ontologia comum	72
Figura 4.6: Exemplo de Mapeamento em LORIS	73
Figura 4.7: Implementação do mapeamento do conceito <i>author</i>	74
Figura 4.8: Representação em OWL do mapeamento do <i>author</i>	75
Figura 5.1: Uma Visão Geral da Arquitetura LORIS	77
Figura 5.2: Camada de Mediação	81
Figura 5.3: Serviços do Mediador Central	82
Figura 5.4: Serviços de Tradução	85
Figura 5.5: Diagrama de Casos de Uso	86
Figura 6.1: Configuração do ambiente.	90
Figura 6.2: Diagrama de classes da mediação principal	92
Figura 6.3: Diagrama de Classes do módulo de sub-mediação	94
Figura 6.4: Diagrama de classes da tradução das fontes.	95
Figura 6.5: Diagrama de classes do subsistema GUI	96
Figura 6.6: Diagrama de classes do pacote <i>utilities</i> .	97
Figura 6.7: Tecnologias utilizadas no desenvolvimento do LORIS	98
Figura 6.8: Documento XML representando a consulta comum.	99
Figura 6.9: Consulta mapeada para as fontes	100
Figura 6.7: Tela da busca simples	104
Figura 6.8: Tela de busca avançada	105

Figura 6.9: Resultados da busca avançada.	105
Figura A.1: Representação de uma mensagem RPC	119
Figura A.2: Representação de um código WSDL	122
Figura A.3: Estrutura de uma mensagem SOAP	124
Figura A.4: Exemplo de um <i>Envelope</i> (XML)	124
Figura A.5: Mensagem SOAP usando chamada RPC	126
Figura A.6: Exemplos de especificações do campo <i>SOAPAction</i> .	127
Figura A.7: Estrutura de dados UDDI	128
Figura B.1: Arquitetura TSIMMIS	130
Figura B.2: Arquitetura GARLIC	132
Figura B.3: Arquitetura MIX	133
Figura C.1: Representação de um XML Schema	158

Lista de Tabelas

Tabela 2.1: Tipos de Heterogeneidade	32
Tabela 2.2.: Nomes de atributo	33
Tabela 2.3: Valores de atributo	33
Tabela 2.4: Tipos de Sistemas de Informação Federados	43
Tabela 2.5: Padrões Fundamentais dos Serviços <i>Web</i>	48
Tabela A.1: Definições de elementos de um documento WSDL	120
Tabela A.8: Tabela de definições de elementos do registro UDDI	128

Abreviaturas e Siglas

ADL	Iniciativa da Secretaria de Defesa dos EUA no sentido de estabelecer um ambiente distribuído de aprendizagem - <i>Advanced Distributed Learning</i>
ARIADNE	<i>Alliance of Remote Instructional Authoring and Distribution Networks for Europe</i>
BD	Banco de Dados
DC	Dublin Core
GaV	<i>Global as View</i>
IEEE	Institute of Electrical and Electronics Engineers
LaV	<i>Local as View</i>
LMS	Sistemas de Gerência da Aprendizagem - <i>Learning Management Systems</i>
LOS	Objetos de Aprendizagem - <i>Learning Objects</i>
LOM	Padrão de metadados proposto pelo IEEE para descrição de LOs - <i>Learning Objects Metadata</i>
LORIS	Sistema de Integração de Repositórios de Objetos de Aprendizagem - <i>Learning Objects Repositories' Integration System</i>
LTSC	Comitê do IEEE responsável pela padronização da tecnologia de aprendizagem - <i>Learning Technology Standards Committee</i>
NSF	National Science Foundation
OWL	Linguagem para definição/especificação de ontologias para a Web - <i>Web Ontology Language</i>
P2P	Par-a-par ou <i>peer-to-peer</i>
PGL	Partnership in Global Learning
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
SBDF	Sistema de Banco de Dados Federado
SGBDR	Sistema de Gerência de Banco de Dados Relacional
SIBM	Sistema de Informação Baseado em Mediador
SIF	Sistema de Informação Federado
SOA	Service-Oriented Architecture
SOAP	Simple Object Access Protocol
TecBD	Laboratório de Tecnologia em Banco de Dados do Departamento de Informática da PUC-Rio
URI	Uniform Resource Identifier
XML	Extensible Markup Language
W3C	World Wide Web Consortium
Web	World Wide Web
WSDL	Web Service Description Language