



**Luiz André Portes Paes Leme**

**Uma arquitetura de software para catalogação automática  
de dados geográficos**

**Dissertação de Mestrado**

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

Orientador: Prof. Marco Antonio Casanova

Rio de Janeiro  
8 de agosto de 2006



**Luiz André Portes Paes Leme**

## **Uma arquitetura de software para catalogação automática de dados geográficos**

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. Marco Antonio Casanova**  
Orientador  
PUC-Rio

**Prof. Daniel Schwabe**  
PUC-Rio

**Prof<sup>a</sup>. Karin Koogan Breitman**  
PUC-Rio

**Prof. José Eugênio Leal**  
Coordenador Setorial do Centro Técnico Científico - PUC-Rio

Rio de Janeiro  
8 de agosto de 2006

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

### **Luiz André Portes Paes Leme**

Engenheiro Eletrônico graduado pela Universidade do Estado do Rio de Janeiro em julho de 1989. Coursou Pós-graduação "latu-sensu" em Gerência e Desenvolvimento de Sistemas Distribuídos no Núcleo de Computação Eletrônica da UFRJ de julho/2000 a jun./2001. Coordenou projetos de Sistemas de Informação para a Golden Cross de 1996 a 2004 nas áreas de análise de sinistros e atendimento a clientes.

#### Ficha Catalográfica

Leme, Luiz André Portes Paes

Uma arquitetura de software para catalogação automática de dados geográficos / Luiz André Portes Paes Leme ; orientador: Marco Antonio Casanova. – Rio de Janeiro : PUC-Rio, Departamento de Informática, 2006.

120 f. : il. ; 30 cm

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

Inclui bibliografia

1. Informática – Teses. 2. Dicionário geográfico. 3. Tesouro. 4. Alexandria digital library. 5. Catálogo. 6. Banco de dados. 7. Indexação. 8. Dado geográfico. 9. Metadado. 10. ISSO 19115. I. Casanova, Marco Antonio. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

## Agradecimentos

Ao meu orientador Professor Marco Antonio Casanova pelo grande estímulo e conhecimento técnico.

À minha mulher pela paciência e carinho com que me estimulou a superar os desafios.

Aos meus colegas da PUC-Rio, em especial à Daniela Brauner pelas inúmeras contribuições a este trabalho.

Ao CNPq e à PUC-Rio, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

Aos professores que participaram da Comissão examinadora.

A todos os professores e funcionários do Departamento pelos ensinamentos e pela ajuda.

A todos os amigos e familiares que de uma forma ou de outra me estimularam ou me ajudaram.

## Resumo

Paes Leme, Luiz André Portes. **Uma arquitetura de software para catalogação automática de dados geográficos**. Rio de Janeiro, 2006. 120p. Dissertação de Mestrado - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Dados geográficos estão disponíveis em quantidade e variedade crescentes à medida que evoluem as tecnologias de informática. Para torná-los úteis, é necessário que mecanismos de busca de dados possam identificar dados apropriados a determinado propósito. Tais mecanismos, comumente, utilizam catálogos de metadados que descrevem cada dado geográfico. Entretanto, a geração de metadados é um processo que pode consumir muito tempo e estar sujeito a muitos erros, caso seja feito manualmente. Essa dissertação apresenta uma arquitetura de software e tecnologias correlatas para aplicações de catalogação automática de dados geográficos.

## Palavras-chave

1. Informática – Teses; 2. dicionário geográfico; 3. tesouro; 4. Alexandria Digital Library; 5. catálogo; 6. banco de dados; 7. indexação; 8. dado geográfico; 9. metadado; 10. ISO 19115

## **Abstract**

Paes Leme, Luiz André Portes. **A software architecture for automated cataloguing of geographic data.** Rio de Janeiro, 2006. 120p. MSc. Dissertation - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

The amount and variety of geographic data increase as technology evolves. To make them useful it is necessary to implement search engines capable of identifying appropriate data. Such engines are usually based on metadata catalogs which describe the geographic data. However, the metadata generation process is time consuming and is not fail safe if it is carried out manually. This dissertation presents a software architecture, and related technologies, for the construction of automated cataloguing applications of geographic data.

## **Palavras-chave**

1. information science – thesis; 2. gazetteer; 3. thesaurus; 4. Alexandria Digital Library; 5. catalog; 6. database; 7. indexing; 8. geographic data; 9. metadata; 10. ISO 19115

# Sumário

1	Introdução	12
1.1.	Motivação	12
1.2.	Trabalhos relacionados	13
1.3.	Organização do trabalho	15
2	Dicionários e catálogos	17
2.1.	Padrões de metadados	19
2.1.1.	Padrão ISO 19115	19
2.1.2.	Padrão FGDC	23
2.2.	Interface de serviço de catálogos	26
2.2.1.	OpenGIS Services Framework	26
2.2.2.	Geospatial Web Services Architecture	28
2.2.3.	OGC Catalogue Service 2.0	29
2.3.	Dicionários geográficos	32
2.3.1.	Exemplos de dicionários geográficos	32
2.3.2.	Classificação de objetos geográficos	34
2.3.3.	Padrões para dicionários geográficos	35
2.3.4.	Alexandria Digital Library Gazetteer	38
3	Arquitetura de software para catalogação automática de dados geográficos	42
3.1.	Estratégia de catalogação	42
3.2.	Requisitos da arquitetura	46
3.3.	Processo de catalogação	50
3.4.	Geração de metadados	52
3.5.	Esquema de metadados	55
4	Implementação do GeoCatalog	58
4.1.	Casos de uso	59
4.2.	Padrão arquitetural da aplicação	67
4.3.	Interface de usuário	68
4.3.1.	Interação com usuário	70

4.3.2. Atualização das informações de interface de usuário	72
4.3.3. Configuração de um processo de catalogação	74
4.4. Inicialização de processo de catalogação	74
4.5. Geração de metadados	77
4.6. Framework de catalogação de dados	80
4.7. Métricas e organização do código	87
4.8. Exemplo de utilização	89
5 Conclusões e trabalhos futuros	94
6 GLOSSÁRIO	96
7 REFERÊNCIAS	99
8 APÊNDICE A – Protocolo de Serviço do ADL Gazetteer (esquemas XML)	101
9 APÊNDICE B – Casos de teste	109

## Lista de figuras

Figura 1 – Esquema do processo de identificação de áreas de alagamento ao redor de um rio	13
Figura 2 – Esquema de organização de dados geográficos definidos pela ISO 19115	20
Figura 3 – Esquema geral de metadados geográficos definidos pela ISO19115	21
Figura 4 – Componentes do <i>OpenGIS Services Framework</i>	27
Figura 5 – Camadas de serviços na arquitetura OWS	29
Figura 6 – Arquitetura interna de um serviço de catálogo	30
Figura 7 – Fragmento da hierarquia de termos do ADL Gazetteer Feature Type Thesaurus	35
Figura 8 – Um fragmento do ADL Feature Type Thesaurus e uma imagem da cidade do Rio de Janeiro	46
Figura 9 - Arquitetura de agentes para o processo de catalogação	48
Figura 10 - Processo de catalogação automática	51
Figura 11 - Recuperação de informações	54
Figura 12 - Esquema de metadados	57
Figura 13 – Diagrama de Casos de Uso	59
Figura 14 - Divisão de componentes do GeoCatalog segundo o padrão MVC	68
Figura 15 – Componentes da interface de usuário do GeoCatalog	69
Figura 16 - Janela de área de trabalho	70
Figura 17 - Estrutura do padrão de projeto <i>Command</i>	70
Figura 18 - Padrão <i>Command</i> no GeoCatalog	71
Figura 19 - Estrutura do padrão de projeto Observer	72
Figura 20 - Estrutura do padrão de projeto Factory Method	75
Figura 21 – Fábrica de objetos do GeoCatalog	76
Figura 22 - Infra-estrutura para criação de agentes	77
Figura 23 – <i>Framework</i> de catalogação automática	83
Figura 24 – Arquivo contendo a descrição de área geográfica do entorno do estádio do Maracanã (Rio de Janeiro)	89
Figura 25 – Arquivo contendo a descrição de uma região vazia	90

Figura 26 – Arquivo com erro de estrutura (falata o identificador)	90
Figura 27 – Configuração de processo de catalogação no GeoCatalog	91
Figura 28 – Metadados produzidos pela catalogação dos arquivos de dados anteriormente especificados	92
Figura 29 – Metadados produzidos pela catalogação dos arquivos de dados anteriormente especificados (cont.)	93
Figura 30 - Estratégia de teste	111

## Lista de tabelas

Tabela 1 – Conjunto de metadados principal definidos pela ISO 19115	22
Tabela 2 – Conjunto de metadados principal definidos pela ISO 19115 (cont.)	23
Tabela 3 – Elementos de metadado obrigatórios no padrão FGDC	24
Tabela 4 – Elementos de metadado obrigatórios no padrão FGDC (cont.)	25
Tabela 5 – Elementos de metadado obrigatórios no padrão FGDC (cont.)	26
Tabela 6 – Conjunto de metadados mínimo	31
Tabela 7 – Conjunto de metadados mínimo (cont.)	32
Tabela 8 – Relacionamentos entre conceitos de um tesouro	34
Tabela 9 – Alguns atributos selecionados do esquema conceitual da ISO 19112	36
Tabela 10 – Alguns atributos selecionados do ADL Gazetteer Content Standard	37