

Objetos Complexos em Bibliotecas Digitais: Analisando o Gerenciamento de Componentes de Imagens

Nádia Puchalski Kozievitch

Este exemplar corresponde à redação final da
Tese/Dissertação devidamente corrigida e defendida
por: NÁDIA PUCHALSKI KOZIEVITCH

é aprovada pela Banca Examinadora.
Campinas, 22 de agosto de 2011

COORDENADOR DE PÓS-GRADUAÇÃO
CPG-IC

Este exemplar corresponde à redação final da Tese devidamente corrigida e defendida por Nádia Puchalski Kozievitch e aprovada pela Banca Examinadora.

Campinas, 08 de Julho de 2011.

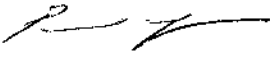
Prof. Dr. Julio César Lopez Hernández
Coord. Subst. de Pós-Graduação
Instituto de Computação/Unicamp
Matr. 28.620-1

Prof. Dr. Ricardo da S. Torres (Orientador
Instituto de Computação - UNICAMP)

Tese apresentada ao Instituto de Computação, UNICAMP, como requisito parcial para a obtenção do título de Doutora em Ciência da Computação.

ERRATA:

Desconsiderar informação no campo "Área de Concentração".


Prof. Dr. Paulo Lício de Geus
Coord. de Pós-Graduação
Instituto de Computação - Unicamp
Matrícula 10.326-8

FICHA CATALOGRÁFICA ELABORADA POR ANA REGINA MACHADO – CRB8/5467
BIBLIOTECA DO INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E
COMPUTAÇÃO CIENTÍFICA – UNICAMP

Unidade BCC
UNICAMP
Cutter K849
v. 1 Ed. 1
Imbo BC 92617
Proc. 16P 13Q 11
C D
Preço 71,00
Data 14/09/11
Cód. tit. 504652

K849o

Kozievitch, Nádía Puchalski

Objetos complexos em bibliotecas digitais :
analisando o gerenciamento de componentes de
imagens / Nádía Puchalski Kozievitch. – Campinas, SP :
[s.n.], 2011.

Orientador: Ricardo da Silva Torres.
Tese (doutorado) - Universidade Estadual de
Campinas, Instituto de Computação.

1. Bibliotecas digitais. 2. Sistemas de recuperação da
informação. 3. Processamento de imagens. 4.
Componente de software. I. Torres, Ricardo da Silva,
1977-. II. Universidade Estadual de Campinas. Instituto
de Computação. II. Título.

Informações para Biblioteca Digital

Título em Inglês: Complex objects in digital libraries : analyzing the management
of image components

Palavras-chave em Inglês:

Digital library

Information storage and retrieval systems

Image processing

Component software

Área de concentração: Banco de dados

Titulação: Doutor em Ciência da Computação

Banca examinadora:

Ricardo da Silva Torres [Orientador]

Marcos André Gonçalves

Islene Calciolari Garcia

André Santanchè

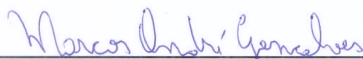
Carmem Satie Hara

Data da defesa: 08-07-2011

Programa de Pós Graduação: Ciência da Computação

TERMO DE APROVAÇÃO

Tese Defendida e Aprovada em 08 de julho de 2011, pela Banca examinadora composta pelos Professores Doutores:



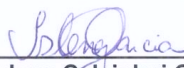
Prof. Dr. Marcos André Gonçalves - membro
DCC / UFMG



Profa. Dra. Carmem Satie Hara - membro
DInf / UFPR



Prof. Dr. André Santanchè - membro
IC / UNICAMP



Profa. Dra. Islene Calciolari Garcia - membro
IC / UNICAMP



Prof. Dr. Ricardo da Silva Torres - orientador
IC / UNICAMP

Objetos Complexos em Bibliotecas Digitais: Analisando o Gerenciamento de Componentes de Imagens

Nádia Puchalski Kozevitch¹

Julho de 2011

Banca Examinadora:

- Prof. Dr. Ricardo da S. Torres (Orientador
Instituto de Computação - UNICAMP)
- Prof. Dr. Marcos André Gonçalves
Departamento de Ciência da Computação - UFMG
- Prof^a. Dr^a. Islene Calciolari Garcia
Instituto de Computação - UNICAMP
- Prof. Dr. André Santanchè
Instituto de Computação - UNICAMP
- Prof^a. Dr^a. Carmem Satie Hara
Departamento de Informática - UFPR

¹Apoio financeiro Capes (processo 4479-09-2 - 2010).

Resumo

Atualmente, com os avanços tecnológicos, há disponível uma quantidade grande e heterogênea de dados multimídia. Estes dados podem ser provenientes de fontes diversificadas, desempenhando um papel fundamental em sistemas de informação, tais como repositórios de Bibliotecas Digitais. Com o objetivo de reutilizar, integrar, agregar e unificar diferentes recursos sob o mesmo conceito, objetos complexos surgiram para facilitar a abstração de agregações, utilizando componentes de diferentes domínios, e unificando-os sob o mesmo conceito. Em especial, os objetos complexos de imagens são um exemplo representativo de fonte de dados que geralmente é integrada com componentes heterogêneos, tais como metadados, links e softwares de manipulação de imagem. A busca de imagens é um serviço importante, amplamente explorado em sistemas de Biblioteca Digital. A recuperação de imagens por conteúdo aborda a busca automática de imagens, considerando propriedades visuais (como textura e cor). Este trabalho trata de três aspectos relacionados a objetos complexos: (i) a formalização de conceitos; (ii) a comparação de tecnologias relacionadas; e (iii) a especificação e a implementação de um protótipo, que encapsula e publica objetos complexos de imagens manipulados em sistemas de busca de imagens por conteúdo. A infraestrutura proposta baseia-se em uma estrutura de componentização específica - *Digital Content Component* (DCC) - para encapsular o processo de busca de imagens por conteúdo e o objeto complexo de imagem. Posteriormente, os dados são publicados como itens de um repositório de dados, usando o protocolo *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH). A solução proposta prevê (i) um acesso homogêneo ao processo de busca de imagens por similaridade, e fontes de dados (imagens, metadados e descritores); (ii) a distribuição e coleta de metadados de objetos complexos de imagens, juntamente com as medidas de similaridade e vetores de características; e (iii) a centralização do processamento, encapsulamento, publicação e coleta do objeto complexo de imagem. Finalmente, é ilustrado o reuso dos componentes da infraestrutura proposta no domínio de biodiversidade, na integração de bibliotecas digitais de impressões digitais, na busca multimodal e em histórias de vídeo.

Palavras-chave: Objeto Complexo, Busca de Imagens por Conteúdo, *Digital Content Component*, *Open Archives Initiative Protocol for Metadata Harvesting*, Biblioteca Digital.

Abstract

The large amount of heterogeneous data from different resources available today play a key role in information systems, such as Digital Library repositories. In order to reuse, integrate, aggregate, and unify different resources under the same concept, complex objects (COs) have emerged to facilitate aggregation abstraction, embracing components from different domains, and unifying them under the same concept. In special, image COs are a representative example of data source which is generally integrated with different components, such as metadata, links, and image manipulation software. Image search is an important service, widely explored in Digital Library systems. The so-called Content-Based Image Retrieval (CBIR) addresses the automatic search of images, trying to retrieve collection images by taking into account their visual properties (such as texture and color). This work is concerned with tackling three aspects related to complex objects: (i) formal definitions; (ii) comparison of related technologies; and (iii) the specification and implementation of a framework, which encapsulates and publishes complex image objects resulting from the CBIR process. Our infrastructure relies on a specific component technology - Digital Content Component (DCC) - to encapsulate the CBIR process and wrap image COs. Later, the data is exposed as items in a data repository, using Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). The proposed solution provides (i) an homogeneous access to CBIR process and heterogenous data sources (image collection, metadata, descriptors); (ii) the harvesting of complex image objects (ICOs), along with the similarity distances and feature vectors; and (iii) the centralization of ICO processing, packaging, publishing, and harvesting. Finally, we illustrate the reuse of our infrastructure components in biodiversity domain, in a fingerprint digital library integration, with multimodal search, and video stories.

Keywords: Complex Object, Content-Based Image Retrieval, Digital Content Component, Open Archives Initiative Protocol for Metadata Harvesting, Digital Library.

Agradecimentos

Eu me sinto abençoada por ter tido a oportunidade de conhecer as várias pessoas inspiradoras, que questionaram e participaram do processo desta tese.

Ao Prof. Ricardo, meu orientador de doutorado, por orientar, ensinar, encorajar e ter paciência, e principalmente por confiar em mim. Obrigada por ter me apresentado ao formalismo matemático e a busca por conteúdo. Especialmente obrigada pelas inúmeras horas utilizadas para refinar idéias e questionar a pesquisa.

À minha família, que mesmo estando longe, foi meu apoio. Em especial, ao meu pai Ricardo, que sempre me incentivou em alcançar algo mais à frente; à minha mãe Maria, por me lembrar que o simples é essencial; à minha irmã Esther, por estar incondicionalmente ao meu lado; à minha irmã Joselene, por sua alegria e paciência; ao André, pela compreensão em muitos momentos difíceis. Obrigada pelos conselhos e exemplos de determinação.

Às minhas amigas, que sempre me incentivaram. Rosana e Dorina, pelo exemplo de perseverança e dedicação. Angela, pela humildade e alegria. Simone, por acreditar em mim. Michele, pelo exemplo de coragem. Samara, por me abrigar e encorajar desde o princípio. Sandra e Miriam, pelo companheirismo. Aos amigos do LIS e Recod, em especial Eduardo Valle, Daniel Pedronette e Jurandy Almeida. Aos amigos do DLIB (*Digital Library Research Laboratory*, Virginia Tech) pelo apoio. Em especial à Uma, Venkat, Seungwon, Monika, Sung Hee, Susan, Nathan, Supratik. Aos alunos da Virginia Tech, que direta ou indiretamente participaram da pesquisa.

Ao Prof. Edward A. Fox, do Departamento de Ciências da Computação, por me receber na Virginia Tech, ensinar, orientar, e encorajar, não só durante os oito meses nos EUA, mas durante toda a pesquisa. Obrigada por me ensinar a avaliar meu trabalho usando outras perspectivas.

À Profa. Claudia Bauzer Medeiros, pelas sugestões, apoio, e por proporcionar várias oportunidades de aprendizado.

Ao Prof. André Santanchè, pelas sugestões, e apoio em aspectos essenciais para o desenvolvimento deste trabalho.

Aos professores e alunos do Departamento de Biologia da Unicamp. Em especial, à Rubens Riscala Madi, Marlene Tiduko Ueta e Silmara Marques Allegratti, por incentivar, ter paciência e participar de pontos importantes no desenvolvimento deste trabalho. Obrigada por me apre-

sentar a Parasitologia.

Aos Professores Michael Hsiao, Lynn Abbott (*Department of Electrical and Computer Engineering* - Virginia Tech), Eric Hallermann (*Department of Fisheries and Wildlife Sciences* - Virginia Tech), por toda a cooperação e paciência. Obrigada por incentivar o uso da tecnologia em áreas diversificadas.

Aos Professores do Instituto de Computação, por instigar e fazer funcionar o que foi aprendido. Aos funcionários do Instituto de Computação, por toda ajuda e apoio.

Aos inúmeros alunos que tive durante este processo, desde iniciação científica júnior (Maria, Scheila, Naira), até os alunos gerenciados durante a estadia na Virginia Tech (Sherley e Jennifer) e alunos de graduação de MC536.

Em especial, agradeço aos membros das bancas examinadoras, tanto de qualificação quanto de defesa de tese, Professores André Santanchè, Claudia Bauzer Medeiros, Cecília Rubira, por todos os questionamentos, sugestões e discussões que me ajudaram a crescer como pesquisadora.

Finalmente, este trabalho de pesquisa seria muito difícil sem o apoio da Bolsa CAPES de Doutorado Sanduíche (processo número 4479-09-2), FAPESP e CNPq. Agradeço ainda ao NIJ (*National Institute of Justice*), BAE Systems, e NSF IIS-0910183, IIS-0916733, DUE-0840719, e CCF-0722259.

Sumário

Resumo	vii
Abstract	ix
Agradecimentos	xi
1 Introdução	1
1.1 Motivação	1
1.2 Metodologia e Aspectos de Pesquisa Envolvidos	3
1.3 Objetivos e Contribuições	5
1.4 Organização da Tese	7
1.5 Principais Termos e Acrônimos Utilizados	8
2 Trabalhos Relacionados	9
2.1 Bibliotecas Digitais	9
2.2 Formalismos para Bibliotecas Digitais	9
2.3 Softwares de Bibliotecas Digitais	13
2.3.1 Fedora	14
2.3.2 Eprints	15
2.3.3 CDS/ISIS	16
2.3.4 Greenstone	16
2.3.5 Harvest	17
2.3.6 Dspace	17
2.3.7 Koha	19
2.4 Objetos Complexos	19
2.5 Tecnologias Relacionadas a Objetos Complexos	20
2.5.1 Digital Content Component	22
2.5.2 Buckets	24
2.5.3 OAI-ORE	25
2.6 Recuperação de Imagens por Conteúdo	28

2.7	Conclusões	29
3	Estudos de Caso Preliminares e Aplicações	31
3.1	<i>Superimposed Image Description and Retrieval Tool - SuperIDR</i>	31
3.2	A Integração de Bibliotecas Digitais de Impressões Digitais	34
3.3	Nou-Rau: Biblioteca Digital de Teses e Dissertações da Unicamp	37
3.4	Conclusões	39
4	Formalização de Objetos Complexos e Busca de Imagens por Conteúdo	41
4.1	Notações Iniciais do Arcabouço 5S	41
4.2	Objetos Complexos	42
4.3	Busca de Imagens por Conteúdo	46
4.4	Estudo de Caso	50
4.5	Conclusões	52
5	Analisando as Tecnologias de Objetos Complexos da Perspectiva do Arcabouço 5S	53
5.1	Analisando o DCC, Buckets, e OAI-ORE	53
5.1.1	Uso do Formalismo 5S para DCC, Buckets e OAI-ORE	54
5.1.2	Explorando as Diferenças entre DCC, Buckets e OAI-ORE	56
5.2	Estudo de Caso	58
5.3	Conclusões	61
6	Uma Infraestrutura baseada em DCC-OAI para Objetos Complexos de Imagem	63
6.1	Visão Geral da Infraestrutura Proposta	63
6.1.1	A Busca de Imagens por Conteúdo	64
6.1.2	A Publicação de Objetos Complexos	69
6.2	Estudo de Caso	70
6.2.1	A Implementação	71
6.2.2	O Encapsulamento de Recursos	72
6.2.3	A Infraestrutura	72
6.3	Conclusões	73
7	Cenários de Reutilização da Infraestrutura Proposta	77
7.1	A Integração de Bibliotecas Digitais de Impressões Digitais	78
7.1.1	O Protótipo	78
7.1.2	A Implementação	82
7.2	Busca Multimodal	82

7.2.1	Utilizando DCCs para a Busca Multimodal	84
7.2.2	A Publicação dos COs	85
7.2.3	Estudo de Caso	85
7.3	Busca em Estórias de Vídeo	87
7.3.1	A Geração da Sumarização de Estórias de Vídeos	87
7.3.2	Utilizando DCCs para a Busca de Estórias	87
7.3.3	Estudo de Caso	88
7.4	Conclusões	90
8	Conclusões	93
	Bibliografia	98

Lista de Tabelas

2.1	Conceitos básicos de CO da perspectiva do METS, SCORM e MPEG-21. . . .	22
2.2	Principais métodos públicos implementados por Buckets [69].	26
3.1	Características das três aplicações utilizadas nesta tese.	39
5.1	Conceitos básicos de CO da perspectiva de DCC, Buckets, e OAI-ORE. . . .	57

Lista de Figuras

1.1	Visão geral da metodologia empregada nesta tese [51].	4
1.2	As perspectivas de CO tratadas nesta tese.	5
1.3	As principais contribuições desta tese.	6
1.4	A organização desta tese de doutorado.	7
2.1	A estrutura de definição do Arcabouço 5S [115].	11
2.2	As ferramentas relacionadas ao Arcabouço 5S [32].	12
2.3	O modelo de referência mínimo para Bibliotecas Digitais [82].	13
2.4	O modelo do objeto digital Fedora [22].	14
2.5	O modelo de dados de <i>Eprints</i> [45].	15
2.6	Exemplo de plugins disponíveis no from Greenstone [131].	17
2.7	O modelo de dados do Dspace [101].	18
2.8	A estrutura do <i>Digital Content Component</i>	24
2.9	A implementação do arquivo Jar Fluid Web e seus principais blocos.	24
2.10	A Arquitetura de Buckets (adaptado de [69]).	25
2.11	Mapeando conceitos no Arcabouço 5S e OAI-ORE [54].	27
2.12	Exemplo de uso do OAI-ORE [123].	27
3.1	A utilização de anotações para comparar duas espécies no SuperIDR.	32
3.2	A navegação taxonômica na Ferramenta SuperIDR.	33
3.3	A Ferramenta SuperIDR: A- Navegação, B- Comparação, C- Anotação e D - Busca Textual.	34
3.4	Exemplos de imagens de parasitas utilizadas no estudo de caso.	34
3.5	As características analisadas em uma impressão digital.	35
3.6	As características analisadas em impressões digitais retiradas de uma cena do crime.	36
3.7	(A) Comparando Imagens. (B) Distorção de Imagem. (C) <i>Blurring</i> . (D) Cena do Crime.	37
3.8	O sistema Nou-Rau (último acesso em 20/07/2011).	38
4.1	A estrutura de definição do Arcabouço 5S [115].	43

4.2	Exemplo de um objeto complexo que representa uma tese.	45
4.3	Exemplo de uma estrutura XML agregando os elementos do objeto complexo tese.	45
4.4	Extensões do Arcabouço 5S para a descrição de imagens por conteúdo e serviços relacionados.	46
4.5	Exemplo de um <i>structured feature vector</i>	47
4.6	(a) O uso de um descritor simples <i>D</i> para computar a similaridade de imagens. (b) Descritor de imagem composto.	48
4.7	Os elementos do objeto digital de imagem.	49
4.8	Busca de imagem por conteúdo para a imagem 01_ancylostoma.jpg.	50
4.9	Exemplo de uma estrutura XML para um objeto complexo de imagem.	51
4.10	O Objeto Complexo de Imagem.	51
5.1	A integração de bibliotecas de impressões digitais.	58
5.2	As classes principais representando uma biblioteca digital de impressões digitais.	59
5.3	Um exemplo de objeto complexo com as quatro bibliotecas digitais: (A) Im- pressões Digitais armazenadas, (B) Imagens distorcidas, (C) Imagens da Cena do crime, and (D) Material de Treinamento.	60
6.1	Arquitetura de uma biblioteca digital baseada em COs.	64
6.2	(a) Estrutura do ImageDCC. (b) Estrutura do ImageCODCC.	65
6.3	Exemplo de uma estrutura XML para o objeto complexo de imagem.	66
6.4	Estrutura do DescriptorDCC.	67
6.5	(a) A Taxonomia dos DCCs para a manipulação de ICOs. (b) As camadas de gerenciamento.	68
6.6	Os objetos digitais disponíveis para publicação.	70
6.7	A resposta da requisição OAI <i>ListFeatureVectors</i>	71
6.8	(a) Resultado da consulta CBIR para a imagem 01_ancylostoma.jpg. (b) Resul- tado da consulta CBIR para a imagem 06_ancylostoma.jpg.	73
6.9	Estrutura para SpeciesDCC.	74
6.10	(a) XML para o CO <i>Ancylostoma caninum</i> . (b) A requisição OAI <i>ListIdentifiers</i> para os ICOs de Parasitas.	75
7.1	A reutilização da infraestrutura baseada em DCC e OAI-PMH.	77
7.2	Exemplos de impressões digitais obtidas a partir do registro da polícia.	79
7.3	Exemplos de impressões digitais que simulam uma cena de crime.	80
7.4	Resultado da consulta CBIR para a Figura 7.2 - parte 11.	81
7.5	Resultado da consulta CBIR para a Figura 7.3 - parte 3.	82
7.6	Estrutura do IndividualDCC.	83

7.7	XML para a agregação do conceito de indivíduo.	84
7.8	A estrutura do MultimodalDCC.	85
7.9	Resultado da consulta CBIR para a imagem <i>01_ancylostoma.jpg</i>	86
7.10	(a) Resultados da Busca Textual. (b) Interface para a Busca Multimodal.	86
7.11	Um exemplo de estórias produzidas para o vídeo <i>A New Horizon, segment 04</i> . .	88
7.12	A estrutura do CBIRStoryDCC.	89
7.13	Resultado da consulta CBIR para uma estória do vídeo <i>A New Horizon, segment 04</i>	90

Capítulo 1

Introdução

1.1 Motivação

Considere uma família. Uma família nada mais é do que um conceito único, inicialmente composto por pais e filhos; posteriormente estendido a todas as pessoas que vivem em uma mesma casa. Juntos são componentes que formam um grupo, resultado da união das partes de um todo.

Para uma família continuar sobrevivendo como um conceito único (uma agregação), precisa se **adaptar** às diferentes perspectivas de seus componentes, aos seus comportamentos heterogêneos, à sua evolução (ao crescimento dos filhos, à união do casal, aos namorados e namoradas que aparecerão com o tempo, netos repentinos, etc.), e alguns padrões sociais de comportamento (ao menos um adulto deve trabalhar, as crianças deverão ir a escola, haverá regras para a limpeza do quarto e para os jogos, etc.).

Um componente pertencente a uma família é um indivíduo único, uma entidade física única. Se uma família se separa, a divisão dos componentes requer uma nova adaptação, uma visão nova para uma realidade diferenciada. O todo torna-se a soma dos segmentos das partes, e certas perspectivas e serviços podem não estar mais disponíveis (aquele quarto que era seu já está alugado, e os quitutes de sua mãe já não lhes pertencem).

Considere a abstração do conceito de **grupo, componentes heterogêneos, componente único, integração, interface**. Suponha os mesmos conceitos, agora no contexto de uma tese de doutorado. Um doutorando pode ter várias publicações em áreas distintas (integradas a uma aplicação, a formalismos e a implementações), em congressos distintos (nacionais, internacionais), e de tipos distintos (posters, seminários, *short paper*, *full paper*, etc.). Considere ainda que esta tese esteja conectada a diversas mídias (vídeo, áudio, imagens, etc.). Além dos conceitos apresentados, há ainda a questão de como agregar estes componentes da tese, como processá-los, encapsulá-los e apresentá-los em aplicações.

No mundo real e em bibliotecas tradicionais, os recursos são entidades físicas (objetos),

com um início e fim (como livros, jornais, CD-ROMS, revistas). E geralmente estes recursos têm índices, catálogos, ou resumos, para facilitar buscas. Em bibliotecas digitais, um recurso não necessariamente é associado a uma entidade física, e torna-se difícil definir um começo ou um fim. Esta necessidade fundamenta o uso de *máquinas de busca*, levando usuários diretamente a componentes como “capítulos” de um “livro” nos quais estejam interessados. Este exemplo ilustra quão difícil é a determinação do que constitui o recurso digital a ser catalogado (**granularidade**). Além disso, é possível relacionar estas sub-partes (criando redes), através de anotações, links, referências, etc.

Cada um destes dois cenários pode ser considerado exemplo de um Objeto Complexo (*Complex Object* - CO), composto por várias partes, conectadas entre si. Em especial, com recursos digitais, o uso do conceito de objetos complexos facilita o gerenciamento da granularidades diferentes, a identificação única dos componentes, e a composição ou o encapsulamento dos recursos. Em um CO podem ser incluídos vários elementos distintos, como links, documentos, anotações, mídia, áudio, etc. Junto com as mídias diversificadas, há a necessidade do suporte adequado aos serviços de busca, navegação, processamento, etc. Conforme [8, 104], a heterogeneidade de objetos complexos apresenta os seguintes desafios:

1. suporte inadequado de COs pelos softwares de Bibliotecas Digitais (*Digital Library* - DL);
2. algumas características específicas das partes do CO (como regras de acesso, segurança e propriedade) dificultam o tratamento do CO como um objeto único; e
3. o sistema de busca tem que permitir não somente a busca dos componentes individuais, mas também a busca do CO como um objeto único.

Em especial, objetos complexos relacionados a imagens cada vez adquirem uma importância maior. Nos últimos anos a quantidade de imagens digitais disponíveis vem aumentando rapidamente. Entre os principais responsáveis estão a popularização das câmeras fotográficas digitais, a disseminação da Internet de alta velocidade e a facilidade de armazenamento. Estes fatores permitiram, entre outros, que a quantidade de imagens geradas aumentasse rapidamente e que estas imagens fossem colocadas em domínio público. Tudo isso significa uma quantidade gigantesca de informação visual presente no cotidiano de um número crescente de pessoas. Estes objetos geralmente estão associados a links, metadados, ou diferentes tipos de software. O serviço mais utilizado para este tipo de mídia é a busca (mais especificadamente a busca de imagens por conteúdo), também amplamente utilizada em sistemas de Biblioteca Digital. A recuperação de imagens por conteúdo aborda a busca automática de imagens, considerando propriedades visuais (como textura e cor).

Ao combinar conceitos de objetos complexos e bibliotecas digitais no domínio de recuperação de imagens, apresentamos uma solução alternativa no suporte para criação, manutenção e integração

destes sistemas heterogêneos. Acreditamos que a comunidade de sistemas de informação (especialmente a de bibliotecas digitais), possa obter proveito das definições formais, exemplos de cenários e integrações, comparação de tecnologias relacionadas, e exemplos em protótipos com utilização dos conceitos na prática. Em especial, como aplicações preliminares, podemos citar:

- CTRnet ¹: o projeto *Crisis, Tragedy, and Recovery Network* objetiva desenvolver melhores abordagens para tornar a tecnologia útil para arquivamento de informações sobre eventos envolvendo crises e tragédias. Este projeto é formado por coleção de eventos, onde cada evento é retratado por uma coleção de imagens, como o terremoto no Haiti [53, 134].
- Sistemas de Informação Geográfica: estes sistemas utilizam modelos *raster* ou vetorial, e podem ser associados com imagens [57] ou anotações [31] para representar a localização de eventos ou serviços;
- Uma biblioteca multilíngue para países de língua portuguesa [51]: este protótipo apresenta a integração de documentos (doc, mp3, áudio) em onze línguas, utilizando o software Greenstone.

Além das aplicações preliminares, outras aplicações foram utilizadas nesta tese, para descrever os estudos de caso, e serão detalhadas no Capítulo 3.

Esta tese objetiva (i) a formalização de conceitos relacionados a Objetos Complexos; (ii) a comparação de tecnologias relacionadas; e (iii) a especificação e a implementação de uma infraestrutura, que encapsula e publica objetos complexos de imagens manipulados em sistemas de busca de imagens por conteúdo. A infraestrutura é baseada na utilização do modelo de componentes *Digital Content Component* (DCC) para encapsular o processo de busca de imagens por conteúdo e o objeto complexo de imagem. Posteriormente, os dados são publicados como itens de um repositório de dados, usando o protocolo *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH). Posteriormente é ilustrada a integração desta infraestrutura em bibliotecas digitais de impressões digitais, consultas multimodais e histórias de vídeo.

1.2 Metodologia e Aspectos de Pesquisa Envolvidos

Com o objetivo de desenvolver uma infraestrutura para gerenciar COs em DLs, verificando a sua caracterização formal, é necessária a compreensão clara de uma DL com tipos diversificados de dados. Acreditamos que a formalização através do arcabouço 5S e o encapsulamento através de DCCs irá fornecer soluções eficazes para COs em DLs. Essa hipótese leva às seguintes considerações (Figura 1.1):

¹<http://www.ctrnet.net/> (último acesso em 05/05/11).

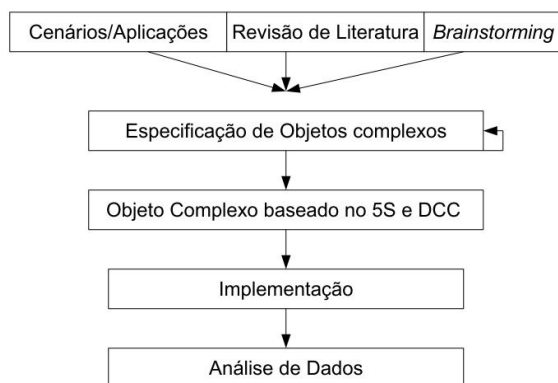


Figura 1.1: Visão geral da metodologia empregada nesta tese [51].

- Explorar CO a partir de uma perspectiva teórica, aprendendo sobre trabalhos relacionados e na literatura de DL e áreas afins. Verificamos a sua definição formal, a integração e agregação das partes do CO, a flexibilidade com novos formatos e padrões, arquitetura e serviços básicos, como a navegação e a busca. A partir disso, desenvolvemos uma especificação preliminar de um modelo de CO utilizando o Arcabouço 5S;
- Refinar estas especificações a partir da aplicação em estudos de caso;
- Especificar um fundamento teórico para objetos complexos e serviços em bibliotecas digitais utilizando o Arcabouço 5S. O Arcabouço 5S foi estendido visando a definição de COs;
- Avaliar este modelo de dados a partir de uma perspectiva prática, com a agregação com DCC, usando imagens, e encapsulamento de dados (sob a perspectiva de CO). Uma infraestrutura foi especificada e implementada, com serviços como a navegação e a busca, com o intuito de se identificar conceitos importantes relacionados à integração de dados.

Os principais desafios tratados nesta tese são:

- Como formalizar uma definição da conexão / agregação entre componentes de COs e seus serviços?
- Como definir formalmente uma estrutura unificada para a gestão de COs em bibliotecas digitais? Como definir um CO mínimo?
- Como tratar o relacionamento das partes de um objeto complexo? Qual tipo de estrutura pode ser utilizada?

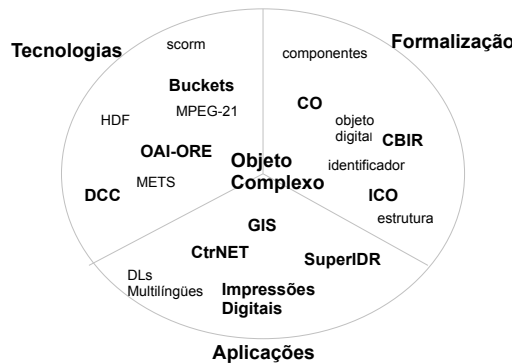


Figura 1.2: As perspectivas de CO tratadas nesta tese.

- Se uma aplicação manipula COs cujos componentes incluem dados de tipos diferentes (como arquivos de texto, áudio, imagem, vídeo e software), como é o acesso aos serviços do CO e de seus componentes?
- Como lidar com os problemas de integração de dados (como a integração de dois diferentes objetos complexos em um único modelo) e serviços a partir de uma perspectiva de CO?
- É possível descrever, comparar e integrar as tecnologias relacionadas a Objetos Complexos?
- Se o enfoque é realizado em um domínio (como o de imagens, por exemplo), como gerenciar e agregar metadados, arquivos, softwares, de modo a facilitar o seu reuso e manutenção?

1.3 Objetivos e Contribuições

O objetivo desta tese é contribuir na solução de problemas de especificação e implementação de sistemas de informação que utilizem objetos complexos, em especial, relacionados a imagens. O trabalho é baseado em dois conceitos básicos: (i) a noção de objetos complexos, utilizada para integrar e agregar conceitos e recursos heterogêneos; (ii) o arcabouço de bibliotecas digitais, para permitir o gerenciamento e manutenção de coleções de dados grandes e heterogêneas. O uso destes dois conceitos permite a criação, manutenção e integração de sistemas sob uma perspectiva de processos diferentes. A partir destes conceitos básicos, três perspectivas são analisadas: formalização, tecnologias de CO e aplicações. A Figura 1.2 ilustra os principais conceitos tratados dentro das perspectivas de aplicações, tecnologias e formalização. Os conceitos em negrito apresentam tópicos que são analisados com maior profundidade nesta tese.

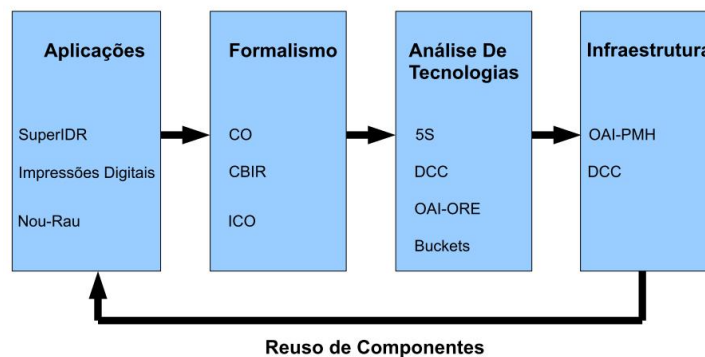


Figura 1.3: As principais contribuições desta tese.

A partir de descrição de aplicações, esta tese trata do formalismo, da análise de tecnologias, a criação de uma infraestrutura e o reuso de seus componentes (como ilustrado na Figura 1.3). Mais especificadamente, as principais contribuições desta tese são:

- formalização de conceitos relacionados a objetos complexos e busca de imagens por conteúdo, ilustrando o objeto complexo baseado em imagens da aplicação SuperIDR;
- comparação e análise de tecnologias de objetos complexos (DCC, Buckets e OAI-ORE), da perspectiva do arcabouço 5S;
- proposta e desenvolvimento de um protótipo que provê:
 - um acesso homogêneo ao processo de busca de imagens por similaridade, e fontes de dados (imagens, metadados e descritores);
 - a distribuição e coleta de metadados de objetos complexos de imagens, juntamente com as medidas de distância e vetores de características;
 - a centralização do processamento, encapsulamento, publicação e coleta do objeto complexo de imagem.
- exemplos de cenários que utilizam integração e agregação de informações, ilustrando serviços e conceitos relacionados a objetos complexos;
- exemplo de integração de metadados, imagem, vetores de características e processamento na aplicação SuperIDR, utilizando imagens de parasitas;
- exemplo de integração de metadados, imagem, vetores de características e processamento de quatro bibliotecas digitais de impressões digitais, utilizando a perspectiva de objeto complexo;

- o uso da infraestrutura desenvolvida para buscas multimodais;
- o uso da infraestrutura para a busca em sumarização de vídeos.

1.4 Organização da Tese

Esta tese é composta por oito capítulos, como mostra a Figura 1.4.

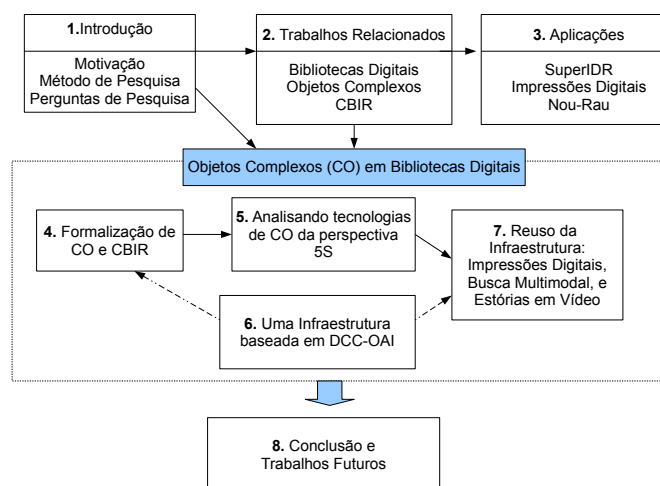


Figura 1.4: A organização desta tese de doutorado.

O próximo capítulo (*Trabalhos Relacionados*) lista os trabalhos relacionados a esta tese, tal como Bibliotecas Digitais, objetos complexos, tecnologias de CO, e CBIR.

O Capítulo 3 (*Estudos de Caso Preliminares e Aplicações*) apresenta aplicações utilizadas nesta tese.

O Capítulo 4 (*Formalização de Objetos Complexos e Busca de Imagens por Conteúdo*) apresenta as definições formais para Objetos Complexos e Recuperação de Imagens por Conteúdo. Em particular, é descrito o objeto complexo de imagem, exemplificado no contexto da ferramenta SuperIDR.

O Capítulo 5 (*Analisando as Tecnologias de Objetos Complexos da Perspectiva do Arcabouço 5S*) apresenta a análise das tecnologias relacionadas a objetos complexos, da perspectiva do arcabouço 5S. Este capítulo (1) analisa três tecnologias utilizadas em objetos complexos (DCC, Buckets, OAI-ORE); (2) utiliza o arcabouço 5S para descrevê-las; e (3) apresenta um estudo de caso, ilustrando como o arcabouço 5S e as tecnologias baseadas em objeto complexo podem ser utilizadas para explorar informações compostas.

O Capítulo 6 (*Uma Infraestrutura baseada em DCC-OAI para Objetos Complexos de Imagem*) apresenta uma infraestrutura baseada em *Digital Content Component* (DCC) e OAI-PMH

para processar, realizar busca e publicar objetos complexos de imagens.

O Capítulo 7 (*Cenários de Reutilização da Infraestrutura Proposta*) descreve três cenários onde a infraestrutura foi reutilizada, como a busca multimodal e integração com histórias sumariadas de vídeo.

O Capítulo 8 apresenta conclusões e trabalhos futuros.

1.5 Principais Termos e Acrônimos Utilizados

Arcabouço 5S (*5S Framework*): é um esquema para a concepção, criação e modelagem de bibliotecas digitais. A interpretação desta construção dá-se sob o conceito ou modelos baseados em *Stream*, *Structures* (Estruturas), *Spaces* (Espaços), *Scenarios* (Cenários) e *Societies* (Sociedades).

Biblioteca Digital (*Digital Library - DL*): organizações que fornecem recursos, incluindo o suporte especializado para selecionar, organizar, interpretar, distribuir, preservar a integridade e garantir a persistência ao longo do tempo de coleções digitais de obras.

Busca de Imagens por Conteúdo - (*Content-Based Image Retrieval- CBIR*): consiste na recuperação automática de imagens, com base nas propriedades visuais, tais como textura, forma ou cor.

Objeto Complexo (*Complex Object - CO*): é a agregação ou composição de objetos digitais. Algumas referências da literatura [9] consideram que esta agregação está interligada por uma rede.

Digital Content Component (*DCC*): modelo baseado em componentes, capaz de encapsular e tratar exclusivamente qualquer tipo de conteúdo.

Objeto Complexo de Imagem (*Complex Image Object - ICO*): nesta tese é representado pelo objeto complexo de imagem, composto por uma imagem, um arquivo de vetores de características e um arquivo XML (com medidas de distância e metadados).

Open Archives Initiative Protocol for Metadata Harvesting(*OAI-PMH*): é um protocolo baseado na arquitetura cliente-servidor, desenvolvido pela *Open Archives Initiative*, utilizado para publicar (*expose*) e coletar (*harvest*) metadados principalmente relativos a descritores de documentos.

Capítulo 2

Trabalhos Relacionados

Este capítulo apresenta os principais trabalhos relacionados a esta tese, referentes a bibliotecas digitais, formalismos em bibliotecas digitais, objetos complexos, tecnologias usadas para o gerenciamento de objetos complexos e recuperação de imagens por conteúdo.

2.1 Bibliotecas Digitais

Grossman [39] define uma Biblioteca Digital (*Digital Library* - DL) como um sistema que é projetado para selecionar, recuperar e gerenciar dados complexos, construído sobre o modelo de mineração de dados. A *Digital Library Federation* - DLF¹ define Bibliotecas Digitais como organizações que fornecem recursos, incluindo o suporte especializado para selecionar, organizar, interpretar, distribuir, preservar a integridade e garantir a persistência ao longo do tempo de coleções digitais de obras. Além disso, promove que sua disponibilidade seja de maneira rápida e eficiente para uso por uma comunidade específica ou conjunto de comunidades [29].

Informalmente, sob a perspectiva do Arcabouço 5S [33], uma biblioteca digital é uma coleção gerenciada de informações com serviços associados, envolvendo comunidades onde informações são armazenadas em formato digital e acessadas através de uma rede. A informação em bibliotecas digitais se manifesta em termos de objetos digitais, com conteúdo textual ou multimídia (como por exemplo, imagens, áudio, e vídeo), e metadados. Os serviços básicos oferecidos pelas bibliotecas digitais são a indexação, a busca e a navegação [33].

2.2 Formalismos para Bibliotecas Digitais

A modelagem, definição, e descrição de bibliotecas digitais ainda é um desafio, como exemplificado em [4, 15, 68, 78, 82]. Há várias abordagens disponíveis, como o *Dexter Hypertext*

¹<http://www.diglib.org/> (último acesso em 05/05/11).

Reference Model [24, 38], o *Amsterdam Model* [40], o *Guidelines for Electronic Text, Encoding and Interchange* (TEI-P5 [117]), o modelo de Referência DELOS [14], o *Open Provenance Model* (OPM) [75], o *Component-Based Authoring* de Park et al. [94], e o Arcabouço 5S [34].

O *Dexter Hypertext Reference Model* [24, 38] define um componente composto, que pode ser um conjunto de quadros, um *iframe*, um link para outro objeto (como um clip de vídeo), ou um componente de ligação. Este modelo não considera a estrutura interna do componente, ou descrição / formato do conteúdo.

O modelo Dexter foi estendido para fornecer suporte sincronizado a hipermídia. Esta extensão é conhecida como *Amsterdam Model* [40]. Neste caso, em contraste com os componentes de hipertexto, componentes de multimídia são representados em uma ordem pré-determinada. Para expressar este item explicitamente, *timelines* são introduzidos.

As Diretrizes TEI-P5 [117], outra iniciativa com base na marcação e hipertexto, nasceu de uma conferência de planejamento promovida pela *Association for Computers and the Humanities* (ACH). TEI-P5 provê uma variedade de mecanismos de links, que podem ser usados para indicar o alinhamento temporal e agregação de fragmentos estruturas, tais como o uso de partes no elemento *<lg>*, o uso dos atributos *next* e *prev* para a união de partes de um elemento, e o uso da estrutura *<join>* para definir um elemento composto de fragmentos. Porém o TEI-P5 não considera a formalização destes conceitos.

Na comunidade de Biblioteca Digital, o modelo de referência DELOS [14] também gerencia recursos que podem ser compostos de recursos menores (com a relação *<hasPart>*) e links a outros recursos (com a relação *<associatedWith>*), para formar objetos compostos. No entanto, o modelo de referência DELOS não considera a formalização detalhada destes conceitos, nem a definição de abstrações, como serviços em termos de cenários [14].

Ainda na comunidade de Biblioteca Digital, podemos citar o *Open Provenance Model* (OPM) [75], e o *Component-Based Authoring Tool* [94]. Em essência, o *Open Provenance Model* é responsável pelo controle de proveniência de objetos, consistindo de gráficos, dependências, objetos (nós) e regras (como “gerado por”, “controlado por”). Mesmo permitindo restrições temporais e anotações, não abrange outros conceitos de biblioteca digital, como serviços. O *Component-Based Authoring Tool*, em contrapartida, utiliza um formalismo baseado em redes de Petri para referenciar a modelagem unificada da infraestrutura bibliotecas digitais.

O Arcabouço 5S

O Arcabouço 5S (*Streams*, *Structures* (Estruturas), *Spaces*, (Espaços), *Scenarios* (Cenários), e *Societies* (Sociedades)) foi proposto em [32–34], como uma teoria formal para descrever bibliotecas digitais. O arcabouço busca descrever vários aspectos de bibliotecas digitais, como mostra a Figura 2.1. A figura ilustra as camadas de apoio das definições: fundamentos

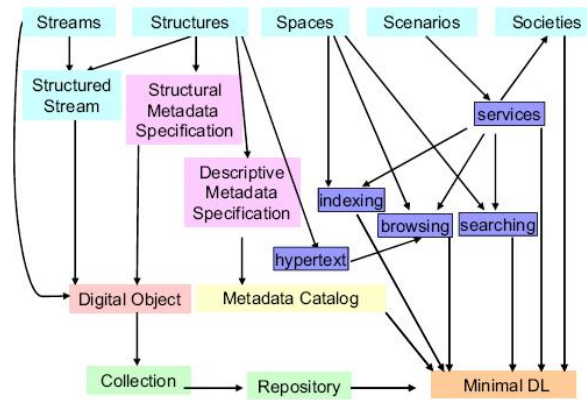


Figura 2.1: A estrutura de definição do Arcabouço 5S [115].

matemáticos (e.g., gráficos, sequências e funções), os 5 Ss (*Streams*, *Structures* (Estruturas), *Spaces* (Espaços), *Scenarios* (Cenários), e *Societies* (Sociedades), e os conceitos-chave de uma biblioteca digital (por exemplo, objeto digital (*Digital Object*), coleção (*Collection*), catálogo dos metadados (*Metadata Catalog*), repositório (*Repository*)). Setas representam dependências, indicando que um conceito é formalmente definido em termos de outros conceitos que apontam para ele.

No formalismo 5S:

- *Streams* modelam conteúdo estático e dinâmico (e.g., texto, vídeo, áudio). O conteúdo estático inclui por exemplo, textos, enquanto o conteúdo dinâmico inclui uma apresentação de vídeo, ou uma sequência temporal e localização para um objeto em movimento (como GPS, por exemplo).
- *Structures* (Estruturas) especificam aspectos organizacionais do conteúdo de uma biblioteca digital. Em bibliotecas digitais estruturas podem representar hipertexto, taxonomias, conexões de sistemas, ou encapsulamento, por exemplo.
- *Spaces* (Espaços) definem visões de apresentação e suportam 2D, 3D, Realidade Virtual, GIS, etc. A combinação de operações sobre os objetos no conjunto é o que distingue os *Spaces* de *Streams* e *Structures*. Uma vez que esta combinação é uma construção poderosa, ela também é utilizada quando uma parte de uma DL não pode ser descrita usando outros itens do arcabouço.
- *Scenarios* (Cenários) consistem de sequências de eventos ou ações que basicamente descrevem o comportamento de um serviço de biblioteca digital tal como busca, navegação, submissão, etc. O modelo de *Scenarios* fornece orientações para construir o custo-benefício de um protótipo, ou ajudar a validar, inferir, e especificar requisitos de suporte

Figura 2.2: As ferramentas relacionadas ao Arcabouço 5S [32].

e fornecer critérios de aceitação para o teste.

- *Societies* (Sociedades) compreendem as entidades e os relacionamentos entre elas. As entidades incluem seres humanos, bem como de hardware e componentes de software, que utilizam ou apoiam serviços de biblioteca digital. Relacionamentos sociais realizam conexões entre as entidades.

Junto com a formalização do 5S, algumas ferramentas [33, 76, 112, 115, 135] foram criadas para trabalhar com o arcabouço 5S: 5SL, 5SGraph, 5SGen, 5SQual, 5SSuite. 5SL é uma linguagem declarativa para a especificação e geração de *DLs*. 5SGraph é uma ferramenta de modelagem de domínio específico da biblioteca digital que foi concebida em um trabalho original com 5SL. 5SGen é uma ferramenta que gera grande parte do código necessário para a implementação de uma *DL*, utilizando como fonte uma biblioteca digital definida no 5SL. 5SQual fornece meios para verificar a qualidade dos objetos digitais, metadados e serviços. E o 5SSuite é usado para cobrir o processo da integração e união de *DLs*, incluindo o levantamento de requisitos, modelagem conceitual, prototipagem, e geração de código. A ferramenta 5SSuite consiste no 5SGraph, 5SGen e SchemaMapper (uma ferramenta que mapeia um esquema local em global para uma união de *DLs*). Mas essas ferramentas ainda não têm suporte para objetos complexos, com sub-partes e anotações, permitindo ligações entre componentes. A Figura 2.2 mostra como o Arcabouço 5S suporta a análise e especificação de *DLs*.

O Arcabouço 5S também foi utilizado para elaborar o conceito de qualidade em *DLs* [35]. O conceito de qualidade foi utilizado em [35] para procurar fatores e variáveis envolvidos na mensuração de conceitos como objeto digital, especificação de metadados, coleção, catálogos e repositórios. Mas a avaliação de *DLs* utilizou outros parâmetros também, como dados, sistemas de coleta, tecnologia, usuários e utilização [29]. Para cada uma destas dimensões, atributos principais foram descritos para o estabelecimento de uma meta-biblioteca. O mesmo conceito de qualidade pode ser usado na avaliação formal para o CO, em relação à sua completude.

Outro problema presente em *DLs* é a especificação de um modelo de referência mínimo. Murphy et al. [78, 82] usaram o Arcabouço 5S para fornecer especificações, sendo o modelo de referência mínimo uma fundação para várias extensões, como mostrado na Figura 2.3. Na primeira extensão há a recuperação de imagens baseado em conteúdo (*Content-Based Image Retrieval* - CBIR), o segundo é um meta-modelo para anotações superimpostas e o terceiro trata da geração *DLs* através de software, tais como o Dspace. Em [82] é realizada a formalização de repositórios, catálogos coletivos, os serviços mínimos de união, esquemas, mapeamentos 1-1, mapeamentos complexo e de uma biblioteca digital de união mínima, integrando um conjunto de *DLs*. Para a identificação de dois elementos semelhantes, uma pontuação de confiança é

utilizada para a função de mapeamento: quanto maior a pontuação de confiança, mais semanticamente similares são dois objetos.

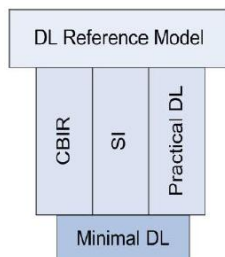


Figura 2.3: O modelo de referência mínimo para Bibliotecas Digitais [82].

Gorton [36] por exemplo, usou o 5SL, 5SGraph e 5SGen para desenvolver um modelo baseado em XML com o objetivo de especificar a natureza da biblioteca digital DSpace, utilizando uma parte da *DL CITIDEL* (*Computing and Information Technology Interactive Digital Educational Library*). O estudo de caso em [36] gerou uma instância DSpace com todo o conteúdo da coleção CSTC (conteúdo para o ensino de Ciências da Computação).

Já o 5SSuite foi utilizado em [115], tratando outro problema relacionado a *DLs*: a formalização da integração em *DLs*. Um metamodelo mínimo foi desenvolvido utilizando o kit de ferramentas de integração 5SSuite, ilustrando as soluções para os principais problemas com um estudo de caso em uma *DL* de arqueologia. Os conceitos do Arcabouço 5S foram utilizados ainda na ferramenta WS-ODL [112], um *framework* baseado na arquitetura web, utilizando Fedora para a construção de bibliotecas digitais.

Entretanto, as ferramentas mencionadas trabalham somente com o objeto digital em si (não considerando software ou possíveis interações entre uma agregação, como links e anotações, por exemplo).

Dos vários modelos disponíveis, o Arcabouço 5S foi escolhido devido a sua flexibilidade para descrever a estrutura interna e componentes do CO, a utilização dos serviços em termos de cenários, coleções, além de conceitos específicos relacionados a recuperação de imagens por conteúdo.

2.3 Softwares de Bibliotecas Digitais

A comunidade de *software livre* promoveu vários softwares relacionados à bibliotecas digitais, fornecendo o armazenamento e o acesso às coleções de todos os tipos de documentos, livros, fotografias, imagens de jornais, metadados, áudio (MP3) e vídeo, bem como coleções mistas. Entretanto, mesmo permitindo a utilização de um banco de dados ou a integração com o protocolo OAI-PMH, estes softwares ainda não permitem o tratamento de particularidades dos obje-

tos digitais em um CO (como direitos autorais), ou serviços mais complexos, como anotações e busca multimodal. Este é um dos principais desafios relacionados a CO [8, 104]. Detalhamos a seguir alguns dos softwares relacionados a bibliotecas digitais, sob ótica do gerenciamento de COs.

2.3.1 Fedora

O *Flexible Extensible Digital Object and Repository Architecture* (Fedora) [10, 22, 64, 121] ² é uma aplicação *open source*, que utiliza APIs públicas como serviços web. A arquitetura do Fedora é baseada em modelos de objetos que por definição são unidades de conteúdo (*data objects*), que podem incluir recursos digitais, metadados, e links para softwares e serviços. O Fedora utiliza o design “*compound digital object*” (ilustrado na Figura 2.4), responsável pela agregação de um ou mais itens em um mesmo objeto digital (conceito bastante similar à agregação disponível no DCC). O Fedora também trabalha com a publicação de objetos através do protocolo OAI-PMH (utilizando o *Fedora OAI Provider Service* ³ ou o PROAI⁴. Os componentes básicos do objeto digital do Fedora são:

- PID: um identificador único e persistente para o objeto.
- propriedades do objeto: um conjunto de propriedades definidas pelo sistema que são utilizadas para o gerenciamento e localização do objeto no repositório.
- *Datastream(s)*: o elemento do objeto digital Fedora que representa o conteúdo de um item.

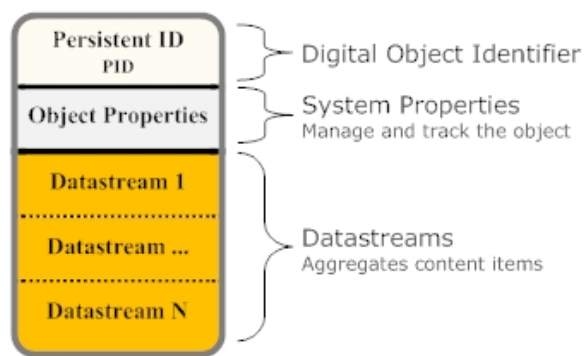


Figura 2.4: O modelo do objeto digital Fedora [22].

²<http://www.fedora-commons.org/> (último acesso em 05/05/11).

³<https://wiki.duraspace.org/display/FCSVCS/OAI+Provider+Service+1.2.1> (último acesso em 05/05/11).

⁴<http://proai.sourceforge.net/> (último acesso em 05/05/11).

2.3.2 Eprints

Eprints [42, 43] é uma plataforma para repositórios, particularmente desenvolvida para documentos provenientes de pesquisa acadêmica. O software eprints utiliza MySQL como base de dados e é utilizado pelo *Functional Requirements for Bibliographic Records* (FRBR), como ilustra a Figura 2.5. Os modelos FRBR utilizam quatro entidades principais: *work*, *expression*, *manifestation* e *item*:

- *work* é uma entidade abstrata, que representa uma criação artística ou intelectual;
- *expression* é a realização artística ou intelectual de um item *work*;
- *manifestation* é a entidade física de um item *expression*;
- *item* é um exemplar singular de *manifestation*.

O *Eprints*⁵ é um software *open source*, e também é compatível com o protocolo OAI-PMH. Para a associação de vários arquivos a um objeto digital, o Eprints habilita 2 plugins (associados ao *Metadata Encoding Transmission Standard* - METS e *MPEG-21 Digital Item Declaration Language* DIDL), porém ainda há falta de suporte para serviços mais complexos (como a busca multimodal).

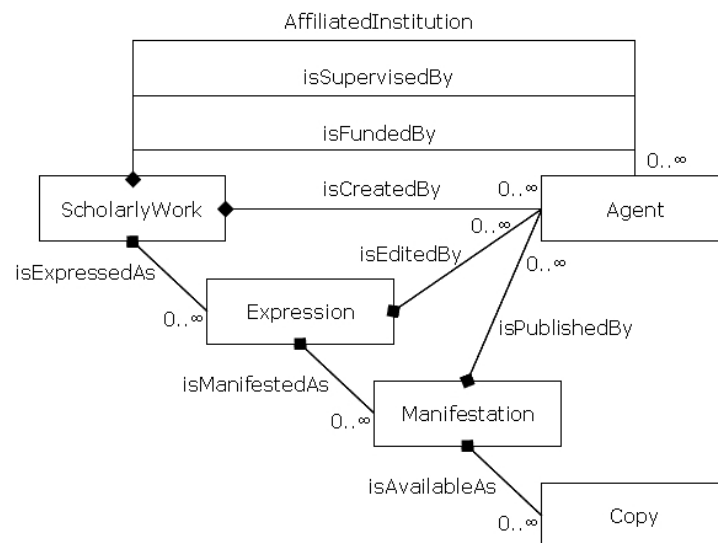


Figura 2.5: O modelo de dados de *Eprints* [45].

⁵<http://www.eprints.org/software/> (último acesso em 05/05/11)

2.3.3 CDS/ISIS

O CDS/ISIS [90] foi inicialmente baseado na versão mainframe do software CDS/ISIS. Iniciado na década de 60 pela UNESCO, foi desenvolvido especialmente para países em desenvolvimento. A definição dos dados do CDS/ISIS consiste nos seguintes componentes, armazenados em arquivos distintos:

- *Field Definition Table* (FDT) : O FDT define os campos e as características que estarão presentes nos registros da base de dados.
- *Data entry worksheet(s)* (FMT) : Um ou mais layouts de tela utilizados para criar e/ou atualizar os registros principais da base de dados. O CDS/ISIS provê um editor especial para a criação destas telas.
- *Display format(s)* (PFT): Define os requerimentos de formatação para a navegação ou busca online dos registros. CDS/ISIS provê uma linguagem de programação poderosa que permite expressar o conteúdo de um registro de várias maneiras.
- *Field Select Table(s)* (FST) : Define os campos da base de dados que serão utilizados para a busca. FSTs adicionais definem as requisições mais utilizadas na ordenação da base de dados.

CDS/ISIS está disponível online ⁶, e apesar de não possuir alguma estrutura específica para a agregação de dados, é possível integrá-lo com o protocolo OAI-PMH [48].

2.3.4 Greenstone

O Greenstone [89, 130, 131] é um software utilizado para a construção e distribuição de coleções em bibliotecas digitais. O Greenstone é um software *open source* e multilíngue, produzido pelo projeto de Biblioteca Digital da Nova Zelândia, na Universidade de Waikato, com cooperação da UNESCO e *Human Info NGO*.

O Greenstone ⁷ utiliza o *GNU database management system* (GDBM) para armazenar e recuperar dados. O GDBM pode ser instalado em Windows, Linux e Macintosh. Possui plugins para XML, Refer, MARC, OAI, CDS/ISIS, METS (subset), ProCite, DSpace, e BibTex, entre outros. A Figura 2.7 lista alguns plugins disponíveis para o Greenstone. O Greenstone já está integrado com o protocolo OAI-PMH, e apesar de não tratar diretamente COs, é possível utilizar o plugin DBPlugin (*Database Plugin*) com arquivos XML.

⁶<http://www.unesco.org/webworld/isis/isis.htm> (último acesso em 05/05/11).

⁷<http://www.greenstone.org/> (último acesso em 05/05/11).

TEXTPlug Plain text.
HTMLPlug HTML, replacing hyperlinks appropriately.
WordPlug Microsoft Word documents.
PDFPlug PDF documents.
PSPlug PostScript documents.
EMAILPlug E-mail messages, recognizing author, subject, date, etc.
BibTexPlug Bibliography files in *BibTex* format.
ReferPlug Bibliography files in *refer* format.
SRCPlug Source code files.
ImagePlug Image files for creating a library of images.
SplitPlug Splits a document file into parts.
ZIPPlug Uncompresses files.
BookPlug Specially marked-up HTML.
GBPlug Project Gutenberg E-text.
TCCPlug E-mail documents from Computists' Weekly.
PrePlug HTML output from the PRESCRIPT program.

Figura 2.6: Exemplo de plugins disponíveis no from Greenstone [131].

2.3.5 Harvest

O *Harvest* [41] é um sistema integrado que coleta, extrai, organiza e busca informações, tornando-as disponíveis através de uma interface web. Harvest possui um módulo que extrai dados de documentos fonte utilizando um *flex parser* para a comparação de padrões. A extração customizada é, em essência, a escrita de um novo conjunto de regras para o *flex parser*.

O *Harvest*⁸ utiliza a licença *GPL*, e apesar de trabalhar com a indexação, coleta e publicação de dados, não fornece suporte para a agregação de informações. Não é possível restringir a busca para campos de metadados específicos.

2.3.6 Dspace

O DSpace [101, 103, 122] foi criado pelo laboratório *Hewlett-Packard*, com a cooperação do MIT (*Massachusetts Institute of Technology*). O Dspace é utilizado para a construção, organização, distribuição e preservação de repositórios digitais.

Dspace utiliza a base de dados postgresql. Os dados são organizados conforme o sistema do Dspace (Figura 2.7). Cada site Dspace é dividido em comunidades, que podem ser posteriormente divididas em sub-comunidades, refletindo a típica estrutura universitária de instituto, departamento, centro de pesquisa ou laboratório.

Comunidades contêm coleções, que são grupos de um conteúdo específico. Uma coleção pode aparecer em mais de uma comunidade.

Cada coleção é composta de itens, que são os elementos básicos de um repositório. Cada

⁸<http://harvest.sourceforge.net/> (último acesso em 05/05/11).

item pertence a uma coleção. Este mesmo item pode aparecer em outras coleções, porém pertence a uma coleção única.

Itens podem ser subdivididos em *bitstreams*. *Bitstreams* são *streams* de bits. *Bitstreams* que são relacionadas, como arquivos HTML e imagens que compõem um mesmo documento são organizados em *bundles*.

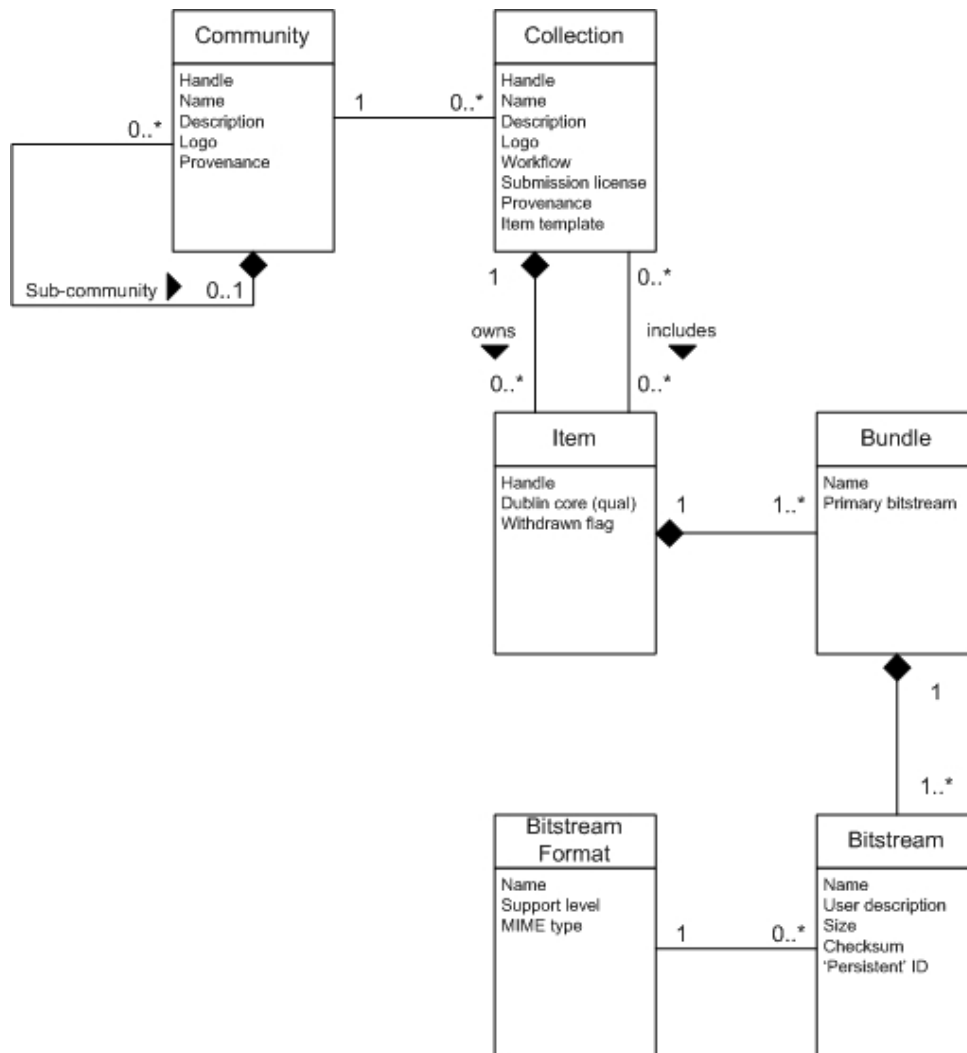


Figura 2.7: O modelo de dados do Dspace [101].

O Dspace já está integrado com o protocolo OAI-PMH, porém não tem suporte direto a objetos complexos e versões (mais detalhes em [104]).

2.3.7 Koha

Koha [27, 128]⁹ é um software *open-source* do *Integrated Library System* (ILS), utilizado pela comunidade de bibliotecários. O Koha inclui desde módulos básicos até avançados, como circulação, catálogo, aquisição, reserva, entre outros. Koha utiliza MySQL como base de dados, e atua somente como provedor de dados com o protocolo OAI-PMH (não pode coletar dados de outros repositórios).

Koha não possui suporte direto a objetos complexos, tratando somente de itens individuais do catálogo, sua estrutura e respectivos metadados.

2.4 Objetos Complexos

Existem modos diferentes de estruturar objetos [9]: atômico, composto e complexo. A abordagem atômica é quando o usuário tem um único arquivo (seja feita a partir de um arquivo de texto simples ou múltiplo) em um formato preferido. A abordagem composta é formada a partir de arquivos de conteúdo, que pode ser de diferentes formatos. Na terceira forma - objeto complexo, o objeto é descrito através de uma rede de objetos digitais dentro do repositório. Da perspectiva do projeto Fedora [9], o objeto complexo tem um objeto de agregação principal que está relacionado à múltiplas partes (também objetos digitais), onde cada objeto tem um fluxo de dados de conteúdo único.

De acordo com Krafft et al., *COs* são entidades individuais que são compostos de vários objetos digitais, cada uma dos quais é uma entidade em si [62]. Cheung et al. definiram o *CO* como o encapsulamento de vários conjuntos de dados e os recursos gerados ou utilizados durante um experimento científico, ou processo de descoberta, dentro de uma única unidade, para a publicação e intercâmbio [17]. Em outras palavras, um *CO* é uma coleção de objetos, que podem ser agrupados e manipulados como um objeto único.

Os *COs* também foram definidos como agregações de unidades de informação distintas que, quando combinadas, formam um único conceito lógico [65]. Santanchè, em contrapartida, utilizou o conceito de *COs* na área de reuso de software [107]. Tal como o conceito de *script* [114], ou o conceito de *frames* [73], os componentes de um *CO* supostamente têm o mesmo comportamento, respeitando as mesmas regras, ou representando o mesmo conceito.

Uma variedade de perspectivas e os parâmetros têm sido apresentadas em diferentes partes da literatura, para explorar os objetos complexos e agregações:

- ontologias: Gerber et al. [3] especificaram por exemplo, uma ontologia para o encapsulamento de recursos digitais e registros bibliográficos;

⁹<http://www.koha.org> (último acesso em 05/05/11).

- granularidade: Fonseca et al. mencionam a navegação vertical, onde o acesso a uma classe imediatamente superior ou inferior implica em uma mudança de nível de detalhe [28];
- estabelecimento de padrões: Candela et al. exploraram a integralidade do *CO* (medir se um *CO* tem um conjunto mínimo de elementos) [14]. Se considerarmos os padrões para agregações, outros parâmetros ainda podem ser incluídos, como o número de componentes, tipos de composições aceitáveis, ou o mínimo / máximo número de elementos que a composição deveria ter;
- prioridade sobre os demais componentes: Candela et al. também examinaram a prioridade de um componente em relação ao conjunto completo, tal que, se esse componente é copiado ou excluído, os outros elementos também são copiados ou apagados junto com ele [14];
- portabilidade para a estrutura do *CO*: Park et al. exploraram a adaptação da estrutura de *CO* para domínios diferentes, tais como dispositivos portáteis, onde alguns componentes (como vídeos) teriam tratamento diferenciado [93];
- acesso aos componentes: Manghi et al. mencionaram os diferentes perfis de acesso para componentes distintos, como sugerido nos serviços de autenticação e autorização [70];
- reutilização e preservação digital: Rehberger et al. examinaram o papel secundário que os repositórios podem desempenhar na preservação e acesso de materiais digitais e patrimônios históricos [102];
- outros: rastreamento de proveniência [75], e variação temporal [40].

COs também têm sido utilizados na preservação e publicação de metadados [71, 105], para reusar objetos existentes com o objetivo de criar novos [108], ou mesmo para o agrupamento de informações que respeitam as mesmas permissões ou operações. Dependendo da agregação, diferentes camadas podem ser expostas, com granularidades diferentes, ou tipo de mídia distintos, por exemplo. Em especial, quatro características do *CO* serão tratadas nesta tese: a identidade, os componentes, a estrutura e a delimitação do objeto.

2.5 Tecnologias Relacionadas a Objetos Complexos

Vários formatos de *COs* surgem de diferentes comunidades [85, 86]. Em computação científica, existem alguns padrões, como *Network Common Data Form*¹⁰ (NetCDF), *Hierarchical Data*

¹⁰<http://www.unidata.ucar.edu/software/netcdf/> (último acesso em 05/05/11).

*Format*¹¹(HDF), *Common Data Format*¹² (CDF) e *Extensible File System* (ELFS) [50]. HDF, CDF e NetCDF, por exemplo, são utilizados no armazenamento e recuperação multi-dimensional, enquanto ELFS é uma abordagem para resolver o problema de alto desempenho de operações de entrada e saída utilizando o tratamento arquivos como objetos tipados.

COs são frequentemente encontrados no armazenamento de banco de dados persistentes. Eles podem ser representadas usando padrões como o MPEG-21 [11] ou METS [18]. Outras tecnologias foram propostas, como o padrão *Moving Picture Experts Group* (MPEG) [99], o arcabouço multimídia MPEG-21 [11] e formatos de objetos digitais como o *Moving Picture Experts Group - 21 Digital Item Declaration Language* (MPEG-21 DIDL) [46] e o *Metadata Encoding and Transmission Standard* (METS) [23].

Embora haja uma série de normas auxiliando a gestão das COs, ainda há incompatibilidade, motivando soluções de integração e interoperabilidade. Como cada padrão é especializado para um determinado domínio, é difícil para interoperar em contextos diferentes. No entanto, é possível combinar algumas delas, como proposto em [25], com um estudo comparativo do *IMS Content Package* (IMS CP) [12] e *Reusable Asset Specification*(RAS) [91].

Novos padrões surgiram, como o *SQL Multimedia and Application Packages* (SQL/MM) [72]. Ele foi definido para descrever o armazenamento e manipulação para objetos complexos. Um certo número de candidatos de domínios multimídia foram sugeridos, incluindo textos, dados geográficos, imagens e outros.

O *Open Archives Initiative* (OAI) [63] é um arcabouço para repositórios institucionais com conteúdo digital (bibliotecas digitais). A infraestrutura técnica do OAI, especificada no *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH) [119], define um mecanismo para publicar metadados. Este protocolo determina que metadados de arquivos sejam mapeados para o padrão *Dublin Core*¹³.

O METS [23] trata da agregação e coleta de metadados de recursos digitais para a submissão em um repositório. É uma iniciativa da *Digital Library Federation*. Um documento METS consiste das seguintes seções: cabeçalho, metadados descritivos, metadados administrativos, seção para o arquivo, mapa estrutural, links estruturais e comportamento. METS utiliza o mapa estrutural como uma estrutura hierárquica para o objeto digital, onde arquivos podem ser agrupados em uma estrutura $\langle fileGrp \rangle$. Esta estrutura é utilizada para delimitar uma única versão eletrônica do objeto digital. A estrutura $\langle FContent \rangle$ foi criada para encapsular o conteúdo real do arquivo dentro de um documento METS, porém é raramente utilizada. METS provê um esquema XML, com o seguinte propósito:

- Criar instâncias do documento XML que expressam a hierarquia estrutural do objeto digital;

¹¹<http://www.hdfgroup.org/> (último acesso em 05/05/11).

¹²<http://cdf.gsfc.nasa.gov/> (último acesso em 05/05/11).

¹³<http://dublincore.org/> (último acesso em 05/05/11).

- Armazenar os nomes e localizações dos arquivos que compõem estes objetos; e
- Armazenar os metadados associados.

O SCORM [2] é uma compilação de especificações técnicas para habilitar a interoperabilidade, acessibilidade e reusabilidade de conteúdo de aprendizado baseado em web. Com um modelo de agregação de conteúdo, os recursos são descritos em um arquivo *imsmanifest.xml*, organizado em arquivos de esquema/definição (.xsd and .dtd), e armazenados em um arquivo zip como um encapsulamento de conteúdo. SCORM define um model de agregação e um ambiente *run-time* para conteúdo de aprendizado baseado em web. No SCORM, o objeto de conteúdo é uma unidade de aprendizado que pode ser transportado via web. Geralmente um objeto de conteúdo é uma simples página HTML ou um documento que pode ser visualizado via um web browser. Um objeto de conteúdo é o menor nível de granularidade dos recursos de aprendizagem, e pode disponibilizar das mesmas tecnologias que uma página web utiliza (como Flash, JavaScript, frames, e imagens).

O MPEG-21 [11] tem por objetivo definir um framework aberto para aplicações multimídia, apoiar a declaração e identificação dos objetos, manutenção de direitos legais digitais, e adaptação. O MPEG-21 é baseado em dois conceitos essenciais: a definição de uma unidade fundamental de distribuição e transação (item digital), e o conceito de usuários interagindo com ele. Itens são agrupados em um container estrutural utilizando o *Digital Item Declaration Language* (DIDL).

A Tabela 5.1 sumariza conceitos básicos de objetos complexos no METS, SCORM e MPEG-21: suas unidades básicas, e o relacionamento, identificação e estruturação de seus componentes.

Nome	Unidade	Componentes Internos	Identificador	Estrutura
METS	Objeto simples	Estrutura <i>FContent</i>	OBJID	Mapa Estrutural
SCORM	<i>Asset</i>	Regras de seqüência	—	Arquivos de esquema/definição
MPEG-21	Item	<i>anchors</i> e <i>fragments</i>	URI	XML-DIDL

Tabela 2.1: Conceitos básicos de CO da perspectiva do METS, SCORM e MPEG-21.

Em particular, três tecnologias foram selecionadas para a análise nesta tese (*Digital Content Component*, Buckets e OAI-ORE), e serão detalhadas a seguir.

2.5.1 Digital Content Component

O *Digital Content Component* (DCC) [110] se refere a um modelo e a tecnologia que o implementa. Funcionando como uma cápsula, tem uma estrutura interna organizada como um objeto complexo e uma estrutura externa construída como um componente de software. O CO interno tem uma estrutura de confinamento hierárquica, capaz de armazenar e relacionar qualquer tipo de conteúdo digital, incluindo software executável. Essa estrutura é baseada em

uma síntese e generalização de padrões relacionados *CO: Reusable Asset Specification* – RAS, MPEG-21 [11], METS [23], e IMS CP [109].

Externamente, qualquer DCC atua como um componente de software, declarando uma interface. Quando um DCC encapsula software executável (DCC de Processo) ele se comporta como um componente de software. No entanto, um DCC também pode encapsular outros tipos de conteúdo, sem software executável incluído (DCC Passivo). Neste caso, ele declara um tipo especial de interface potencial. É o potencial no sentido de que, embora as operações declaradas não são implementadas, poderiam ser aplicados ao conteúdo - por exemplo, um DCC contendo um vídeo pode declarar as operações de tocar, parar, pausar, etc. Tal como os objetos digitais complexos, um DCC agrega e encapsula um ou mais artefatos digitais e representa internamente as suas relações; tal como os componentes de software, um DCC esconde detalhes internos da sua representação e expõe uma interface pública.

O DCC já foi utilizado na gestão de dados de sensores para aplicações científicas [49] (manipulando a gestão de acesso, operações de processamento, e metadados), e em uma ferramenta de autoria baseada em componentes [111]. Quanto às vantagens da DCC sobre outras tecnologias, podemos citar o uso de ontologias, o encapsulamento de software, e a descrição de uma interface para manipular as informações.

A DCC é composto de quatro subdivisões distintas (Figura 2.8):

- **(a) conteúdo:** o conteúdo em si - os dados ou código em seu formato original, ou outro DCC;
- **(b) estrutura:** a declaração de uma estrutura de gestão que define como os componentes dentro de um DCC se relacionam entre si, em XML;
- **(c) interface:** especificação da interface do DCC usando padrões abertos para a descrição da interface - WSDL e OWL-S (semântica);
- **(d) metadados:** metadados para descrever versão, a funcionalidade, aplicabilidade, e restrições de uso - usando OWL.

Nesta tese a estrutura XML do DCC foi simplificada para representar a agregação dos componentes do ICO. O DCC é utilizado nesta tese devido à vantagem de gerir de forma homogênea fontes de dados e software, juntamente com a habilidade de construir recursivamente componentes e serviços (facilitando a implementação de serviços como a busca multimodal). Estas propriedades serão exploradas nesta tese, para combinar o acesso, tratamento e implementação de serviços de busca de imagens por conteúdo para COs.

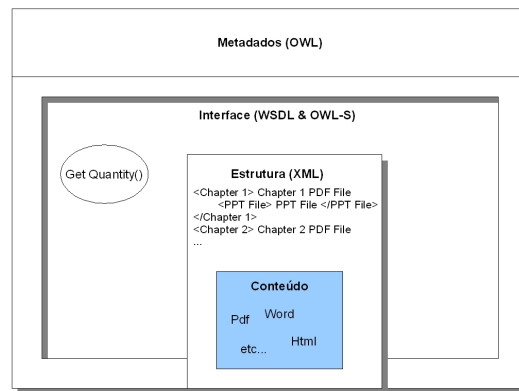


Figura 2.8: A estrutura do *Digital Content Component*.

O DCC não estabelece métodos públicos específicos para o acesso aos dados. No entanto, na versão Java, uma biblioteca Jar chamada FluidWeb é sugerida [109], como ilustra a Figura 2.9. Santanchè chamou de “Fluid Web” a perspectiva dinâmica da web, onde o conteúdo digital pode viajar e ser replicado, adaptado, decomposto, fundido e transformado. A implementação da biblioteca FluidWeb é composta por seis classes: (i) Anotação - é a parte responsável pela anotação do componente; (ii) Componente - é a base utilizada para criar cada DCC, (iii) Conector - é usado para conectar DCCs, (iv) Fábrica - é a parte responsável pela criação do DCC, (v) Mensagem - é responsável pela troca de mensagens entre os DCCs; e (vi) Contexto - é responsável por definir o contexto do componente. O DCC não usa um formato específico para o armazenamento dos dados e execução dos componentes, uma vez que não impacta a interoperabilidade.

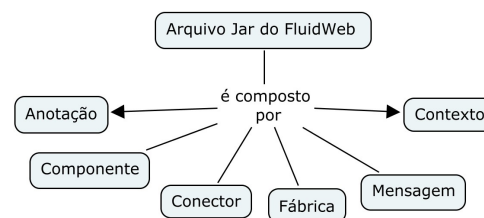


Figura 2.9: A implementação do arquivo Jar Fluid Web e seus principais blocos.

2.5.2 Buckets

Buckets [69, 87, 88] trabalha com a construção de um container independente de repositórios, no qual todos os objetos e tipos de dados estão relacionados de forma semântica e/ou sintática,

podendo ser agrupados, arquivados, e manipulados juntos de forma lógica, como um objeto único. Buckets são objetos ativos de repositórios, e podem comunicar-se entre si, em forma de rede, ou em forma de serviços arbitrários. Buckets se baseiam no padrão de recursos da *World Wide Web* (WWW) para funcionar, sendo gerido por duas ferramentas. Uma delas é o *author tool*, que permite ao autor a construção de um bucket sem conhecimentos de programação. A segunda é a ferramenta *management tool*, que fornece uma interface para permitir aos gestores de site a definição das configurações padrão para o mesmo.

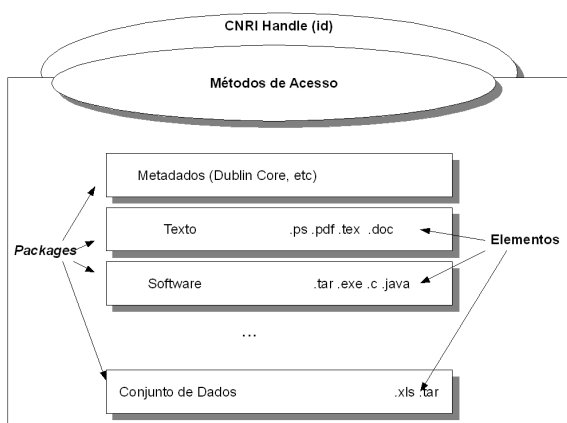


Figura 2.10: A Arquitetura de Buckets (adaptado de [69]).

O protótipo de Buckets é escrito em Perl 5, e faz uso do HTTP como um protocolo de transporte, armazenando os metadados na RFC 1807 (um formato para registros bibliográficos), com todos os arquivos reunidos em um único diretório Unix. Um bucket segue o mesmo modelo e tem todos os arquivos relevantes agregados juntos utilizando diretórios para representar a semântica. Um único bucket pode ter vários pacotes. Um pacote único pode ter vários elementos, como mostra a Figura 2.10 (adaptada de [69]).

Buckets têm seus métodos públicos (mostrados na Tabela 2.2), porém métodos alternativos podem ser adicionados através de sua interface (utilizando a função *Add_method*).

2.5.3 OAI-ORE

O *Open Archives Initiative Object Reuse and Exchange* (OAI-ORE) [65, 67, 105, 123] define normas para a descrição e intercâmbio de agregações de recursos da Web, e desenvolve mecanismos para interoperabilidade para expressar COs na web. Com o OAI-ORE é possível reconstituir os limites lógicos de objetos complexos, as relações entre os seus componentes

Tabela 2.2: Principais métodos públicos implementados por Buckets [69].

Método	Descrição
Metadata	retorna o metadado
Display	mostra o conteúdo do bucket
Id	retorna o id do bucket
List_methods	lista todos os métodos públicos reconhecido por um bucket
List_owners	lista os usuários com acesso de administrador
Add_package	adiciona um package a um bucket existente
Delete_package	remove um package de um bucket existente
Add_element	adiciona um elemento a um package existente
Delete_element	remove um elemento de um package existente
Add_method	adiciona um método a um bucket existente
Delete_method	remove um método de um bucket

internos, e suas relações com os outros recursos. A informação é encapsulada utilizando *named graphs*: um conjunto de asserções RDF identificado por uma URI. A Figura 2.11 ilustra o mapeamento dos principais conceitos do Arcabouço 5S e OAI-ORE [54]. Note que conceitos como recurso, objeto digital, e objeto complexo podem ser mutuamente mapeados, além de permitir o foco em outros aspectos importantes como a composição, identificação e unidade de encapsulamento. Um *named graph* pode ser descrito por um mapa de recursos.

O OAI-ORE não possui métodos públicos (como os apresentados em uma interface de acesso), mas é congruente com a arquitetura Web [65], constituída essencialmente por:

- URIs para identificar objetos;
- recursos - itens que serão mapeados;
- protocolo padrão, como HTTP, habilitando o acesso aos recursos;
- links via referências a URIs;
- *named graphs* para encapsular a informação em objeto composto.

O OAI-ORE já foi utilizado, por exemplo, como uma extensão que permite a aceleração do processo de importação dos dados, apresentando interfaces a repositórios para o armazenamento remoto, onde o conteúdo é representado e acessado através de um *resource map* do OAI-ORE. Outro exemplo do uso do protocolo OAI-ORE é o *Literature Object Re-use and Exchange* (LORE) [3], uma ferramenta simples, projetada para permitir que acadêmicos e professores de literatura possam gerir, editar e publicar objetos.

Tarrant et al. [123] modelaram um repositório simplificado de três fases (adaptado na Figura 2.12), mostrando exemplos de interfaces, ou plug-ins, que podem ser utilizados em cada

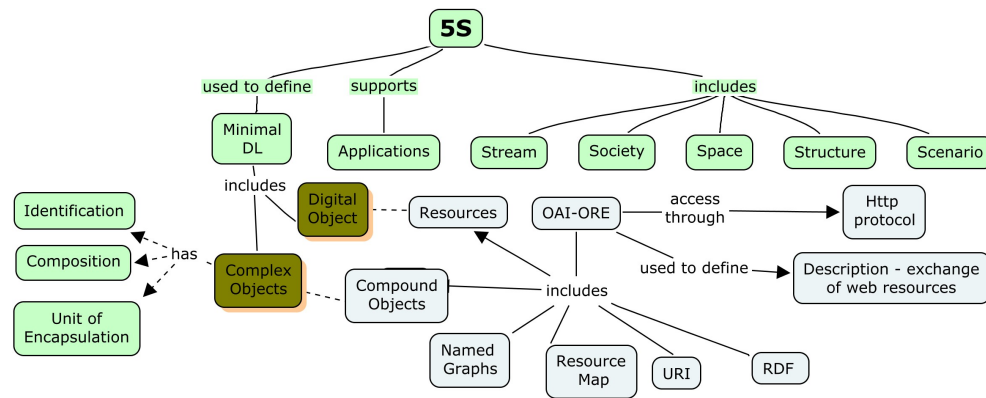


Figura 2.11: Mapeando conceitos no Arcabouço 5S e OAI-ORE [54].

fase, onde o protocolo OAI-ORE é utilizado para obter e publicar conteúdo. O foco principal é sobre os objetos e os relacionamentos entre esses objetos. Em [123] por exemplo, o objeto de publicação tem duas partes, a própria publicação (PDF) e um registro XML que representa os metadados armazenados. O modelo de dados OAI-ORE também trabalha com metadados específicos, como por exemplo, a identidade do agente de criação do *resource map*, a última modificação do *resource map*, etc.

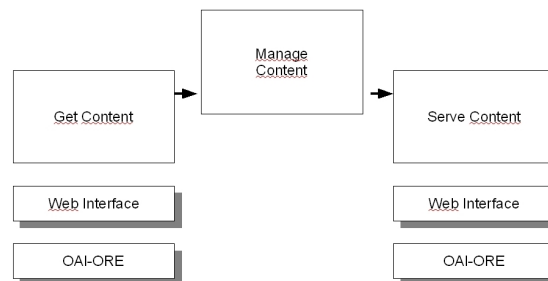


Figura 2.12: Exemplo de uso do OAI-ORE [123].

O OAI-ORE também foi usado como base do *D-Net Software Toolkit* [70], no qual foi chamado de *Aggregative Digital Library Systems* (ADLSs). O objetivo era construir ADLSs seguindo uma abordagem “Lego”, isto é, pelos dados implantação, serviços e áreas de funcionalidade e combinando-os em fluxos de trabalho, baseada nas exigências de aplicação.

Apesar dos exemplos citados nesta seção, o OAI-ORE ainda possui poucas aplicações online de fácil acesso que permitam a integração com outras aplicações e a visualização dos componentes do CO.

2.6 Recuperação de Imagens por Conteúdo

Em geral, duas abordagens diferentes têm sido utilizadas para permitir a busca em coleções de imagens: uma baseada em metadado textual da imagem e outra com base no conteúdo da imagem. A segunda abordagem, chamada Recuperação por Conteúdo (*Content-Based Image Retrieval* - CBIR) [20] propõe a recuperação automática de imagens, com base nas propriedades visuais, tais como textura, forma ou cor. Nesse sentido, imagens são usadas para procurar outras imagens em sistemas de CBIR [19, 129]. Esta solução exige a construção de descritores de imagens, que são caracterizadas por (i) um algoritmo de extração (descritor) para codificar características da imagem em vetores de características, e (ii) uma medida de similaridade para comparar duas imagens com base na distância entre os vetores de características. A medida de similaridade é uma função correspondente (como por exemplo, a distância euclidiana), que fornece o grau de semelhança para um dado par de imagens representadas por vetores de suas característica. Quanto maior o valor da distância, menor a semelhança das imagens.

Existem várias aplicações que suportam serviços baseados em conteúdo da imagem, permitindo a integração em domínios distintos [1, 84]. Um exemplo é a coleção digital do projeto *Crisis, Tragedy, and Recovery Network* (CTRnet) [53]. O CTRnet tem como objetivo desenvolver melhores abordagens para tornar a tecnologia útil em situações de crise e tragédia, além de apoiar a análise de informações para salvamento, socorro e recuperação, da perspectiva de uma biblioteca digital (DL). O CTRnet tem diversos módulos, incluindo o rastreamento de dados, a filtragem de informações, um aplicativo do Facebook, visualização para usuários, busca através de metadados e recuperação de imagens por conteúdo. O módulo CBIR se baseia no teste e avaliação de descritores diferentes.

Em um contexto diferente, imagens também podem ser utilizadas como base para realizar a busca de impressões digitais [59]. Neste caso, os algoritmos utilizados para processamento de imagens não são apenas relacionadas com a implementação de mecanismos de CBIR, mas também utilizando softwares específicos para representar adequadamente os detalhes de impressões digitais, tais como direção, o fluxo, o contraste e a curvatura. Junto com esses parâmetros, detalhes específicos (minúcias) são analisados, para comparar se duas imagens diferentes (de registros policiais e cena do crime, por exemplo) pertencem à mesma pessoa.

Vários sistemas de CBIR têm sido propostos [19], como QBIC, Chabot e Photobook. Alguns tornaram-se produtos comerciais, outros foram propostos como protótipos de investigação [60], sendo desenvolvidos em universidades e laboratórios de pesquisa. Existem ainda outros projetos que comparam descritores CBIR para melhorar o desempenho [98], para visualizar os resultados [126], combinar o conteúdo visual de imagens com metadados [21] ou mesmo explorar diferentes domínios [26, 80, 129]. Nenhuma dessas iniciativas porém, têm-se centrado em mecanismos para gerenciar objetos complexos de imagem.

Embora estes sistemas CBIR fornecem um conteúdo adicional útil, ainda encontram-se dis-

persos por fontes heterogêneas, e às vezes precisam de integração, para evitar o chamado “gap” semântico entre as características extraídas dos sistemas e a consulta do usuário. A causa da diferença semântica é a incapacidade de representar semântica real a partir do uso de descritores de propriedades visuais de baixo nível. Para resolver isso, alguns sistemas usam os textos ao redor da imagem [133], anotações [80], marcas d’água [66], metadados [113], o uso de vários descritores [98], modelos de comportamento [37] ou técnicas multimodais [127, 132]. No caso das técnicas multimodais, as aplicações podem tirar proveito de todas as modalidades disponíveis. Em especial, se considerarmos imagens, a abordagem textual pode ser considerada complementar, e a abordagem multimodal permite o equilíbrio entre os prós e contras de cada abordagem [13].

Existem ainda técnicas avançadas para recuperação multimodal utilizando métodos de pós-processamento, tais como (i) a re-classificação de imagens, tendo em conta a informação contextual [97] e (ii) a correlação de suas distâncias para a re-classificação [96].

2.7 Conclusões

Este capítulo apresentou a revisão de literatura relacionada aos principais temas desta tese. Os conceitos básicos abrangem várias áreas: formalismos matemáticos (Arcabouço 5S, Delos, etc.), tecnologias de aplicações (DCC, Buckets, OAI-ORE, Scorm, MPEG-21), conceitos básicos (CO, CBIR, DL, etc.) e aplicações relevantes na literatura que utilizam estes conceitos. Em especial, da área de formalismo, nos focaremos no Arcabouço 5S. Da área de tecnologias, nos focaremos em DCC, Buckets e OAI-ORE. E dentre os vários conceitos apresentados, nos focaremos em DLs, serviços de DLs, CO, CBIR e ICO.

Capítulo 3

Estudos de Caso Preliminares e Aplicações

Neste capítulo descrevemos três aplicações que serão utilizadas ao longo desta tese para compreender e exemplificar cenários de estudo de caso.

3.1 *Superimposed Image Description and Retrieval Tool* - SuperIDR

O projeto *Deployment and Assessment of an Image Annotation and Retrieval Tool, including for Biodiversity*, com suporte financeiro da Microsoft é uma colaboração entre o Instituto de Computação da Unicamp e a Virginia Tech, USA. Este projeto é parte do programa *Microsoft Research under the Tablet PC Technology and Higher Education*. O principal objetivo deste projeto é a especificação e implementação de uma aplicação para o suporte de busca e anotação de imagens. Para mais detalhes sobre o projeto, verifique a página do projeto ¹.

O *Superimposed Image Description and Retrieval Tool* - SuperIDR é uma ferramenta de recuperação e descrição de imagens [55, 56, 78–81], desenvolvida com o objectivo de ajudar os usuários a trabalhar com as partes das imagens “in situ”, onde é possível seleccionar, anotar e recuperar partes de imagens no contexto da a imagem original. SuperIDR foi avaliada em dois cenários diferentes. No primeiro cenário, utilizamos a *identificação das espécies de peixes* como tarefa acadêmica específica para testar o uso desta ferramenta. No segundo cenário, a ferramenta foi avaliada na Unicamp [55, 56], comparando as características morfológicas de diferentes espécies de parasitas. A comparação de espécies (Figura 3.1) é importante no domínio da parasitologia porque detalhes como formas (ovais, bastonete), fases (grávido ou maduro), número de dentes, etc. são decisivos para a sua classificação. No entanto, a ferramenta pode ser usada em qualquer tarefa que envolva imagens com um número significativo de detalhes

¹<http://si.dlib.vt.edu/> (último acesso em 05/05/11)

importantes, como a análise de pinturas da história da arte, do exame de um edifício de estilo na arquitetura, a compreensão de árvores em dendrologia, etc.

SuperIDR foi populado com detalhes de 13 espécies e 49 imagens de parasitas, geradas a partir da análise do microscópio. Além de fazer anotações, o SuperIDR permite pesquisar e navegar nas descrições de espécies, imagens, marcas, imagem e anotações. Um usuário pode pesquisar em uma de duas maneiras: 1) executar a busca textual (texto completo e pesquisa de campo, alimentado por Lucene.NET²) sobre as descrições de espécies e anotações, onde a consulta pode incluir combinações booleanas de termos, e frases; e 2) realizar pesquisa de imagens por conteúdo de imagens e marcas, onde a consulta pode ser uma imagem completa ou parcial. Finalmente, no SuperIDR, um usuário pode navegar através de informações sobre as espécies, seja através de uma organização taxonômica das espécies com base na família e gêneros (ilustrada na Figura 3.2) ou através de uma versão eletrônica da chave dicotômica. A Figura 3.3 mostra *screenshots* da ferramenta.

Especificadamente, o estudo de caso da aplicação SuperIDR citado nesta tese utilizou metadados e 76 imagens de parasitas (ilustradas na Figura 3.4), agrupadas em 17 espécies (*Ancylostoma caninum*, *Ancylostoma duodenale*, *Ancylostoma braziliense*, *Fasciola hepatica*, *Entamoeba histolytica*, *Iodamoeba butchlii*, *Dipylidium caninum*, *Entamoeba coli*, *Eurytrema pancreaticum*, *Hymenolepis nana*, *Necator americanus*, *Taenia saginata*, *Taenia solium*, *Diphyllobothrium latum*, *Echinococcus granulosus*, *Hymenolepis diminuta*, e *Trichuris trichiura*).

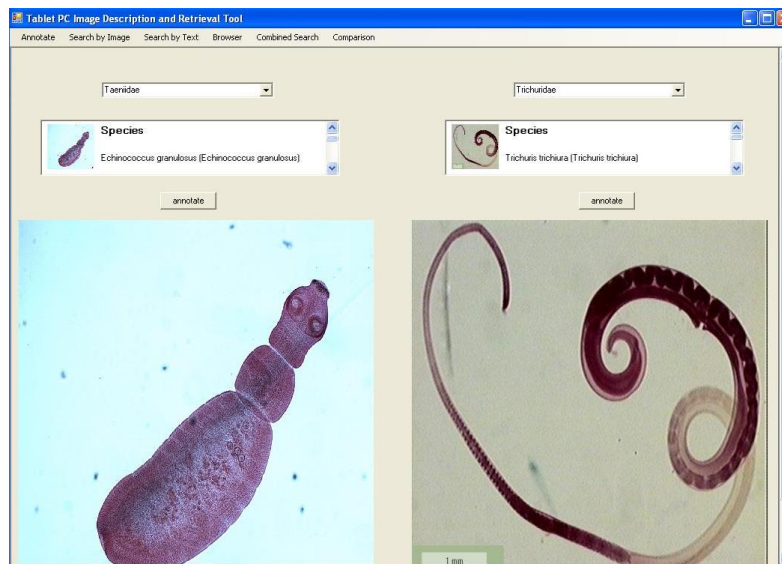


Figura 3.1: A utilização de anotações para comparar duas espécies no SuperIDR.

O objetivo dos experimentos realizados com o SuperIDR no segundo semestre de 2008 e primeiro semestre de 2009 era validar o uso do tablet PC em um novo domínio, como um

²<http://incubator.apache.org/lucene.net/> (último acesso em 05/05/11).

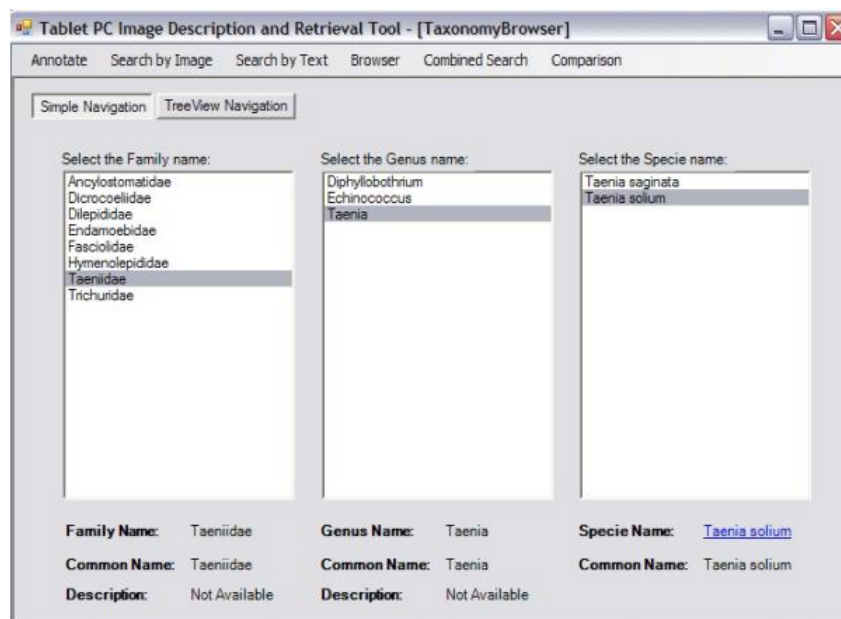


Figura 3.2: A navegação taxonômica na Ferramenta SuperIDR.

método alternativo para ensinar conceitos relacionados com imagens e anotações. Entre os benefícios da ferramenta, podemos citar:

- O uso da caneta do tablet PC para:
 1. Criar marcas em imagens.
 2. Criar anotações e associá-las com marcas e imagens.
- Criar links de marcas a outras informações multimídias tais como anotações e descrições, áudio/vídeo, outras imagens, e marcas em outros documentos multimídia. Além disso, é possível realizar a busca e navegação em marcas e informações associadas, utilizando os mecanismos de busca textual e visual.

O relacionamento de imagens com marcas, anotações, metadados e partes de imagens na ferramenta SuperIDR foram fundamentais para compreender a complexidade de agregação de informações diferenciadas, e como partes de documentos podem ser conectadas, anotadas, e acessadas. Em consequência, o serviço de busca textual, visual e multimodal (texto + CBIR) exemplificaram como os serviços individuais dos componentes do CO podem ser utilizados para realizar uma busca integrada.



Figura 3.3: A Ferramenta SuperIDR: A- Navegação, B- Comparação, C- Anotação e D - Busca Textual.

Figura 3.4: Exemplos de imagens de parasitas utilizadas no estudo de caso.

3.2 A Integração de Bibliotecas Digitais de Impressões Digitais

Impressões digitais têm sido utilizadas para a identificação de indivíduos desde 1900. O processo de identificação e comparação de impressões digitais baseia-se na análise da posição de detalhes (*minutiae*), como terminações e bifurcações de sulcos. Estes padrões são importantes uma vez que são formados no feto pela quarta mês de gravidez e não mudam até a morte. Estes padrões não podem ser alterados, exceto por acidente, mutilação ou doença de pele muito graves, já que são formados em camadas profundas da derme.

Os padrões comuns em impressões digitais - loops, espirais e arcos - são utilizados para classificar a impressão digital, pré-alinhando algoritmos de acordo com estas singularidades [47]. É um processo semelhante aos sistemas de busca baseada em conteúdo, com a análise adicional de características como qualidade, direção, ângulos, etc.

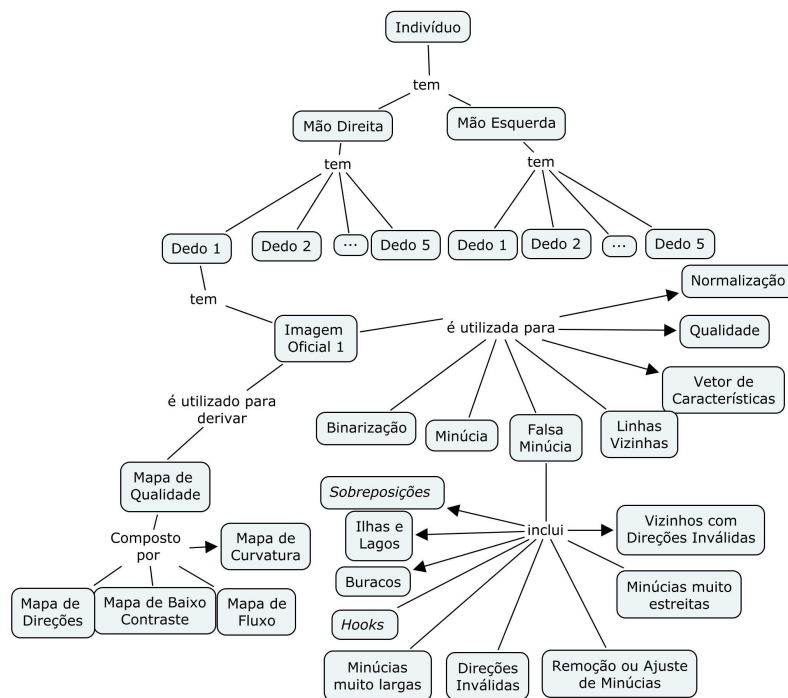


Figura 3.5: As características analisadas em uma impressão digital.

Neste contexto, considere uma DL de Impressões Digitais [58, 59, 74, 92, 116], que unifica quatro bibliotecas digitais diferentes, utilizando a perspectiva de objeto complexo. Utilizamos este exemplo para mostrar como abordagens de modelagens de banco de dados podem ser enriquecidas quando baseadas na teoria de objetos complexos, de modo a melhorar a análise de requisitos, o projeto e implementação destas bases e / ou aplicações de biblioteca digital. Para mais detalhes sobre do projeto *Fingerprint Sufficiency and Training*, verifique a página do projeto³.

O primeiro tipo de Biblioteca Digital (DL1) está associada com o armazenamento de impressões de digitais em bancos de dados. O segundo domínio está relacionado com a criação de materiais de treinamento para examinadores de impressões digitais (DL2). Um terceiro tipo de biblioteca digital se refere às provas e dados que descrevem uma cena de crime (DL3). O quarto e último tipo de DL está relacionado ao projeto NIJ (*National Institute of Justice*) de pesquisa da Virginia Tech, implementando experimentos, análise, medidas de qualidade, e métodos de pareação em bibliotecas digitais (DL4). A combinação destas quatro bibliotecas em uma biblioteca única e integrada (como mostra a Fig. 5.1), utilizando objetos complexos como um meio de navegar e agregar estas informações, resulta em uma biblioteca digital de grande porte e interessante do ponto de vista da pesquisa.

³<http://collab.dlib.vt.edu/runwiki/wiki.pl?Fingerprint> (último acesso em 05/05/11)

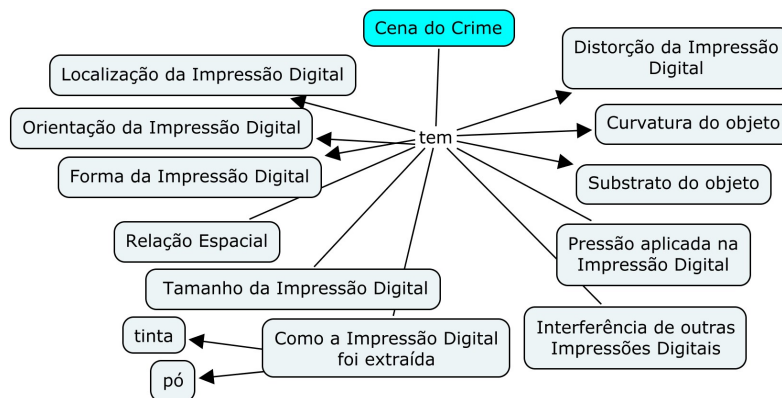


Figura 3.6: As características analisadas em impressões digitais retiradas de uma cena do crime.

Na DL1, a informação é utilizada para identificar uma pessoa (como ilustra a Figura 3.7-A). Ela gerencia bases de dados de grande porte, que podem possuir milhões de indivíduos, onde cada um possui 10 impressões para os dedos da mão, mais 10 impressões digitais para os dedos do pé, mais uma impressão digital para a palma da mão, etc. Para cada imagem, várias características são analisadas, como o mapa de qualidade e falsas minúcias (Figura 3.5). Uma das maiores bases de dados relacionadas à identificação de impressões digitais é o do *Federal Bureau of Investigation*⁴. Ele possui aproximadamente 66 milhões de indivíduos cadastrados com uma ficha criminal, além de outros 25 milhões de impressões digitais de civis.

A DL2 tem um propósito diferente: educação e treino dos usuários. Em princípio, para testar os examinadores de impressões digitais, a combinação de exemplares identificados poderiam ser utilizados para a avaliação, por isso cada caso de exame é única, reduzindo as oportunidades de reprodução ou cópia dos resultados. Os módulos de formação terão exemplos utilizados na instrução, e outros ainda para exercícios e exames, provenientes de todas as outras DLs.

Na DL3, as imagens são utilizadas para identificar ou excluir pessoas. A evidência de uma cena de crime (ilustrada na Figura 3.7-D) pode listar milhares de pessoas que visitaram um local popular, ou um objeto tocado, criando evidências que podem ser posteriormente comparadas com um registro de antecedentes criminais. Cada um dos dez dedos da mão de um indivíduo pode produzir imagens diferentes dependendo de vários parâmetros (Figura 3.6), como tipo de distorção, como por exemplo, a pressão excessiva ou o deslizamento de um dedo. Além disso, há sobreposições de várias camadas de impressões digitais, ou seja, combinações das imagens sob o mesmo substrato.

Na DL4, o foco está nos algoritmos de impressões digitais (como *Blurring*- Figura 3.7-B e distorção de pele - Figura 3.7-C), variando os parâmetros de distorção da pele e foco da imagem. Imagens distorcidas ou sintéticas são criadas por algoritmos que simulam o movimento e / ou distorção da pele. Uma única impressão digital, por exemplo, combinada com 10 parâmetros

⁴<http://www.fbi.gov/hq/cjisd/iafis.htm> (último acesso em 05/05/11).

de distorção pode gerar sinteticamente cerca de 10.000 imagens.

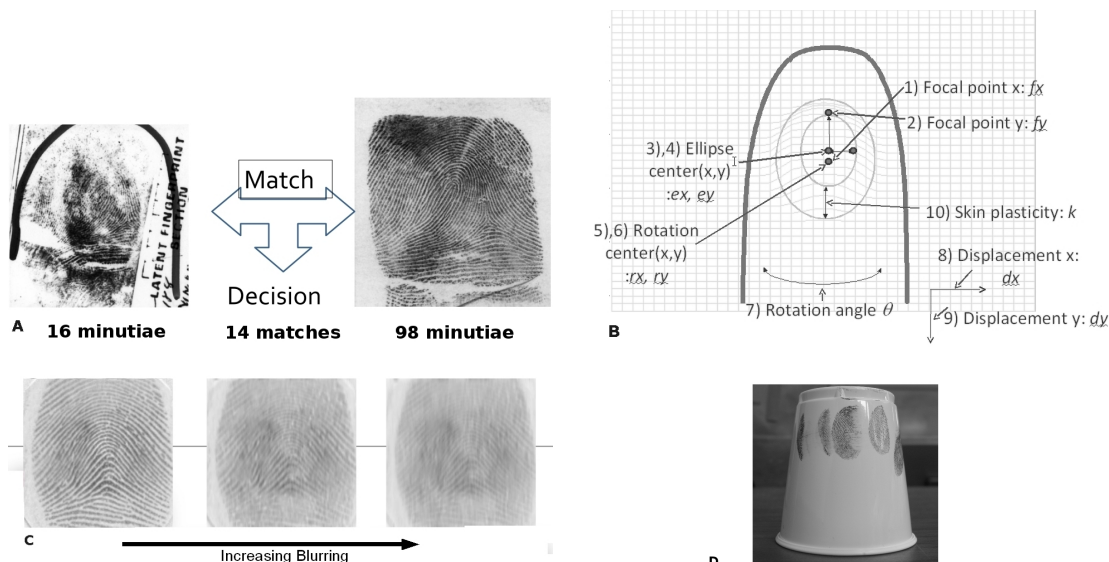


Figura 3.7: (A) Comparando Imagens. (B) Distorção de Imagem. (C) *Blurring*. (D) Cena do Crime.

Para apresentar uma noção de escala, suponha que uma imagem gera 100 imagens distorcidas. Multiplique por 25 milhões de possíveis suspeitos. Em seguida, compare com uma imagem da cena do crime, que possui 55 impressões digitais parciais. Finalmente, selecione e relacione bons exemplos para uso em um treinamento.

Com a integração, a biblioteca digital unifica quatro comunidades diferentes, permitindo que cada uma delas avalie as informações de diferentes perspectivas, e explorem o sistema como um todo, ou com o foco em uma determinada área. Além disso, podemos tirar vantagem dos serviços da biblioteca digital (por exemplo, navegação e buscas), formalismos, e soluções na área de preservação digital.

3.3 Nou-Rau: Biblioteca Digital de Teses e Dissertações da Unicamp

O sistema Nou-Rau⁵ centraliza as teses digitais de estudantes da Unicamp para a comunidade em geral. Com a utilização dos metadados principais relacionados a tese (como autor, resumo, título, ano, etc.), o usuário pode navegar no website da instituição, realizando buscas em teses, dissertações, trabalhos de conclusão de curso, eventos, hemeroteca, revistas eletrônicas,

⁵<http://libdigi.unicamp.br/> (último acesso em 05/05/11).

Biblioteca Digital da UNICAMP

Catálogo da UNICAMP | Programa de Acesso à Informação Eletrônica | Faculdades e Institutos | PRPG | SIPEX | Plataforma Lattes | +A A -A A

Bases Disponíveis

- Dissertações e Teses
- Trabalhos de Conclusão de Curso
- Eventos
- Hemeroteca - CMU
- Revistas Eletrônicas
- Produção Técnico-Científica Digital
- Coleções Raras e Especiais

Sobre a Biblioteca Digital

- Apresentação
- Instrução para autores
- Instrução para downloads

Relatórios Estatísticos

- Indicadores
- Estatística do Site
- Administração do Sistema

Pesquisa

Entrar com uma ou mais palavras:

Pesquisar: Pesquisar

Use as aspas para expressões. Exemplo: "biblioteca"

☒ Todas as palavras (AND) ☐ Qualquer uma das palavras (OR)

Pesquisar em:

- Eventos
- Hemeroteca - CMU
- Revistas Eletrônicas
- Produção Técnico-Científica Digital
- Coleções Raras e Especiais

Mostrar: 50 resultados **por página** **Formato de Apresentação:** Com resumo

Pesquisa por Índices

Pesquisar: Pesquisar

Selecionar Índice:

☒ Palavra-Chave ☐ Autor ☐ Orientador ☐ Título

☐ Área de Concentração ☐ Ano

Pesquisar em:

- Todos
- Dissertações e Teses
- Trabalhos de Conclusão de Curso
- Eventos
- Hemeroteca - CMU

Documentos

Total:	44.485
Mestrado:	22.771
Doutorado:	12.651
Livre Docência:	230
Outros:	8.833

Teses mais baixadas

- O conhecimento matemático e o uso de jogos na sala...
- (In)disciplina na escola: cenas da complexidade d...
- As bases fisiológicas da estrutura tridica da sem...
- As dificuldades de aprendizagem em leitura e escri...
- O professor de alunos com deficiência mental: co...

Teses mais visitadas

- O conhecimento matemático e o uso de jogos na sala...
- As dificuldades de aprendizagem em leitura e escri...
- O professor de alunos com deficiência mental: co...
- O brincar na pré-escola
- Análise de 152 pacientes portadores de polipos nas...

Teses com restrição

Figura 3.8: O sistema Nou-Rau (último acesso em 20/07/2011).

produções técnico-científicas digitais e coleções raras. Também é possível navegar na biblioteca digital através de índices (como autor, orientador, título, etc.), como ilustra a Figura 3.8. Posteriormente os usuários podem realizar o download dos documentos, após o preenchimento de um cadastro.

Após o término da graduação (além de mestrado, PhD e especializações), os alunos enviam o documento digital resultante dos seus estudos à universidade. A cada ano, mais teses são adicionadas ao sistema, e algumas possuem restrições de acesso, devido à direitos legais específicos ou restrições de alguns produtos. A diferente qualidade dos metadados, o uso de arquivos distintos (como áudio e vídeo), além da integração com outras bibliotecas digitais são problemas que atualmente estão presentes no sistema. A má qualidade dos metadados está presente em teses antigas, já que informações específicas (como professores presentes na banca) não são fáceis de serem obtidas. O uso de arquivos distintos (como arquivos de áudio, geralmente utilizados pelo Instituto de Artes), ainda não é suportado pelo sistema. Já o problema de integração ocorre quando o sistema publica e coleta dados, através da rede brasileira de bibliotecas digitais (centralizada pelo Instituto Brasileiro de Informação em Ciência e Tecnologia - IBICT⁶).

O sistema Nou-Rau carece da agregação de objetos digitais distintos (como áudio, texto e imagem) e seus serviços, no contexto de teses e dissertações. Em casos semelhantes, o uso de objetos complexos com teses [8, 104] apresentaram os seguintes desafios: 1) os softwares de

⁶<http://www.ibict.br/> (último acesso em 05/05/11).

DLs não suportavam objetos complexos; 2) como os componentes do CO tinham direitos legais distintos, eles não poderiam ser tratados como um objeto único; 3) o sistema de busca teria que suportar não somente a busca dos componentes individuais, mas também a busca em que os componentes estão relacionados como um objeto único. Note que estes mesmos desafios estão presentes em vários aspectos nesta tese. Estas características foram fundamentais na compreensão inicial de requerimentos de objetos complexos, e sua integração com os serviços disponibilizados.

As três aplicações apresentadas neste capítulo exemplificam como dados heterogêneos podem ser integrados e posteriormente explorados em serviços. A Tabela 3.1 resume as principais características das três aplicações apresentadas neste capítulo.

Característica	SuperIDR	Impressões Digitais	Nou-Rau
Tipo de Multimídia	Imagem	Imagem	Texto
Tipo de Imagem	Parasitas	Impressões Digitais	–
Tipo de Metadado	Descrições de espécies	Descrição do indivíduo e da biblioteca digital	Descrições autorais da tese
Tipo de Serviço	Comparação, Busca, Navegação	Comparação, Navegação	Busca, Navegação

Tabela 3.1: Características das três aplicações utilizadas nesta tese.

3.4 Conclusões

Várias aplicações foram analisadas durante o doutorado (CtrNET, DL multilíngue, aplicações GIS, etc.). Em especial, este capítulo focou nas aplicações Nou-Rau, SuperIDR e a integração de bibliotecas digitais de Impressões Digitais.

O sistema Nou-Rau exemplifica a integração de vários documentos relacionados a teses e dissertações. Seus principais serviços incluem a busca e a navegação. A aplicação SuperIDR trata de imagens, metadados, marcas e anotações, relacionados ao domínio de parasitologia. Seus principais serviços incluem a navegação, a busca textual, a busca visual e a busca multi-modal. A integração de bibliotecas digitais de Impressões Digitais utiliza imagens e metadados relacionados a quatro bibliotecas digitais distintas. Estas aplicações serão utilizadas ao longo da tese para exemplificar estudos de caso na formalização, no uso das tecnologias e na infraestrutura proposta nesta tese.

Capítulo 4

Formalização de Objetos Complexos e Busca de Imagens por Conteúdo

Serviços avançados em bibliotecas digitais têm sido desenvolvidos e amplamente utilizados com o objetivo de gerenciar, dar manutenção e acesso a informações heterogêneas. A fim de reutilizar, integrar, unificar, e gerenciar esses recursos heterogêneos, objetos complexos surgiram para facilitar a agregação. Além disso, podem ser empregados por desenvolvedores para gerenciar informações heterogêneas e seus componentes. Em particular, objetos complexos de imagem (junto com o seu serviço mais utilizado - Busca de Imagens por Conteúdo) têm um papel fundamental nos sistemas de informação, devido à sua grande disponibilidade e integração com bancos de dados, metadados e softwares de manipulação de imagem.

Devido à falta de padrão sobre definições teóricas para esses serviços, a implementação muitas vezes envolve esforço adicional do desenvolvimento, levando a problemas de duplicação e interoperabilidade. Neste capítulo será utilizado o Arcabouço 5S para propor uma descrição formal de objetos complexos e busca de imagens por conteúdo, definindo os conceitos fundamentais e as relações de domínio em bibliotecas digitais. Em particular, utilizaremos estes conceitos para descrever o objeto complexo de imagem.

Esses conceitos podem ser usados para classificar, comparar e destacar as diferenças entre os componentes, tecnologias e aplicações, impactando os pesquisadores, designers e desenvolvedores de biblioteca digital. As extensões teóricas da funcionalidade de biblioteca digital aqui apresentados são reforçadas com um estudo de caso prático, para exemplificar a utilização integrada de serviços no equilíbrio de teoria e prática.

4.1 Notações Iniciais do Arcabouço 5S

O Arcabouço 5S busca descrever vários aspectos de bibliotecas digitais, como mostra a Figura 4.1. A figura ilustra as camadas de apoio das definições: fundamentos matemáticos (como

gráficos, seqüências e funções), os 5 Ss (*Streams*, *Structures* (Estruturas), *Spaces* (Espaços), *Scenarios* (Cenários), e *Societies* (Sociedades), e os conceitos-chave de uma biblioteca digital (por exemplo, objeto digital (*Digital Object*), coleção (*Collection*), catálogo dos metadados (*Metadata Catalog*), repositório (*Repository*). Setas representam dependências, indicando que um conceito é formalmente definido em termos de outros conceitos que apontam para ele.

No formalismo 5S:

- *Streams* descrevem propriedades de conteúdo, para material textual ou formas particulares de dados multimídia;
- *Structures* (Estruturas) especificam aspectos organizacionais (formatos de metadados, hipertextos, ordem temporal, taxonomias, esquemas de classificação);
- *Spaces* (Espaços) definem visões lógicas e de apresentação de vários componentes;
- *Scenarios* (Cenários) detalham o comportamento de serviços;
- *Societies* (Sociedades) definem gerentes (responsáveis pela execução de serviços), atores (usuários desses serviços), e relacionamentos.

Informalmente uma biblioteca digital é uma coleção de dados associada a serviços para uma comunidade, onde os dados são armazenados em formatos específicos e acessados através de uma rede. Os dados em uma biblioteca digital são manifestados através de objetos digitais, que podem ter como conteúdo texto ou multimídia (como imagem, vídeo ou áudio), e metadados. Os metadados geralmente descrevem características diversas do objeto e podem ser divididos entre várias categorias (como o metadado descritivo). Uma outra característica de objetos digitais e metadados é a presença de estruturas (como grafos ou árvores), que podem ser exploradas para prover melhores serviços de bibliotecas digitais. Dentre os serviços básicos de bibliotecas digitais, podemos citar a indexação, busca e navegação. Estes serviços podem ser adaptados à diferentes comunidades, dependendo de perfis, como autores, bibliotecários, administradores, etc. Adotaremos as mesmas definições apresentadas em [32–34], e as estenderemos para conceitos de objetos complexos e busca de imagens por conteúdo.

Notação: Seja DL1 uma biblioteca digital; sejam $do_1, do_2, \dots, do_n$ o conjunto de objetos digitais presentes em DL1; seja H o conjunto de identificadores universais únicos; seja SM um conjunto de *streams*; seja ST o conjunto de *Structural Metadata Specifications* [34].

4.2 Objetos Complexos

Em [34], é proposto que um **objeto digital** é uma tupla $do = (h, SM, ST, StructuredStreams)$, onde:

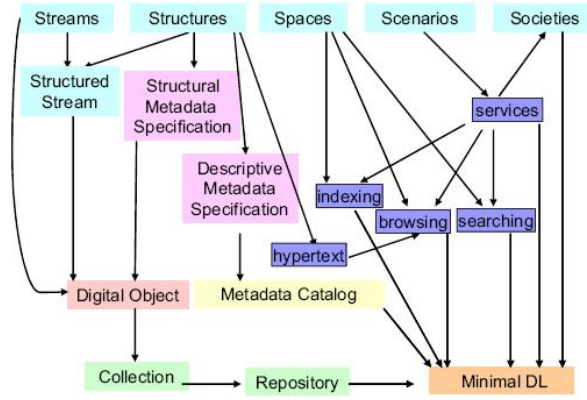


Figura 4.1: A estrutura de definição do Arcabouço 5S [115].

1. $h \in H$, onde H é um conjunto de identificadores universais únicos (*labels*);
2. $SM = \{sm_1, sm_2, \dots, sm_n\}$ é um conjunto de *streams*;
3. $ST = \{st_a, st_b, \dots, st_z\}$ é um conjunto de *Structural Metadata Specifications*;
4. $StructuredStreams = \{stsm_i, stsm_i, \dots, stsm_x\}$ é um conjunto de funções de *StructuredStream* definidas a partir de *streams* no conjunto SM do objeto digital e de estruturas no conjunto ST .

Streams são sequências de elementos de um tipo arbitrário (como bits, characters, imagens, etc.). *Structural Metadata Specifications* são relações entre o objeto e suas partes (como os capítulos de um livro). *Structured Streams* definem o mapeamento de uma estrutura para *streams* (como capítulos, seções, introdução, etc. são organizados para definir um livro). Mais detalhes estão disponíveis em [34].

Um **objeto complexo** é uma tupla $cdo = (h, SCDO, S)$, onde:

1. $h \in H$, onde H é um conjunto universal de identificadores únicos (*labels*);
2. $SCDO = \{DO \cup SM\}$, sendo $DO = \{do_1, do_2, \dots, do_n\}$, onde do_i é um objeto digital ou um objeto complexo; e $SM = \{sm_a, sm_b, \dots, sm_z\}$ é um conjunto de *streams*;
3. S é uma estrutura que compõe o objeto complexo cdo em suas partes em $SCDO$.

Na definição formal do Arcabouço 5S os metadados do objeto digital são armazenados em um catálogo separado [34]. Por consequência, os metadados do objeto complexo também serão armazenados em um catálogo, portanto não é referenciado na definição da tupla cdo . Os componentes DO e SM são conjuntos finitos, portanto, a estrutura S também é finita, delimitando

o que pertence ao objeto complexo ou não. Note que o componente *DO* pode ser outro objeto complexo, agregando COs dentro de outros COs.

A estrutura *S* no objeto complexo não é especificada, portanto, pode ser estendida a toda a estrutura que representa as partes de um todo, como uma lista, uma árvore, ou mesmo um grafo. Esta flexibilidade na formalização é estendida à implementação: é possível utilizar a estrutura de árvore e agregar vários arquivos em um CO (implementado como um arquivo compacto zip), ou realizar agregações mais específicas e complexas (como as implementadas na tecnologia OAI-ORE). Como exemplo prático, podemos citar a abordagem de *Fedora Commons* [9], onde listas representam vários arquivos que foram embalados em conjunto, e gráficos representam arquivos que estão relacionados, criando redes de objetos digitais. Se considerarmos ainda os arquivos organizados em HTML5 [93], a estrutura *S* poderia ser estendida para um grafo cíclico.

Um objeto complexo mínimo é uma tupla $cdo = (h, SCDO, S)$, onde:

1. $h \in H$, onde H é um conjunto universal de identificadores únicos (*labels*);
2. $SCDO = \{DO \cup SM\}$, sendo $DO = \{do_1\}$, onde do_1 é um objeto digital; e $SM = \{sm_1, sm_2, \dots, sm_n\}$ um conjunto de *streams*;
3. S é uma estrutura que indica que $\{do_1\}$ é um componente de cdo .

O CO mínimo têm por objetivo definir a menor granularidade permitida. Caso não sejam necessárias agregações, a definição em [34] pode ser aplicada. O objetivo deste capítulo não é explorar em detalhes estes conceitos, porém considerar uma abordagem de alto nível: agregar logicamente e até fisicamente, objetos distintos, de modo que a estrutura do CO que conecta suas subpartes possa ser explorada, e que o CO em si possa ser representado como uma unidade única.

Recursos também podem ser estruturados utilizando o conceito de coleção [34]. A principal diferença presente é que uma coleção é um simples conjunto de objetos [34], enquanto em um CO, componentes representam partes de um conceito único, e podem possuir relações específicas os conectando.

Considere a aplicação Nou-Rau (citada na Seção 3.3) e uma tese (Figura 4.2), com 12 capítulos, metadados, uma apresentação Powerpoint descrevendo o primeiro capítulo, e um vídeo referenciando o capítulo dois.

O objeto complexo que representa esta tese tem a estrutura $co = (h, SCDO, S)$, onde:

- h_1 é um identificador único que representa co ;
- $SCDO = \{DO \cup SM\}$, sendo $DO = \{do_1, do_2, do_3\}$, onde do_1 é a tese, do_2 é um arquivo Powerpoint, e do_3 é um vídeo; e $SM = \{sm_a, sm_b, \dots, sm_z\}$ é um conjunto de *streams*;
- S é a estrutura que identifica como do_1 , do_2 , e do_3 estão agregados.

Note que pela definição previamente apresentada de objeto digital, cada componente da tese também possui um identificador único. Portanto o exemplo apresentado possui 3 identificadores para os componentes, e um identificador adicional para o CO. Caso a implementação dos serviços precise navegar ou realizar buscas nos componentes, basta utilizar o primeiro identificador citado na definição de objeto digital. Caso seja necessário realizar buscas integradas, basta utilizar o identificador do objeto complexo.

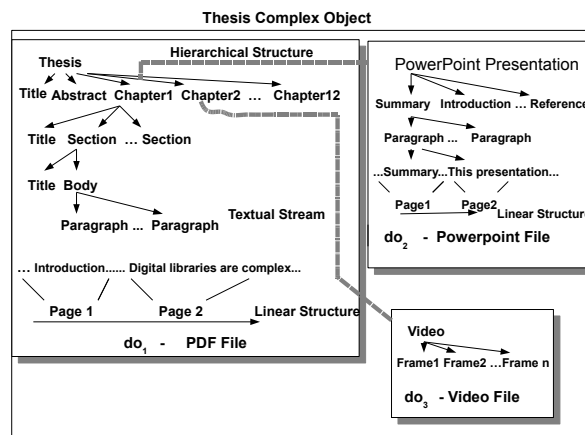


Figura 4.2: Exemplo de um objeto complexo que representa uma tese.

```
<metadata>
  <title>      Theses    </title>
  <creator>    Creator   </creator>
  <subject>    Subject   </subject>
  <description> Description </description>
  <publisher>  Publisher </publisher>
  <contributor> Contributor </contributor>
  <date>       Date      </date>
  ...
</metadata>

<Files>
  <Chapter 1>
    <PDF File> Chapter 1 PDF File </PDF File>
    <PPT File> PPT File </PPT File>
  </Chapter 1>
  <Chapter 2>
    <PDF File> Chapter 2 PDF File </PDF File>
    <Video File> Video File </Video File>
  </Chapter 2>
  ...
  <Chapter 12>
    <PDF File> Chapter 12 PDF File
    </PDF File>
  </Chapter 12>
</Files>
```

Figura 4.3: Exemplo de uma estrutura XML agregando os elementos do objeto complexo tese.

Neste caso o componente *S* pode ser representado formalmente como um grafo, porém na prática vários formatos podem ser utilizados (como a estrutura de árvore). A Figura 4.3 apresenta um exemplo simples para estruturar uma tese e seus respectivos elementos, utilizando XML. No exemplo, o arquivo lista inicialmente os metadados gerais a respeito da tese (como

título, autor, etc.). A seção seguinte apresenta uma estrutura para cada capítulo e seus respectivos arquivos. O capítulo hum, por exemplo, é composto por um arquivo PDF e um arquivo Powerpoint. O capítulo dois referencia um arquivo PDF e um vídeo. A estrutura XML também já foi utilizada para combinar esquemas e o Arcabouço 5S [36].

4.3 Busca de Imagens por Conteúdo

A Figura 4.4 apresenta os conceitos propostos baseados no Arcabouço 5S para tratar descrições de conteúdo para a manipulação de imagens e serviços relacionados a bibliotecas digitais. Estes conceitos serão definidos a seguir.

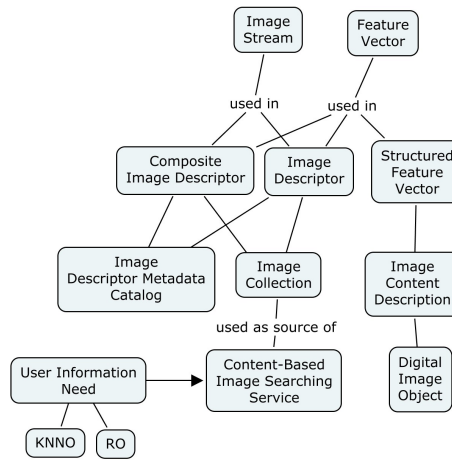


Figura 4.4: Extensões do Arcabouço 5S para a descrição de imagens por conteúdo e serviços relacionados.

Alguns destes conceitos foram introduzidos em [19]. A seguir, eles serão estendidos, considerando aspectos relacionados a bibliotecas digitais.

Uma **stream** de imagem (ou simplesmente **imagem**) $\hat{I}$ é um par $(D_I, \vec{I})$, onde:

- D_I é um conjunto finito de *pixels* (pontos em N^2 , em que $D_I \subset N^2$), e
- $\vec{I}: D_I \rightarrow D'$ é uma função que assigna cada pixel p em D_I a um $\vec{I}(p)$ de valores em um espaço arbitrário D' (por exemplo, $D' = \mathbb{R}^3$ quando uma cor no sistema RGB é assignada a um pixel).

Um **vetor de características** $\vec{f}_{v_{\hat{I}}}$ de uma imagem $\hat{I}$ é um ponto no espaço R^n : $\vec{f}_{v_{\hat{I}}} = (f_{v_1}, f_{v_2}, \dots, f_{v_n})$, onde n é a dimensão do vetor.

Exemplos de possíveis vetores de característica incluem o histograma de cor [120], a curva fractal multi-escala [124], um conjunto de coeficientes Fourier [100] e um conjunto de características locais [5]. Eles essencialmente codificam as propriedades de imagens, como cor,

forma e textura. Note que vetores de características diferentes podem trabalhar com funções de similaridades distintas.

Um **objeto digital vetor de características** é um objeto digital $do = (h, SM, ST, StructuredStreams)$ que representa o vetor de características.

Dada a estrutura $(G, L, \mathcal{F})$, $G = (V, E)$ e um vetor de características $\vec{f}_{v_{\hat{I}}}$, um **Structured-FeatureVector** é uma função $V \rightarrow R^n$ que associa cada nodo $v_k \in V$ com $f_{v_i} \in \vec{f}_{v_{\hat{I}}}$.

A Figura 4.5 apresenta um exemplo do uso de uma função **StructuredFeatureVector**. Neste caso, uma estrutura XML (*structural metadata specification*) é mapeada para um vetor de características obtido através da aplicação do descritor de imagem *Contour Multiscale Fractal Dimension* [124] na imagem *stream* definida pelo arquivo “fish0.pgm”.

```
<?xml version='1.0' encoding='UTF-8'?>
-<feature_vector:Feature_Vector xmlns:feature_vector='http://feathers.dlib.vt.edu/~rtorres/'
  xmlns:xsi='http://www.w3.org/2001/XMLSchema-instance'
  xsi:schemaLocation='http://feathers.dlib.vt.edu/~rtorres/
    http://feathers.dlib.vt.edu/~rtorres/feature_vector.xsd'>
  <feature_vector:ImageName>fish0.pgm</feature_vector:ImageName>
  <feature_vector:DescriptorName>ContourMSFractalDimension</feature_vector:DescriptorName>
  <feature_vector:Type>1</feature_vector:Type>
  <feature_vector:Curve>
    <feature_vector:Nelements>25</feature_vector:Nelements>
    <feature_vector:CurveID>
      <feature_vector:X>
        <feature_vector:value>0.95105259594482394192</feature_vector:value>
        <feature_vector:value>0.98551214588154611995</feature_vector:value>
        <feature_vector:value>1.00415492765507829986</feature_vector:value>
        <feature_vector:value>1.00931032237937512441</feature_vector:value>
        <feature_vector:value>1.00583781572741104426</feature_vector:value>
        (...)
        <feature_vector:value>0.93810555611087775851</feature_vector:value>
        <feature_vector:value>0.87275204902189629230</feature_vector:value>
        <feature_vector:value>0.81066432563100665476</feature_vector:value>
        <feature_vector:value>0.75224263059381879515</feature_vector:value>
      </feature_vector:X>
    </feature_vector:CurveID>
  </feature_vector:Curve>
</feature_vector:Feature_Vector>
```

Figura 4.5: Exemplo de um *structured feature vector*.

Um **descritor de conteúdo de imagem simples** (em resumo, **descritor de imagem**) D é definido como uma tupla $(h_{desc}, \varepsilon_D, \delta_D)$, onde:

- $h_{desc} \in H$, onde H é o conjunto de identificadores universais únicos (*labels*);
- $\varepsilon_D : \{\hat{I}\} \rightarrow R^n$ é uma função que extrai um *vetor de características* $\vec{f}_{v_{\hat{I}}}$ de uma *imagem* $\hat{I}$.
- $\delta_D : R^n \times R^n \rightarrow R$ é uma *função de similaridade* (por ex., baseada na distância métrica) que computa a similaridade entre duas imagens como uma função da distância entre os correspondentes *vetores de características*.

A Figura 4.6(b) ilustra o uso de um descritor simples D para computar a similaridade entre duas imagens $\hat{I}_A$ e $\hat{I}_B$. Inicialmente o algoritmo de extração ε_D é usado para computar os vetores de características $\vec{f}_{v_{\hat{I}_A}}$ e $\vec{f}_{v_{\hat{I}_B}}$ associados com as imagens. Em uma segunda etapa a função de similaridade δ_D é utilizada para determinar o valor de similaridade d entre as imagens.

Um **descritor de imagem composto** $\hat{D}$ é uma tupla $(h_{desc}, \mathcal{D}, \delta_{\mathcal{D}})$ (Figure 4.6(b)), onde:

- $h_{desc} \in H$, onde H é um conjunto dos identificadores universais únicos (*labels*);
- $\mathcal{D} = \{D_1, D_2, \dots, D_k\}$ é um conjunto de k descritores de imagens simples pré-definidos.
- $\delta_{\mathcal{D}}$ é a função de similaridade que combina o valor de similaridade obtido de cada descritor $D_i \in \mathcal{D}$, $i = 1, 2, \dots, k$.

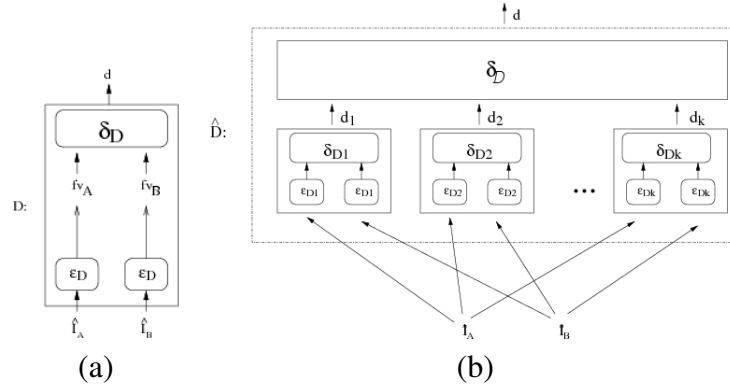


Figura 4.6: (a) O uso de um descritor simples D para computar a similaridade de imagens. (b) Descritor de imagem composto.

Uma **descrição de conteúdo de imagem ICD** é uma tupla $(FV, ST_{FVS}, Structured_{FVS})$, onde

- $FV = \{\vec{fv}_1, \vec{fv}_2, \dots, \vec{fv}_k\}$ é um conjunto de vetores de características;
- $ST_{FVS} = \{stfv_1, stfv_2, \dots, stfv_m\}$ é um conjunto de *structural metadata specifications*;
- $Structured_{FVS} = \{strfv_1, strfv_2, \dots, strfv_m\}$ é um conjunto de funções *Structured-FeatureVector* definidas a partir dos *vetores de características* no conjunto FV (o primeiro componente) da descrição de conteúdo de imagem e das estruturas do conjunto ST_{FVS} (o segundo componente).

Um **objeto digital imagem ido** é um objeto digital com as seguintes restrições:

- ido é um *objeto digital* $= (h, SM, ST, StrStreams, ICD, StrICDStreams)$, onde:
 1. $h \in H$, onde H é um conjunto de identificadores universais únicos (*labels*);
 2. $SM_{sd} = \{sm_{sd}[i, j]\} \in SM$, onde $sm_{sd}[i, j] = \langle a_i, \dots, a_j \rangle, 0 \leq i \leq j \leq n$. $sm_{sd}[i, j]$ refere-se a *substreams* (regiões) de uma imagem *stream*.
 3. $ST = \{st_1, st_2, \dots, st_m\}$ é um conjunto de *structural metadata specifications*;

4. $StrStreams = \{stD_1, stD_2, \dots, stD_m\}$ é um conjunto de funções *StructuredStream* definidas de *substreams* da imagem no conjunto SM (o segundo componente) do objeto digital e das estruturas do conjunto ST (o terceiro componente).
5. ICD é uma *descrição do conteúdo da imagem*.
6. $StrICDStreams = \{stimgD_1, stimgD_2, \dots, stimgD_m\}$ é um conjunto de funções *StructuredStream* definidas a partir da *stream de imagem* no conjunto SM (o segundo componente) do objeto digital de imagem e das estruturas do conjunto $ST_{FVS} \in ICD(2)$.

A Figura 4.7 ilustra as relações entre os conceitos utilizadas para definir um objeto digital de imagem.

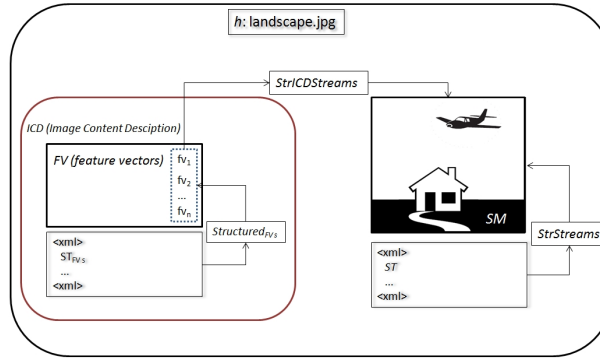


Figura 4.7: Os elementos do objeto digital de imagem.

A definição de $StrICDStreams$ permite a associação de vetores de características às partes de *stream* de imagens (objetos, regiões).

Uma **coleção de imagens** $ImgC$ é uma tupla $(C, S_{imgdesc}, FV_{imgdesc})$, onde C é uma coleção (Def. 17 em [34]), $S_{imgdesc}$ é um conjunto de descritores de imagem, e FV_{desc} é uma função $FV_{desc} : \{C \times S_{imgdesc}\} \rightarrow ICD(1)$, onde ICD é $ido(5)$ e $ido \in C$.

A função FV_{desc} define como um vetor de características foi obtido, dado um objeto digital de imagem $ido \in C$ e um descritor de imagem $\hat{D} \in S_{imgdesc}$.

Seja $S_{imgdesc}$ um conjunto de descritores de imagem com k identificadores em H . Um **catálogo de metadados de descritores de imagem** $DM_{S_{imgdesc}}$ para $S_{imgdesc}$ é um conjunto de pares $\{(h, \{dmdesc_1, \dots, dmdesc_{k_h}\})\}$, onde $h \in H$ e $dmdesc_i$ são *Descriptive metadata specifications* para descritores de imagem.

Descriptive metadata specifications de descritores poderiam incluir, por exemplo, dados sobre o autor (responsável pela implementação da extração e das funções de similaridade), data

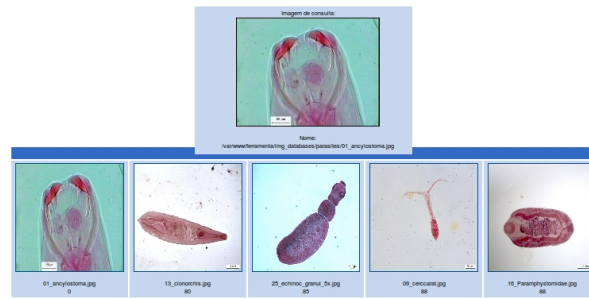


Figura 4.8: Busca de imagem por conteúdo para a imagem 01_ancylostoma.jpg.

da implementação, e a publicação associada. Em geral, um catálogo de metadados é utilizado para assignar especificações de metadados descritivas para objetos digitais de imagem (Def. 18 em [34]).

4.4 Estudo de Caso

Neste estudo de caso utilizaremos o Arcabouço 5S para definir e analisar o conteúdo e o comportamento de uma ferramenta de recuperação e descrição de imagens na aplicação SuperIDR (descrita na Seção 3.1). Os componentes dessa ferramenta não são diferentes de uma biblioteca digital com funcionalidade estendida.

Para exemplificar o conceito de objeto complexo de imagem, considere a espécie *Ancylostoma caninum*, com a imagem 01_ancylostoma.jpg (imagem da cabeça), e seus respectivos metadados.

A busca de imagem por conteúdo da imagem 01_ancylostoma.jpg gera um vetor de características, e suas medidas de distância para outras imagens da coleção. A Figura 4.8 mostra o ranking para a imagem 01_ancylostoma.jpg de acordo com o descritor de cor (BIC [118]). As 5 imagens da região inferior são as mais similares se comparadas com 01_ancylostoma.jpg (se considerarmos a comparação baseada em cor). A Figura 4.9 mostra um exemplo de uma estrutura XML para identificar os componentes deste objeto complexo de imagem. Note que a estrutura organiza informações como o nome da imagem, o seu respectivo vetor de características, o descritor, e as medidas de distância para as cinco imagens mais similares.

Para o propósito de formalização, adotaremos a definição de objeto complexo para descrever uma imagem e seus componentes. A diferença entre um **objeto complexo de imagem** ICO (baseada na definição de CO presente na seção 4.2) e um **objeto digital de imagem** IDO (seção 4.3) é que a definição do objeto complexo considera a imagem e o vetor de característica como objetos digitais atômicos, para facilitar a manipulação, a agregação, e os serviços relacionados (como referência, publicação, etc.).

```

<?xml version="1.0" encoding="UTF8"?>
<image>01_ancylostoma.jpg
  <image_name>01_ancylostoma.jpg</image_name>
  <image_path>/home/nadiapk/data/uploads/01_ancylostoma.jpg</image_path>
  <image_feature_vector_name>/data/fv/01_ancylostoma.txt
  </image_feature_vector_name>
  <image_descriptor>Bic
    <image_name>01_ancylostoma.jpg<image_dist_value>0
    </image_dist_value></image_name>
    <image_name>13_clonorchis.jpg<image_dist_value>80
    </image_dist_value></image_name>
    <image_name>25_echinoc_granul_5x.jpg<image_dist_value>85
    </image_dist_value></image_name>
    <image_name>09_cerccarat.jpg<image_dist_value>88
    </image_dist_value></image_name>
    <image_name>16_Paramphystomidae.jpg<image_dist_value>88
    </image_dist_value></image_name>
  </image_descriptor>
</image>
</image>

```

Figura 4.9: Exemplo de uma estrutura XML para um objeto complexo de imagem.

Se considerarmos a definição de objeto complexo (Seção 2.3), o objeto complexo de imagem (ICO) tem a estrutura $ico = (h, SCDO, S)$, onde:

- $h1$ é um identificador único que representa ico;
- $SCDO = \{DO \cup SM\}$, sendo $DO = \{do_1, do_2, \dots, do_k, do_{k+1}, \dots, do_{k+3}\}$, onde do_1 é uma **imagem**, k é o número de descritores, $do_2, \dots, do_k$ é um conjunto de **objetos digitais do tipo StructuredFeatureVector**, e $do_{k+1}, \dots, do_{k+3}$ é um conjunto de **StructuredFeatureVectors** (com as medidas de distância, de acordo com um descritor k); e $SM = \{sm_a, sm_b, \dots, sm_z\}$ é um conjunto de *streams*;
- S é a estrutura que identifica como $do_1, do_2, \dots, do_k$, e $do_{k+1}, \dots, do_{k+3}$ estão agregados.

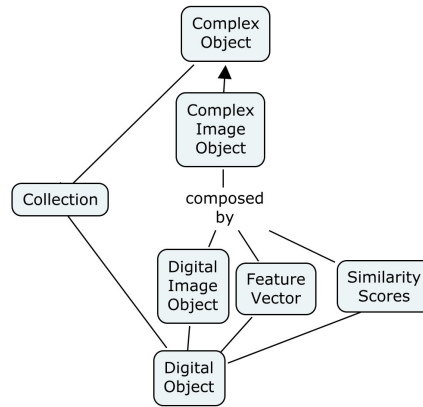


Figura 4.10: O Objeto Complexo de Imagem.

Uma **coleção de objetos complexos de imagem** $ImgCO$ é uma tupla $(C_{ICO}, S_{imgdesc})$, onde C_{ICO} é uma coleção (Def. 17 em [34]), e $S_{imgdesc}$ é um conjunto de descritores de imagem. A

Figura 4.10 apresenta o ICO, e seus componentes correspondentes: a imagem, o vetor de características e as medidas de distância. O vetor de características é obtido através de FV_{desc} , dado um objeto complexo de imagem $ico \in C_{ICO}$ e um descritor de imagem $\hat{D} \in S_{imgdesc}$. Note que cada componente do ICO é um **objeto digital**, com seu próprio identificador. A identificação única permite que a coleção seja explorada não somente através dos seus componentes individuais, como também através das agregações em ICOs.

Na ferramenta SuperIDR, cada ICO tem a estrutura $ico = (h, SCDO, S)$, onde:

- $h1$ é um identificador único que identifica ico;
- $SCDO = \{DO \cup SM\}$, sendo $DO = \{do_1, do_2, do_3\}$, onde do_1 é uma **imagem de parasita**, $k = 1$ (descritor BIC [118]), do_2 é um **objeto digital do tipo StructuredFeatureVector**, e do_3 é um **StructuredFeatureVector** (com as medidas de distância, de acordo com o descritor k); e $SM = \{sm_a, sm_b, \dots, sm_z\}$ é um conjunto de *streams*;
- S é uma estrutura XML (Figura 4.9) que identifica como do_1 , do_2 e do_3 estão agregados.

A **coleção de objetos complexos de imagem** representada na ferramenta SuperIDR têm 43 objetos complexos de imagem e um descritor.

4.5 Conclusões

Muitas implementações em bibliotecas digitais requerem serviços mais especializados, se comparados com aplicações tradicionais. Exemplos destes serviços incluem o suporte a imagem, objetos multimídia, partes de documentos e anotações.

Neste capítulo o Arcabouço 5S foi estendido para conceitos iniciais de agregações em objetos complexos, além do tratamento de busca de imagens por conteúdo. Este formalismo pode ser utilizado para descrever agregações e vários conceitos disponíveis no Arcabouço 5S. O conjunto de definições também pode impactar aplicações futuras (como as apresentadas em [33, 76, 112, 115, 135]), já que servem de suporte para aplicações baseadas em objetos complexos e conteúdo de imagem. Os conceitos propostos foram ilustrados a partir de um estudo de caso com a ferramenta SuperIDR.

Trabalhos futuros nesta área incluem a formalização de serviços que utilizam estas estruturas. Exemplos incluem serviços de busca multimodais, serviços mínimos em objetos complexos (como agregação, versões), serviços de recomendação, e o serviço de navegação de imagens baseado em conteúdo.

Capítulo 5

Analizando as Tecnologias de Objetos Complexos da Perspectiva do Arcabouço 5S

Existem várias aplicações que precisam de suporte para objetos complexos. Este suporte carece de mecanismos para a gestão de integração de dados; ajuda para a criação de referências, links e anotações; e serviços para agrupar, organizar e reutilizar objetos complexos e seus componentes. Poucas tentativas foram feitas para caracterizar formalmente os objetos complexos, os serviços relacionados, e suas tecnologias. Neste capítulo tratamos da descrição e da interação de tecnologias para tratamento de objetos complexos utilizando a formalização proposta pelo Arcabouço 5S. Este capítulo (1) analisa três tecnologias que manipulam objetos complexos (DCC, Buckets, e OAI-ORE); (2) utiliza o Arcabouço 5S para descrevê-las; (3) e apresenta um estudo de caso, ilustrando como as tecnologias de objetos complexos e o Arcabouço 5S podem ser utilizadas juntas para apoiar a exploração de objetos complexos.

5.1 Analizando o DCC, Buckets, e OAI-ORE

Com os principais conceitos de DCC, Buckets, OAI-ORE, e o Arcabouço 5S resumidos no Capítulo 2, e a formalização de CO no Capítulo 4, esta seção apresenta, inicialmente, o tratamento formal para descrever COs utilizando as tecnologias DCC, Buckets e OAI-ORE.

Das várias tecnologias disponíveis, o DCC foi escolhido devido ao suporte a diferentes aspectos do CO, como o encapsulamento de software. O OAI-ORE provê o agrupamento e mapeamento dos componentes do CO, e tem sido amplamente utilizado recentemente para agrupar recursos. E Buckets é amplamente utilizado pela comunidade de bibliotecários, permitindo links a pacotes remotos, redes ou base de dados.

Apesar de seus contextos de utilização, as três tecnologias manipulam uma quantidade grande de conceitos. O DCC foca em (i) uma identidade única; (ii) uma delimitação persistente do objeto, definindo o que é parte dele ou não; (iii) a concessão da composição diversificada dos componentes; e (iv) a composição de COs, criando hierarquias e camadas. Buckets, em contrapartida, requer *handles*, componentes (chamados de *packages*), elementos, métodos de acesso e containers. Finalmente, a arquitetura do OAI-ORE essencialmente consiste de: (i) identificadores; (ii) recursos; (iii) protocolo padrão, como o HTTP, para o acesso ao dado; (iv) links através da referência URI; e (v) *named graphs* para encapsular a informação em um objeto composto.

Considerando todos os aspectos mencionados, quatro características estão sempre presentes: a identidade, os componentes, a estrutura e a delimitação do objeto. A identidade é responsável pela identificação única do CO. Os componentes representam as partes do objeto complexo. A estrutura representa como os componentes estão relacionados. E a delimitação do objeto distingue os componentes que fazem parte do objeto (como a página de um livro), e os componentes que não fazem parte do objeto [65]. Focaremos nossa análise nestes quatro conceitos, junto com conceitos presentes no Arcabouço 5S para explorar as diferenças entre o DCC, Buckets e OAI-ORE.

5.1.1 Uso do Formalismo 5S para DCC, Buckets e OAI-ORE

Nesta seção utilizamos a definição formal de CO apresentada no Capítulo 4 para a descrição das tecnologias DCC, Buckets e OAI-ORE.

DCC

Um objeto complexo no DCC pode ser descrito como um elemento $cdo = (h, SCDO, S)$, onde:

1. $h \in H$, e H é um conjunto de identificadores únicos, representados por URIs e URIs relativas;
2. $SCDO = \{DO \cup SM\}$, sendo $DO = \{do_1, do_2, \dots, do_n\}$, onde do_i é um DCC. Um CO pode também ser representado por um DCC composto por outros DCCs; e $SM = \{sm_a, sm_b, \dots, sm_z\}$ é um conjunto de *streams*;
3. S é uma estrutura XML que compõe o objeto complexo cdo em suas partes presentes em $SCDO$.

As diferentes sociedades (*Societies*) podem ser exploradas de acordo com diferentes grupos de usuários (que precisem de acesso a DCCs individuais ou compostos), além dos componentes de software. As regras de acesso para membros com papéis distintos também podem ser mapeadas (e que eram descritas através do uso de OWL nos DCCs).

Um cenário (*Scenario*) é definido como uma descrição de interações entre o usuário e uma aplicação, e será definido pelas funções disponíveis na interface do DCC. Cenários diferentes podem ser posteriormente combinados em um único DCC. O software encapsulado também pode ser mapeado no Arcabouço 5S através de um cenário estruturado.

A estrutura (*Structures*) representa o modo que o XML organiza os componentes no DCC. O conteúdo encapsulado no DCC define o que pertence ao CO ou não (*boundary*).

O conceito de *Streams* representa os diferentes tipos de multimídia manipulados. Serviços baseados em um tipo específico (como busca textual ou busca baseada em conteúdo) podem ser reutilizados e combinados em um único DCC (busca multimodal).

O vocabulário utilizado nos diferentes DCCs para a descrição do conteúdo, estrutura, interface, metadados, versões, funcionalidade, e aplicabilidade está relacionado ao conceito de *Space*.

Note que, dependendo da implementação, não necessariamente uma estrutura S do Arcabouço 5S é a mesma estrutura utilizada no DCC. Em particular, a estrutura do DCC foi adaptada na infraestrutura implementada nesta tese para representar a agregação de componentes.

Buckets

Um objeto complexo em Buckets pode ser representado por um elemento $cdo = (h, SCDO, S)$, onde:

1. $h \in H$, onde H é um conjunto de identificadores únicos;
2. $SCDO = \{DO \cup SM\}$, sendo $DO = \{do_1, do_2, \dots, do_n\}$, onde do_i é um elemento, um package ou um bucket; e $SM = \{sm_a, sm_b, \dots, sm_z\}$ é um conjunto de *streams*;
3. S é uma estrutura de diretório Unix que unifica todos os componentes de uma agregação.

As diferentes sociedades (*Societies*) podem ser exploradas de acordo com diferentes grupos de usuários (que precisem de acesso a métodos, packages e buckets) e administradores. As regras de acesso para membros com papéis distintos também podem ser mapeadas (definidas pelas ferramentas *author* e *management*). Cada grupo de usuários pode interagir com serviços específicos, utilizando os métodos disponíveis para interagir com os packages e elementos. Cenários novos podem ser adicionados com buckets aninhados.

A estrutura (*Structures*) representa o modo que diferentes arquivos são organizados em um diretório unix, packages e buckets. O conteúdo de um bucket e package define o que pertence ao CO ou não (*boundary*), incluindo links para serviços em rede, chaves estrangeiras em bases de dados, etc.

O conceito de *Streams* representa os diferentes tipos de multimídia manipulados. Serviços específicos para um tipo de componente podem ser representados por métodos (como *Add_package*, *List_methods*, etc.).

O vocabulário utilizado nos diferentes buckets, packages, elementos, termos e condições de um bucket estão relacionados com o espaço (*Space*) conceitual.

OAI-ORE

Um objeto complexo em OAI-ORE pode ser representado por um elemento $cdo = (h, SCDO, S)$, onde:

1. $h \in H$, where H é um conjunto de identificadores únicos, representado por URIs;
2. $SCDO = \{DO \cup SM\}$, sendo $DO = \{do_1, do_2, \dots, do_n\}$, onde do_i é um recurso. O objeto complexo pode ser representado por seus componentes; e $SM = \{sm_a, sm_b, \dots, sm_z\}$ é um conjunto de *streams*;
3. S é um *resource map* presente no OAI-ORE que compõe o objeto complexo cdo a partir de suas partes em $SCDO$.

Entre os exemplos de sociedades (*Societies*), podemos citar as comunidades que interagem com o intercâmbio e reuso dos diferentes recursos digitais de repositórios. Sociedades específicas podem ser exploradas através do uso de *named graphs*.

Cenários (*Scenario*) podem ser utilizados para descrever os processos de criação, intercâmbio, visualização, reuso e preservação das agregações. Cenários também podem ser utilizados para descrever as relações entre comunidades.

Estruturas (*Structures*) são utilizadas para modelar e organizar recursos, representadas no OAI-ORE pelo RDF/XML e RDFa [71]. O conteúdo do *named graph* delimita o que pertence ao CO ou não (*boundary* [65]).

O conceito de *Streams* representa os diferentes tipos de recursos manuseados. Serviços (tal como a publicação do CO) podem tirar vantagem do *resource map*, utilizando conceitos básicos do OAI-ORE, como a identidade (URI), os elementos da agregação (recursos) e o método de intercâmbio (protocolo HTTP).

O vocabulário para expressar os diferentes tipos de links entre recursos e propriedades de recursos (como tipo semântico, tipo de mídia e formato) está relacionado ao conceito de espaço (*Space*).

5.1.2 Explorando as Diferenças entre DCC, Buckets e OAI-ORE

As tecnologias DCC, Buckets e OAI-ORE têm sido utilizadas com finalidades diferentes, porém o foco é o mesmo: agregar recursos. As diferentes vantagens surgem: a partir da perspectiva de *Spaces*, o DCC trabalha com ontologias, enquanto os fluxos de perspectiva, e a estrutura HTML do protocolo OAI-ORE facilita a sua integração com outras aplicações. As operações e as restrições são diferentes, pois eles têm finalidades diferentes dentro do contexto de CO. A agregação de informações podem utilizar diversas abstrações para diferenciar os componentes

internos, como *named graphs*, arquivos XML e diretórios Unix. Perspectivas diferentes de uma mesma entidade pode ser explorados através de interfaces, métodos, ou *named graphs*.

Para destacar ainda mais as suas diferenças, selecionamos itens adicionais, relacionados aos conceitos de identidade, componentes, estrutura, delimitação do objeto e manipulação: (i) identificador único, (ii) a divisão de componentes, (iii) como os componentes são compostos, (iv) o que é encapsulado; (v) utilização, (vi) o formato e estrutura interna; (vii) instrumentação e regras de acesso; (viii) vantagens; (ix) como é feita a manipulação de software; (x) o uso na literatura, e (xi) a forma como eles lidam com a preservação a longo prazo de objetos digitais.

Tabela 5.1: Conceitos básicos de CO da perspectiva de DCC, Buckets, e OAI-ORE.

Descrição	DCC	Buckets	OAI-ORE
Identificador único	URI	Handle	URI
Divisão interna	DCC Passivo e de Processo	Diretórios Unix	<i>Resource map</i> , agregações
A Composição do CO	Partes acessadas através da URI relativa, DCCs	Packages e Buckets agregados	<i>Resource map</i> e agregações
O que pode ser encapsulado?	Metadados, conteúdo, processos	Metadados, conteúdo	Metadados
Utilização	Encapsulamento de conteúdo	Construção de Container independente do repositório	Reuso de objetos, intercâmbio
Formato	Conteúdo, estrutura, interface e metadados	RFC-1807, Diretórios Unix	Mapeamento através de <i>resource map</i> , URIs, agregação
Implementação	Arquivo Jar, extensível para outras linguagens	Acesso através do <i>Author e Management Tool</i>	Mapeamento de recursos através do <i>resource map</i>
Vantagens	Ontologia, interface, encapsula software	Ponteiros para package remotos, redes ou base de dados, log	Intercâmbio de repositórios, utilizado como padrão entre sistemas distintos
Administração de software?	Encapsula conteúdo e software	Como um arquivo normal	Como um arquivo normal
Principal uso na literatura	Encapsula conteúdo em aplicações e web	Objetos de Repositórios ativos	Para descrever objetos e agregações
Preservação	Encapsula conteúdo executável e não executável, estrutura, metadados, e permite reuso	Diretórios podem ser zipados para backup ou transporte	A descrição facilita o transporte e reuso

As tecnologias podem apresentar vantagens diferentes, como o encapsulamento executável no DCC, o ponteiro para packages remotos em Buckets, ou a maneira fácil de visualizar agregações no OAI-ORE. A reutilização de informações pode explorar os diferentes componentes de uma agregação. Eles podem usar as agregações de metadados, através de gráficos como descrições (como o OAI-ORE), ou gerir conteúdos com metadados (como o Buckets), ou mesmo realizando o encapsulamento de processos, conteúdo e metadados, no DCC. O objetivo não é compará-los, mas entender os seus diferentes cenários e suas diferenças, para que possam ser mais bem explorados, mapeados, ou até mesmo combinados posteriormente.

Mesmo apresentando vantagens diferentes, da perspectiva de conceitos verificados neste trabalho (identidade, estrutura, componentes, e a delimitação do objeto), o DCC, Buckets e OAI-ORE apresentam e gerem as mesmas informações, podendo então ser mutuamente mapeados. Um exemplo são os mapeamentos apresentados na Seção 5.1.

5.2 Estudo de Caso

Nesta seção apresentamos um estudo de caso para ilustrar uma melhor compreensão de como os conceitos de identidade, estrutura, delimitação do objeto e componentes podem ser explorados em aplicações reais de bibliotecas digitais, em particular, em uma biblioteca digital de impressões digitais [58, 59, 74, 92, 116] (apresentada no Capítulo 3).

Com a integração, a biblioteca digital unifica quatro comunidades diferentes, permitindo que cada uma delas avalie as informações de diferentes perspectivas, e explorem o sistema como um todo, ou com o foco em uma determinada área. Além disso, podemos tirar vantagem dos serviços da biblioteca digital (por exemplo, navegação e buscas), formalismos, e soluções na área de preservação digital.

De acordo com [106], o processo de integração é dividido em quatro etapas: (i) descoberta: os sistemas “aprendem” sobre a existência um do outro; (ii) identificação: os sistemas identificam os itens disponíveis; (iii) acesso: os sistemas acessam seus itens; e (iv) utilização: os sistemas sintetizam os seus itens. Nosso estudo de caso explora as duas primeiras etapas citadas acima.

Utilizamos o conceito de COs para facilitar a abstração de agregação (como mostra a Figura 5.1), abrangendo componentes de domínios diferentes, e unificando-os com um único conceito.

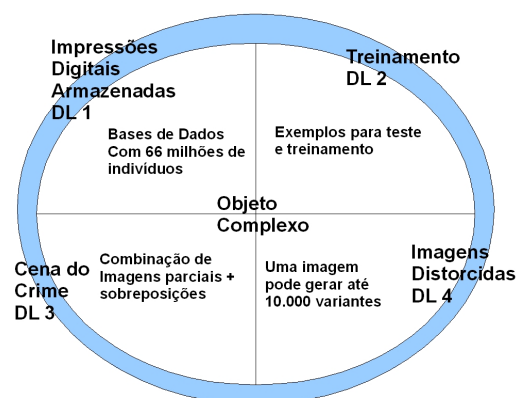


Figura 5.1: A integração de bibliotecas de impressões digitais.

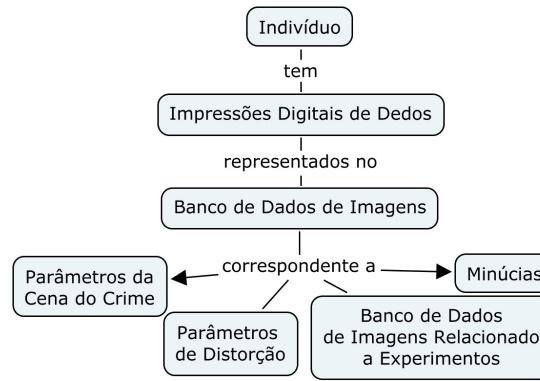


Figura 5.2: As classes principais representando uma biblioteca digital de impressões digitais.

A Figura 5.2 apresenta o mapa conceitual das principais classes, sendo o resumo do diagrama de entidade-relacionamento [58]. A classe indivíduo, por exemplo, agrega todas as informações dos dedos das mãos esquerda e direita, junto com as imagens, minúcias, e outros metadados para uma única pessoa. Posteriormente, o usuário pode explorar se a mesma pessoa tem imagens distorcidas por algoritmos, ou extraídas de uma cena de crime, ou manipuladas por uma delegacia.

A integração dos quatro sub-sistemas pode ser exemplificado pela Figura 5.3. O objeto complexo 1 (CO1) tem os seguintes componentes: a imagem do polegar direito proveniente do sistema A, uma imagem distorcida do sistema B, uma imagem da cena do crime do sistema C, e um link para o material de treinamento do sistema D. Os componentes podem ser identificados por CO1.A.1, CO1.B.1, CO1.C.1 e CO1.D.1, respectivamente. A estrutura CO1 pode ser representada e descrita por RDF no OAI-ORE, enquanto que o conteúdo pode ser encapsulado em DCC. A interface de CO1 pode incluir as informações de união de seus quatro componentes, juntamente com a união de seus respectivos vocabulários (como as palavras indivíduo, dedos, polegar, qualidade, distorção, parâmetros, etc.)

Se considerarmos o tratamento CO formal da Figura 5.3, temos $CO1 = (h, SCDO, S)$, onde:

1. h é um identificador único que representa CO1, e $h \in H$, onde H é um conjunto de identificadores únicos universais (labels);
2. $SCDO = \{DO \cup SM\}$, sendo $DO = \{A.1, B.1, C.1 \text{ e } D.1\}$; e $SM = \{sm_1, sm_2, \dots, sm_n\}$ é um conjunto de streams;
3. S é uma estrutura XML que compõe o objeto complexo cdo em suas partes em $SCDO$.

Além disso, o Arcabouço 5S pode ser usado para mapear outros aspectos de uma biblioteca digital. Considere a descrição de *Society* e *Structures*. Como exemplos de sociedades em uma biblioteca de impressão digital podemos citar as agências de justiça criminal, os acadêmicos, os estudantes e pesquisadores. Como exemplos de estruturas em uma biblioteca de impressão

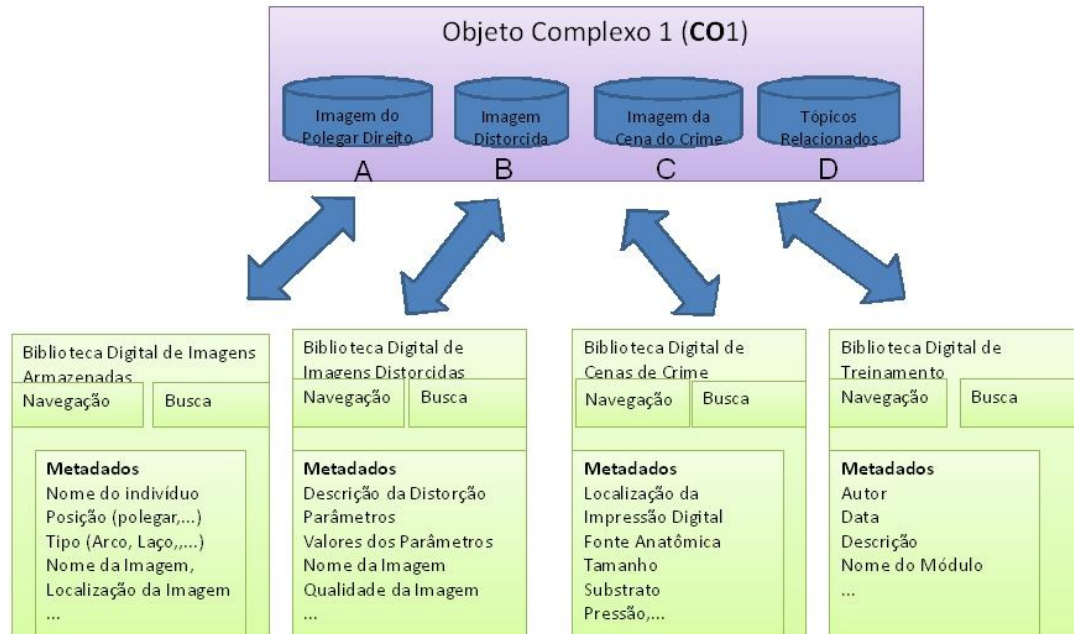


Figura 5.3: Um exemplo de objeto complexo com as quatro bibliotecas digitais: (A) Impressões Digitais armazenadas, (B) Imagens distorcidas, (C) Imagens da Cena do crime, and (D) Material de Treinamento.

digital podemos citar a organização da informação presente na biblioteca digital de impressões digitais mencionada. Esta estrutura pode relacionar uma imagem armazenada de um dedo polegar, com as imagens correspondentes provenientes da delegacia de polícia, ou da cena de crime. A estrutura interna do objeto pode ser descrita e implementada através de XML.

A exploração dos conceitos de CO de um projeto em estágio inicial de desenvolvimento é importante para destacar a quantidade de informação e detalhes necessários para gerenciar, integrar e agregar. Posteriormente, o DCC pode ser usado para encapsular os objetos, ou mesmo OAI-ORE com um parser RDF, em um serviço de integração de bibliotecas digitais, proporcionando o pareamento entre impressões digitais de uma cena de crime e impressões digitais armazenadas em uma base, ou o agrupamento das evidências para convencer o júri sobre um suspeito. Outros serviços de bibliotecas digitais integradas podem considerar diferentes versões do objeto (como a composição de distorções em uma imagem, por exemplo), a correspondência de versões em ordem de proveniência, ou a publicação e busca em um processo de integração

de uma biblioteca digital.

Para o processo de publicação de dados, o *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH) pode ser usado, definindo um mecanismo para os fornecedores de dados exporem os seus metadados. Para disseminar o conteúdo de acordo com este protocolo, alguns passos são necessários [71]: (i) agregar os dados em um formato de dados aceito, (ii) incluir os metadados, (iii) adicionar as referências aos arquivos; e (iv) publicar o pacote. Para isso, OAI-ORE ou DCC pode ser utilizado, representando os objetos e agregações.

A complexidade do mapeamento e da atualização do processo de integração pode ser afetada por vários fatores, tais como o conhecimento do domínio da aplicação, o número de elementos no esquema local, e o tamanho da coleção [115].

No caso das tecnologias de objetos complexos, tais como Buckets e OAI-ORE, o processo de mapeamento também depende de outros fatores, tais como a agregação dos componentes, sua granularidade, o vocabulário que cada tecnologia está usando, como os componentes são identificados e estruturados, ou como eles são organizados em um esquema.

Em resumo, nosso estudo de caso explorou duas etapas do processo de integração: a “descoberta” de cada sistema, bem como a identificação dos itens individuais das possíveis agregações. Para isso foi utilizado o Arcabouço 5S juntamente com as tecnologias de CO para analisar a biblioteca digital integrada, a partir dos conceitos de identidade, componentes, estrutura e delimitação do objeto. Finalmente, discutimos como os componentes podem ser acessados, juntamente com seus metadados individuais.

5.3 Conclusões

Existem diversas aplicações que precisam direcionar as deficiências de manipulação e gestão, definindo e integrando os objetos complexos e os seus serviços em bibliotecas digitais. Exemplos incluem a integração de bibliotecas digitais distintas ou a gestão de vários tipos de informação.

Este capítulo discutiu como objetos complexos podem ser combinados com o Arcabouço 5S e as tecnologias de CO. Uma vez que existem várias dimensões de informações relacionadas, nós nos concentramos em quatro aspectos das COs: identidade, estrutura, componentes, e delimitação do objeto. Nossa contribuição é analisar as tecnologias DCC, Buckets e OAI-ORE, investigando suas semelhanças, além de explorá-los sob a perspectiva do Arcabouço 5S. Finalmente, apresentamos um estudo de caso utilizando uma biblioteca digital de impressões digitais para ilustrar os benefícios da nossa abordagem de modelagem.

Existem vários tópicos de pesquisa que podem ser exploradas para ampliar ainda mais nosso trabalho atual. Eles incluem o estudo do impacto das COs em outros aspectos do Arcabouço 5S, e o uso de COs em outros domínios (como os sistemas de informação de biodiversidade) e os serviços específicos (como a recuperação de imagens baseado em conteúdo e anotação).

Capítulo 6

Uma Infraestrutura baseada em DCC-OAI para Objetos Complexos de Imagem

Objetos complexos têm sido usados para reutilizar, integrar e unificar recursos diferentes sob uma mesma perspectiva, e assim facilitar a agregação e gerenciamento de recursos heterogêneos. Em particular, objetos complexos de imagem desempenham um papel fundamental em diferentes domínios, devido à sua grande disponibilidade e integração com bancos de dados, metadados e softwares de manipulação de imagem. Aplicações que gerenciam estes dados ainda carecem de mecanismos para processamento e gerenciamento, tanto no nível de componentes individuais quanto componentes integrados em um CO [8, 104]. Exemplos de serviços que precisam de apoio nestas aplicações incluem: (i) a busca de imagens por conteúdo; (ii) e a publicação dos objetos complexos, para intercâmbio e reuso de dados e serviços. Este capítulo apresenta a implementação de uma infraestrutura para estes dois serviços, baseada no conceito de objeto complexo para a integração dos serviços como componentes. Esta infraestrutura baseia-se uma tecnologia de componentes específica - *Digital Content Component* (DCC) - para agregar o objeto complexo de imagem e encapsular os processos relacionados a busca de imagem por conteúdo. Posteriormente esta estrutura é integrada com *Open Archives Initiative* para permitir a publicação de metadados e facilitar a exploração destes componentes por outras aplicações.

6.1 Visão Geral da Infraestrutura Proposta

Os recursos disponíveis em bibliotecas digitais podem variar de acordo com vários parâmetros, como seu tipo de mídia (texto, áudio, vídeo, imagem, etc.), tipo semântico (artigo, vídeo, documento, livro, software, etc.), ou formato (XML, MP3, DOC, EXE, PDF, etc.) [65]. Note ainda que cada recurso pode ter permissão específica (como no caso de direitos legais de teses, apresentado no Capítulo 3). Em especial, as aplicações em bibliotecas digitais carecem de suporte específico devido à grande integração e intercâmbio de seus recursos. Neste contexto, COs

podem ser utilizados para integrar e combinar dados heterogêneos em um objeto lógico único. A Figura 6.1 apresenta a arquitetura para uma biblioteca digital baseada em COs. A camada inferior possui as diferentes fontes de dados, tipos de mídia, tipos semânticos e formatos. As fontes de dados são agregadas em objetos complexos, que podem ser posteriormente acessados por serviços distintos (como processamento, encapsulamento, publicação, navegação e busca). Estes serviços são então acessados por aplicações de bibliotecas digitais (que podem requerer acesso tanto aos componentes individuais quanto ao CO como entidade única).

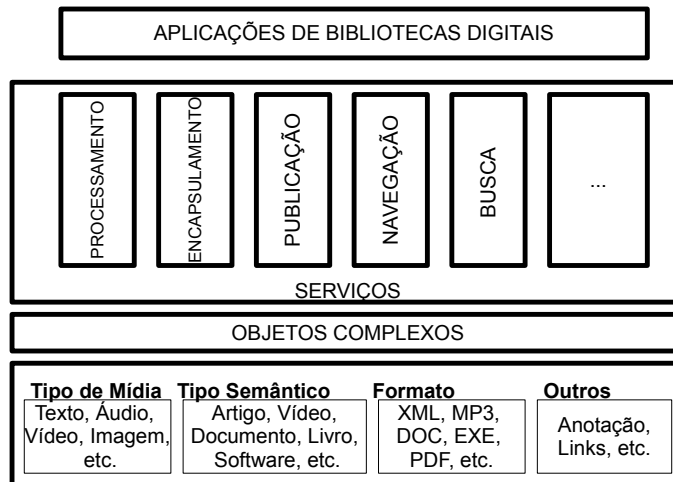


Figura 6.1: Arquitetura de uma biblioteca digital baseada em COs.

Esta seção apresenta uma visão geral da infraestrutura proposta. Como fontes de dados utilizaremos arquivos texto, imagens, software e seus respectivos metadados. Além disso, também serão tratadas as informações resultantes do processo CBIR. Estes objetos digitais são agregados em objetos complexos de imagem, utilizando os serviços de busca de imagens por conteúdo (focando em processamento e encapsulamento dos COs) e publicação dos objetos complexos.

6.1.1 A Busca de Imagens por Conteúdo

Utilizando o *Digital Content Component* para Objetos Complexos de Imagem

Considere a descrição dos DCCs apresentada na Seção 2.5.1. Nesta seção apresentaremos os três principais DCCs desenvolvidos na nossa solução (ImageDCC, ImageCODCC, e DescriptorDCC), com o objetivo de exemplificar como o encapsulamento e composição foram explorados.

O ImageDCC é um DCC passivo simples, responsável por encapsular uma imagem. Este DCC atua como uma camada de interoperabilidade (retornando informações como o nome do arquivo e localização). Como exemplo, considere a Figura 6.2-a, com a imagem *06_ancyllostoma.jpg*.

A figura representa o encapsulamento da imagem, junto com as funções e metadados relacionados. Este DCC tem quatro subdivisões distintas:

- **(a) conteúdo:** a imagem *06_ancylostoma.jpg*;
- **(b) estrutura:** a estrutura delimitando a imagem;
- **(c) interface:** a interface do ImageDCC com a operação *GetImage*;
- **(d) metadados:** descrição sobre a imagem cujo nome é *06_ancylostoma*.

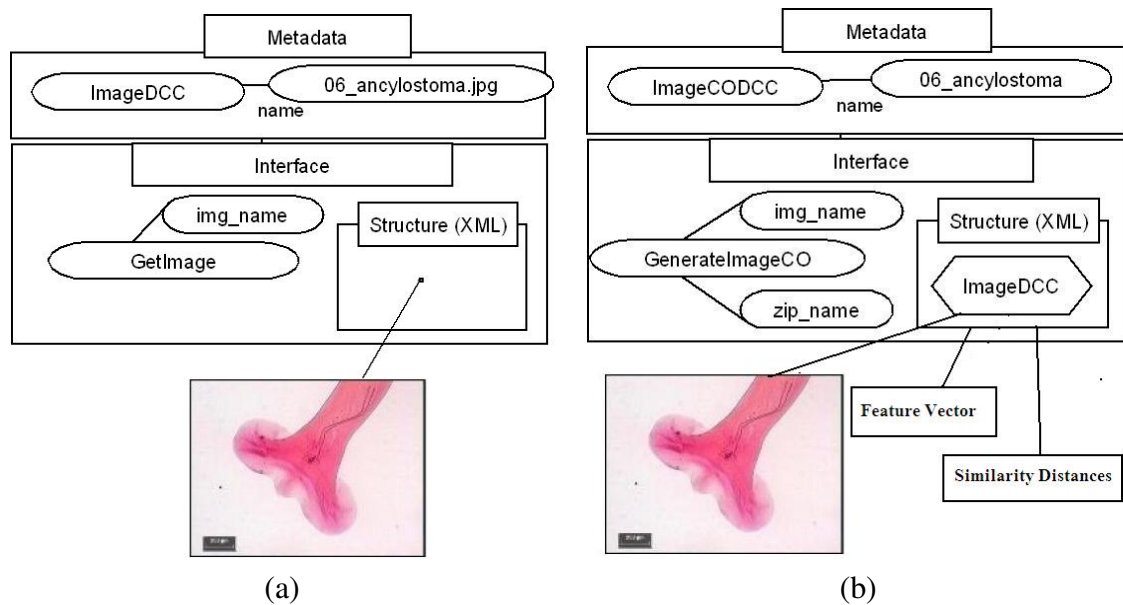


Figura 6.2: (a) Estrutura do ImageDCC. (b) Estrutura do ImageCODCC.

O ImageCODCC é também um exemplo de DCC passivo (Figura 6.2-b), porém pode encapsular vários itens que compõem o objeto complexo de imagem. No exemplo, as quatro subdivisões distintas representam:

- **(a) conteúdo:** o ImageDCC (outro DCC) encapsulando a imagem *06_ancylostoma.jpg*, o vetor de características (resultante da codificação das características da imagem por um algoritmo de extração), e as medidas de distância (que fornece o grau de semelhança para um dado par de imagens);
- **(b) estrutura:** a estrutura relaciona a imagem contida no ImageDCC, o vetor de característica, e as medidas de distância em um arquivo XML;
- **(c) interface:** oferece operações para acessar o vetor de características, as medidas de distância e o ImageDCC;
- **(d) metadados:** descrição sobre o processo de busca de imagens por conteúdo.

```

<?xml version="1.0" encoding="UTF8"?>
<image>01_ancylostoma.jpg
  <image_name>01_ancylostoma.jpg</image_name>
  <image_path>/home/nadiapk/data/uploads/01_ancylostoma.jpg</image_path>
  <image_feature_vector_name>/data/fv/01_ancylostoma.txt
</image_feature_vector_name>
  <image_descriptor>Bic
    <image_name>01_ancylostoma.jpg<image_dist_value>0
    </image_dist_value></image_name>
    <image_name>13_clonorchis.jpg<image_dist_value>80
    </image_dist_value></image_name>
    <image_name>25_echinoc_granul_5x.jpg<image_dist_value>85
    </image_dist_value></image_name>
    <image_name>09_cerccarat.jpg<image_dist_value>88
    </image_dist_value></image_name>
    <image_name>16_Paramphystomidae.jpg<image_dist_value>88
    </image_dist_value></image_name>
  </image_descriptor>
</image>
</image>

```

Figura 6.3: Exemplo de uma estrutura XML para o objeto complexo de imagem.

A Figura 6.3 apresenta o arquivo XML utilizado na estrutura para armazenar as medidas de distância, e os elementos do ICO. Note que esta estrutura é diferenciada da estrutura do DCC, já que não informa o encapsulamento do ImageDCC.

Enquanto o primeiro DCC (ImageDCC) provê o encapsulamento da imagem, o segundo (ImageCODCC) provê o encapsulamento de vários componentes resultantes do processo CBIR. O terceiro DCC (DescriptorDCC) é um DCC de processo, responsável por encapsular o descritor (o algoritmo de extração que codifica as características da imagem em vetores de características). O DescriptorDCC possui quatro divisões distintas (Figura 6.4):

- **(a) conteúdo:** a composição de ExtractionDCC e DistanceDCC como ilustra a Figura 6.4;
- **(b) estrutura:** a estrutura delimita e conecta o ExtractionDCC e DistanceDCC;
- **(c) interface:** compreende operações derivadas dos dois componentes encapsulados;
- **(d) metadados:** a união dos metadados relacionados às respectivas operações do ExtractionDCC e DistanceDCC.

No exemplo, o DescriptorDCC encapsula o descritor de cor *Border/Interior pixel Classification* (BIC) [118]. O BIC é um descritor compacto e eficiente, apropriado para coleções heterôgeneas de imagens. O DescriptorDCC é composto por duas funções: Extração e Distância. A Extração origina-se no ExtractionDCC, um DCC de processo que encapsula a biblioteca para a extração do vetor de característica. A operação Extração recebe um caminho de uma imagem e gera o respectivo vetor de característica. No caso do descritor BIC, a extração do vetor de característica inicialmente caracteriza o pixel como *borda* ou *interior*, e então computa dois histogramas de cor, um para cada tipo de pixel. Um pixel é classificado como *borda* se ao menos um dos seus quatro vizinhos tem uma cor diferente. Se os quatro vizinhos têm a mesma

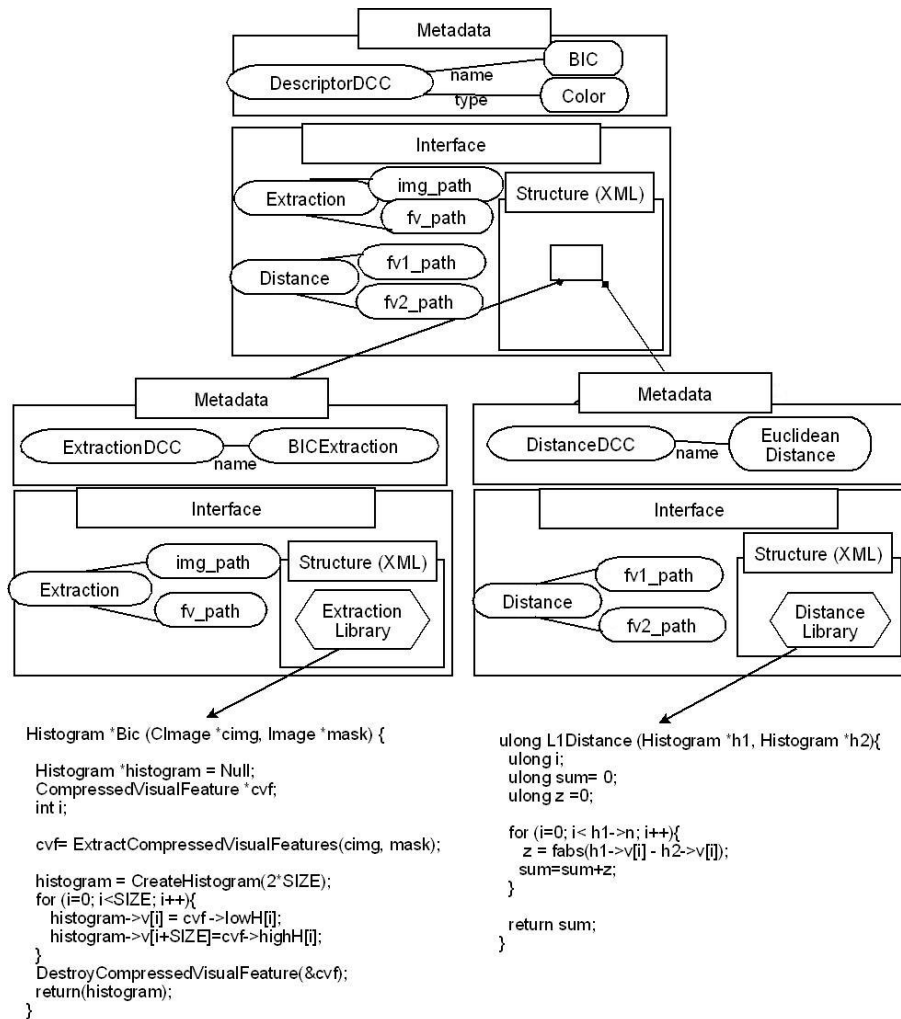


Figura 6.4: Estrutura do DescriptorDCC.

cor, é classificado como interior. A função Distância origina-se no DistanceDCC, outro DCC de processo que encapsula a biblioteca para a medida de distância (no caso do descritor BIC, uma função de distância chamada L1). A operação de Distância L1 recebe dois vetores de características e retorna a distância entre eles (é ilustrada no canto inferior direito da Figura 6.4).

A Infraestrutura para Busca de Imagens por Conteúdo

Outros DCCs foram criados para encapsular o processo de busca de imagens por conteúdo, o gerenciamento do banco de dados, e as configurações básicas. A taxonomia dos DCCs para a manipulação de objetos complexos de imagem é apresentada na Figura 6.5-a, dividindo os DCCs em Passivos e de Processo.

O DistanceDCC e o ExtractionDCC são DCCs de processo que compõem o DescriptorDCC.

A coleção de DescriptorDCCs é manipulada pelo DCC de processo DescriptorLibraryDCC. Seu esquema é similar ao DescriptorDCC, porém com uma estrutura adaptada para tratar múltiplos descritores. Suas operações incluem a consulta de descritores disponíveis, e a inclusão/exclusão de descritores. O DescriptorLibraryDCC é particularmente útil quando uma aplicação precisa manipular vários descritores ao mesmo tempo.

O CBIRProcessDCC oferece operações de manipulação de alto nível na busca de imagens por conteúdo. Estas operações são responsáveis pela extração de todos os vetores de características em uma coleção de imagens, além do cálculo da medida de distância entre cada par de imagens. O CBIRProcessDCC é particularmente útil quando uma aplicação precisa utilizar busca por conteúdo, sem se preocupar com os seus parâmetros e detalhes específicos.

O ManagerDCC oferece operações básicas do ambiente, permitindo a configuração de diretórios e variáveis utilizadas para o acesso dos arquivos fontes e arquivos temporários.

Finalmente, o ImageCODCC, RetrievalDCC, e ImageDCC são utilizados para o encapsulamento de fonte de dados. O ImageCODCC é um DCC passivo, encapsulando a imagem com os arquivos resultantes do processo CBIR (vetor de características e a medida de distância) em um arquivo zip.

O vetor de característica pode ser um arquivo texto ou binário, criado pelo DescriptorDCC. O DCC passivo RetrievalDCC atua como uma camada de interoperabilidade, acessando e recuperando dados do banco de dados.

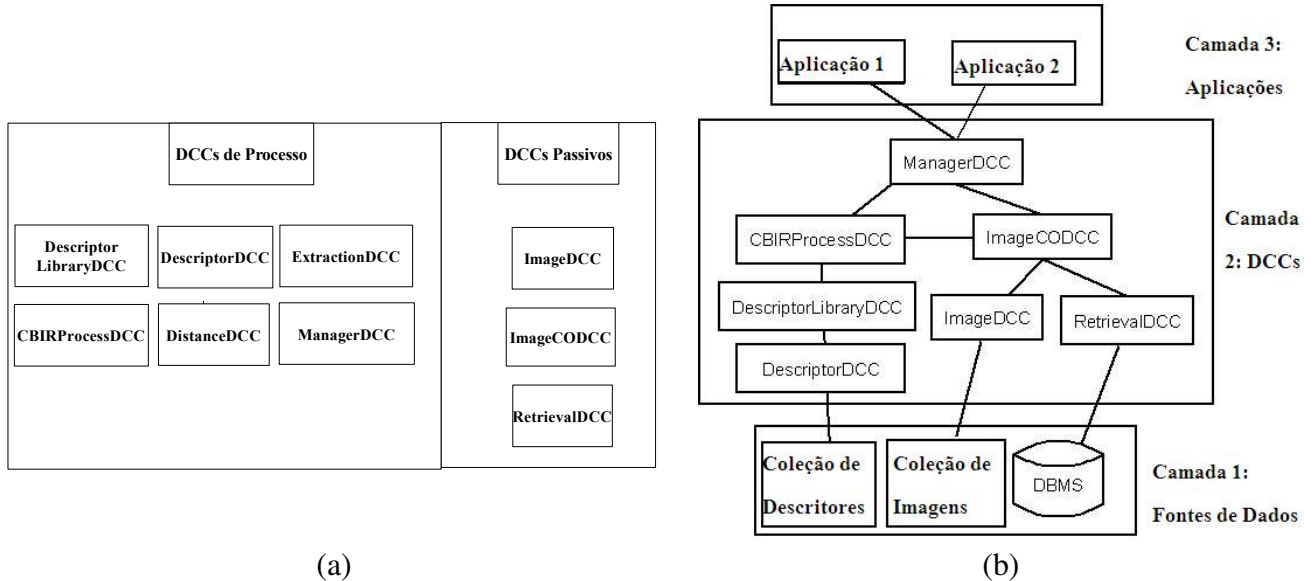


Figura 6.5: (a) A Taxonomia dos DCCs para a manipulação de ICOs. (b) As camadas de gerenciamento.

A Figura 6.5-b mostra uma visão geral da arquitetura da infraestrutura proposta, com a camada 2 utilizando os DCCs da Figura 6.5-a. A camada inferior contém as fontes de dados:

a biblioteca de descritores, a coleção de imagens, e o banco de dados. A segunda camada contém os DCCs que provêm o acesso e a manipulação dos dados. A terceira camada apresenta as aplicações que manipulam os objetos complexos de imagens, o acesso a recuperação por conteúdo, ou até as fontes de dados. Aplicações diferentes podem requerer visões diferentes das informações, processos ou componentes, resultando em um acesso distinto de DCCs.

Os DCCs passivos e de processo, mesmo tendo propósitos diferentes, encapsulam recursos com regras de acesso, metadados descritivos, e estrutura. Uma das vantagens exploradas aqui reside no encapsulamento de dados e software, sob o mesmo esquema de interface.

Note que a mesma infraestrutura é utilizada para manipular as fontes de dados (a biblioteca de descritores, a coleção de imagens e o banco de dados) e suas funções de manipulação. As fontes de dados podem ser divididas entre dois tipos: estáveis e dinâmicas [49]. Fontes de dados estáveis são caracterizadas por atualizações eventuais, e são exemplificadas aqui pela biblioteca de descritores (DescriptorDCC e DescriptorLibraryDCC). Já que cada descritor pode ser implementado em uma linguagem diferente, o seu encapsulamento facilita o acesso homogêneo por outras aplicações e ambientes. Fontes de dados dinâmicas são sempre atualizadas, e são representadas na infraestrutura pela coleção de imagens e o banco de dados. Como várias coleções de imagens podem ser utilizadas e modificadas na aplicação, os DCCs são utilizados como uma camada de interoperabilidade para gerenciá-las, junto com seus respectivos metadados e arquivos temporários gerados através da recuperação por conteúdo.

Cada DCC que manipula as fontes de dados pode explorar diferentes métodos em sua interface. Entretanto, o gerenciamento dos processos em alto nível é necessário, e é representado pelos seguintes DCCs: CBIRProcessDCC, ImageCODCC e ManagerDCC.

6.1.2 A Publicação de Objetos Complexos

O protocolo OAI-PMH tem sido amplamente adotado por muitas bibliotecas digitais e repositórios institucionais para adquirir e realizar intercâmbio de recursos. Em nossa abordagem, usaremos o protocolo OAI-PMH para outros aplicativos “descobrirem os diferentes objetos digitais disponíveis em um repositório”.

A Figura 6.6 ilustra os diferentes objetos disponíveis para publicação: a imagem, o objeto complexo de imagem (com vários componentes de imagem, incluindo o vetor de características), o descritor, e a biblioteca de descritores. Note que a publicação dos COs permite tanto a exploração dos componentes individuais (a imagem, o descritor), quanto as agregações (o ICO, a biblioteca de descritores). Estes objetos podem ser diferenciados com a requisição OAI-PMH *ListSets*. Na abordagem apresentada, além dos metadados textuais (nome da imagem, localização do arquivo, etc.), também estão disponíveis os metadados visuais (medidas de distância, vetores de característica) relacionados a outras imagens da coleção. Note que mesmo havendo recursos de tipos diferentes, todos utilizaram a mesma infraestrutura do DCC para

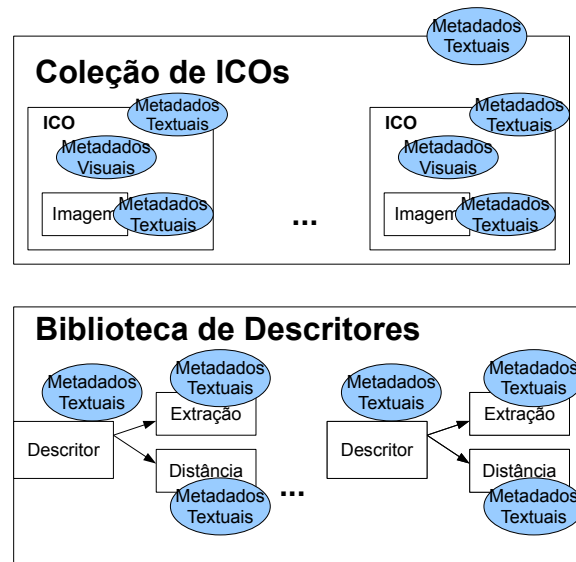


Figura 6.6: Os objetos digitais disponíveis para publicação.

manipular as imagens, seus metadados, software, e funções de agregação.

Como uma alternativa para os usuários, duas requisições (“verbos”) específicos foram implementados, para diferenciar e agilizar a publicação dos metadados relativos a recuperação por conteúdo:

1. *ListFeatureVectors*: Este verbo é utilizado para recuperar a lista de vetores de características para todas as imagens disponíveis na coleção. O único argumento requerido é o nome do descritor utilizado na criação do vetor de características. Na Figura 6.7, por exemplo, é ilustrado a resposta da requisição OAI *ListFeatureVectors* para a imagem *ancylostoma_caninum.jpg*. Os vetores de características podem ser utilizados, por exemplo, para indexação e busca de imagens, considerando suas propriedades visuais.
2. *GetFeatureVector*: Este verbo é utilizado para recuperar os vetores de características. Os argumentos incluem o objeto complexo de imagem, o descritor, e o nome da imagem.

Em uma abordagem inicial, o OAI-PMH foi adotado (ao invés do OAI-ORE [65]) por duas razões: há mais aplicações relacionadas, e a estrutura do objeto complexo poderia ser igualmente mapeada em um arquivo XML (ao invés do *named graph* presente no OAI-ORE).

6.2 Estudo de Caso

Esta seção apresenta o uso de ICOs em um estudo de caso da aplicação SuperIDR (apresentada no Capítulo 3), exemplificando as etapas de encapsulamento, processo e publicação dos ICOs,

```

<OAI-PMH xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/
http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
  <responseDate>2010-12-27T20:49:56Z</responseDate>
  <request verb="ListFeatureVectors" metadataPrefix="zip">
    http://localhost:8180/oai/provider</request>
  -
  <ListFeatureVectors>
    -
    <header>
      <feature_vector::ImageName>Ancylostoma_caninum.jpg</feature_vector::ImageName>
      <feature_vector::DescriptorName>ContourMSFractalDimension</feature_vector::DescriptorName>
    <feature_vector::X>
      <feature_vector::value>0.9509999999999999</feature_vector::value>
      <feature_vector::value>1.0009999999999999</feature_vector::value>
      <feature_vector::value>3.9509999999999999</feature_vector::value>
      (...)
      <feature_vector::value>0.9509999999999999</feature_vector::value>
      <feature_vector::value>0.9509999999999999</feature_vector::value>
      <feature_vector::value>0.9509999999999999</feature_vector::value>
      <feature_vector::value>0.9509999999999999</feature_vector::value>
      <feature_vector::value>0.9509999999999999</feature_vector::value>
    </feature_vector::X>
  </header>
</ListFeatureVectors></OAI-PMH>
</OAI-PMH>

```

Figura 6.7: A resposta da requisição OAI *ListFeatureVectors*

incluindo metadados, imagens, e software.

6.2.1 A Implementação

A infraestrutura proposta permite a instalação de descritores distintos, mas para os testes apresentados, o descritor *Border/Interior pixel Classification* (BIC) [118] foi utilizado. A biblioteca foi implementada em linguagem C, e os DCCs em Java. As funções e parâmetros de cada DCC, além dos dados da aplicação foram armazenados em um banco de dados PostgreSQL.

As imagens, ICOs, e descritores são publicados utilizando o software jOAI¹. Este provedor permite que arquivos XML possam ser expostos como itens de dados em um repositório OAI, e disponibilizados para a utilização de outras aplicações, usando o protocolo OAI-PMH. Após apontar o software para um ou mais diretórios de arquivos, o software monitora os arquivos XML, permitindo adicionar, atualizar ou excluir arquivos XML assim que os arquivos são adicionados, atualizados ou excluídos dos diretórios. O provedor jOAI é usado para recuperar registros de metadados de provedores de dados remotos, e salvar no sistema de arquivos local. Além disso, os registros que tenham sido coletados são empacotados em arquivos zip que podem ser baixados e abertos através de sua interface web.

¹http://www.dlese.org/dds/services/joai_software.jsp (último acesso em 05/05/11).

6.2.2 O Encapsulamento de Recursos

Esta seção descreve dois tipos de DCCs implementados neste estudo de caso: o DescriptorDCC e o SpeciesDCC.

O DescriptorDCC (Figura 6.4) centraliza as operações originárias do ExtractionDCC e DistanceDCC. Ambos os DCCs manuseiam bibliotecas binárias em linux, desenvolvidas na linguagem C. Existe um DescriptorDCC para cada descritor adicionado a infraestrutura. Os metadados do DCC contêm o nome do descritor e seu tipo (cor, textura, forma).

O SpeciesDCC permite a agregação de todos os dados de uma espécie, neste caso, uma coleção de ImageCODCCs. Como cada espécie pode ter tipos diferentes de imagens (focando em uma parte específica do corpo, como mostra a Figura 3.4), este DCC é particularmente útil quando um usuário precisa do gerenciamento em alto nível de todos os arquivos XML, vetores de característica, e imagens para uma determinada espécie. Suas operações incluem o número de ICOs agregados, o nome da espécie, e sua respectiva, família e gênero.

6.2.3 A Infraestrutura

A implementação da infraestrutura para a geração do ICO foi dividida nas seguintes fases: (i) a “descoberta” e definição de cada parte do CO; (ii) a identificação das partes pertencentes ao CO; (iii) o processo CBIR; (iv) o encapsulamento da imagem e os metadados relacionados; e (vi) a publicação dos objetos (imagem, ICO, e descritores).

A aplicação tem imagens e metadados disponibilizados por professores. As imagens (Figura 3.4) foram criadas a partir da análise em microscópio e os metadados extraídos de livros especializados no assunto. Os metadados incluem vários detalhes das espécies, como família, gênero, classe, hábitos, etc. Uma espécie pode ter várias imagens, já que podem referenciar diferentes fases ou partes do corpo de um parasita.

Na fase dois (a identificação das partes pertencentes ao CO) foi definida que a agregação seria realizada a nível de espécie. Para isso, foi criado o SpeciesDCC (Figura 6.9), utilizando ImageCODCCs. O banco de dados foi utilizado para a identificação das partes do CO, relacionando imagens às suas respectivas espécies, metadados e informações CBIR.

O processo CBIR compreende a execução de um algoritmo de extração em uma coleção de imagens, seguido pela comparação das medidas de distância. Neste processo outras partes do ICO tornam-se disponíveis, como o vetor de característica e as medidas de distância.

Como exemplo, considere a espécie *Ancylostoma caninum*. A espécie tem metadados e duas imagens representando partes diferentes de um parasita: *01_ancylostoma.jpg* (cabeça) e *06_ancylostoma.jpg* (corpo).

O processo CBIR da imagem *01_ancylostoma.jpg* gera um vetor de características e as medidas de distância relacionadas a outras imagens da coleção. A Figura 6.8-a ilustra o *ranking* para a imagem *01_ancylostoma.jpg*, considerando o descritor de cor. As cinco figuras da parte

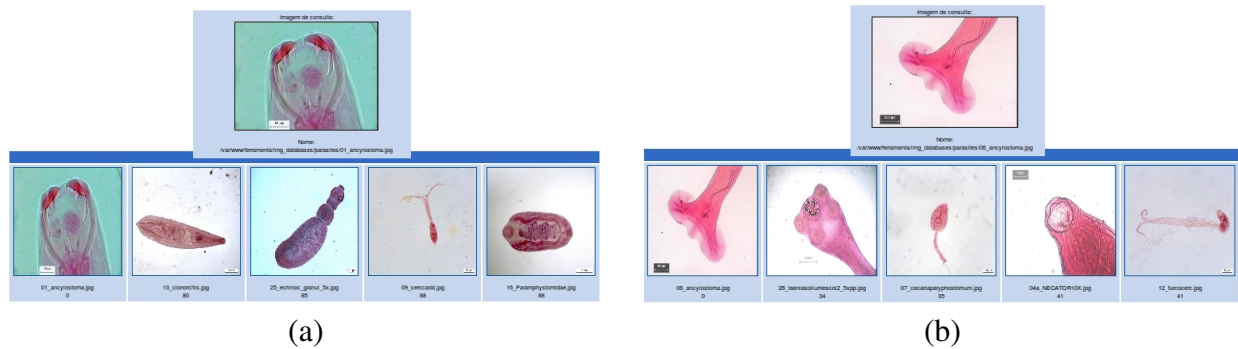


Figura 6.8: (a) Resultado da consulta CBIR para a imagem *01_ancylostoma.jpg*. (b) Resultado da consulta CBIR para a imagem *06_ancylostoma.jpg*.

inferior (*01_ancylostoma.jpg*, *13_clonorchis.jpg*, *25_echinoc_granul_5x.jpg*, *09_cerccarat.jpg* e *16_Paramphystomidae.jpg*) são as mais similares se comparadas com *01_ancylostoma.jpg*.

O processo CBIR da segunda imagem (*06_ancylostoma.jpg*) gera um segundo vetor de características, e um segundo conjunto de medidas de distância. A Figura 6.8-b apresenta o *ranking* para *6_ancylostoma.jpg*, considerando a comparação baseada em cor.

Posteriormente os dois ICOs (*06_ancylostoma.jpg* e *01_ancylostoma.jpg*) são agregados por espécie (*Ancylostoma caninum*) no SpeciesDCC, como mostra a Figura 6.9. O SpeciesDCC agrega toda a informação (metadados, imagem, vetor de característica, medidas de distância) relacionada a espécie. O SpeciesDCC é composto por um ou vários ImageCODCCs, que por sua vez encapsulam o ImageDCC. O CO resultante contém os arquivos de dois ICOs com o respectivo metadado da espécie e XML. A Figura 6.10-a apresenta o XML para esta agregação: o bloco inicial apresenta os metadados da espécie (nome, família, gênero, e classe), o segundo bloco apresenta as informações para o primeiro ICO, e o último bloco apresenta o XML para o segundo ICO. Note que cada ICO começa com a tag `<image>`, seguida por metadados e a medidas de distância de cinco imagens.

Na fase final da implementação da infraestrutura (a publicação dos objetos), os seguintes objetos são publicados: 17 ICOs de espécie, 76 imagens, 76 arquivos de descritores, e seus respectivos metadados com seus 17 arquivos XML.

Os ICOs são publicados com o software jOAI. A Figura 6.10-b ilustra o resultado após a operação *ListIdentifiers*. Como alternativa, os usuários também podem utilizar as requisições *ListFeatureVectors* e *GetFeatureVector* (Figura 6.7).

6.3 Conclusões

Este capítulo apresentou a especificação e implementação de uma infraestrutura baseada em CO para CBIR, permitindo o processamento, encapsulamento e publicação de objetos complexos

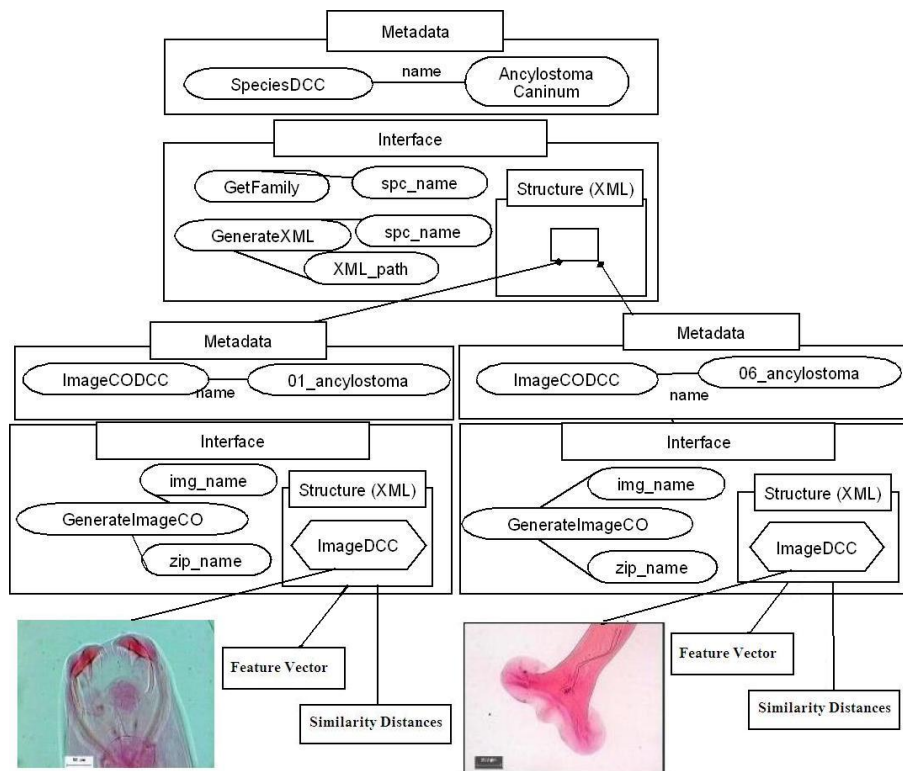


Figura 6.9: Estrutura para SpeciesDCC.

de imagem.

A infraestrutura apresentada baseia-se em uma tecnologia de componentes específica - *Digital Content Component* (DCC) - para encapsular o processo CBIR, e agregar o ICO. Outras contribuições incluem a integração desta infraestrutura com o protocolo *Open Archives Initiative* para suportar a publicação e “descoberta” de objetos digitais por outras aplicações. Parte desta infraestrutura foi implementada e validada em um estudo de caso utilizando imagens de parasitas.

Entre os benefícios de nossa infraestrutura, podemos citar:

- o gerenciamento uniforme da coleção de imagens, metadados, e descritores, a partir de uma infraestrutura baseada em CO;
- a disponibilidade de integração de CBIR em outros domínios, já que a interface do DCC provê uma interface homogênea para o acesso de software;
- as aplicações podem manipular diretamente os DCCs, sem se preocupar com configurações, arquivos temporários ou detalhes;
- a flexibilidade apresentada pela infraestrutura baseada em CO para combinar componentes e serviços (direcionando desafios já enfrentados na busca individual e integrada dos

```

<?xml version="1.0" encoding="UTF8"?>
<species>Ancylostoma caninum
</species>
<species_family>Ancylostomatidae</species_family>
<species_genus>Ancylostoma</species_genus>
<species_class>Nematoda</species_class>

<image>01_ancylostoma.jpg
<image_name>01_ancylostoma.jpg</image_name>
<image_feature_vector_name>/home/nadiapk/data/fv/01_ancylostoma.txt</image_feature_vector_name>
<image_descriptor>Bic
  <image_name>01_ancylostoma.jpg<image_dist_value>0</image_dist_value></image_name>
  <image_name>13_clonorchis.jpg<image_dist_value>80</image_dist_value></image_name>
  <image_name>25_echinoc_granul_5x.jpg<image_dist_value>85</image_dist_value></image_name>
  <image_name>09_cercarar.jpg<image_dist_value>88</image_dist_value></image_name>
  <image_name>16_Paramphystomidae.jpg<image_dist_value>88</image_dist_value></image_name>
</image_descriptor>
</image>

<image>06_ancylostoma.jpg
<image_name>06_ancylostoma.jpg</image_name>
<image_feature_vector_name>/home/nadiapk/data/fv/06_ancylostoma.txt</image_feature_vector_name>
<image_descriptor>Bic
  <image_name>06_ancylostoma.jpg<image_dist_value>0</image_dist_value></image_name>
  <image_name>26_tacniasoliumescol2_5xpp.jpg<image_dist_value>34</image_dist_value></image_name>
  <image_name>07_cecariaparyphostomum.jpg<image_dist_value>35</image_dist_value></image_name>
  <image_name>04a_NECAATOR10X.jpg<image_dist_value>41</image_dist_value></image_name>
  <image_name>12_furcocerc.jpg<image_dist_value>41</image_dist_value></image_name>
</image_descriptor>
</image>
</species>

```

```

<OAI-PMH xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/
http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
<responseDate>2010-12-27T20:49:56Z</responseDate>
<request verb="ListIdentifiers" metadataPrefix="zip">
http://localhost:8180/oai/provider</request>
</ListIdentifiers>
<header>
<identifier>Ancylostoma caninum</identifier>
<timestamp>2010-11-27T20:43:10Z</timestamp>
<identifier>Ancylostoma duodenale</identifier>
<timestamp>2010-11-27T20:44:10Z</timestamp>
<identifier>Ancylostoma braziliense</identifier>
<timestamp>2010-11-27T20:45:10Z</timestamp>
<identifier>Fasciola hepatica</identifier>
<timestamp>2010-11-27T20:46:10Z</timestamp>
<identifier>Entamoeba histolytica</identifier>
<timestamp>2010-11-27T20:47:10Z</timestamp>
<identifier>Iodamoeba butchlii</identifier>
<timestamp>2010-11-27T20:48:10Z</timestamp>

```

(a)
(b)

Figura 6.10: (a) XML para o CO *Ancylostoma caninum* . (b) A requisição OAI *ListIdentifiers* para os ICOs de Parasitas.

componentes de um CO [8, 104]);

- o intercâmbio e disponibilidade de informações que estavam disponíveis anteriormente somente como recursos estáticos (como os livros que continham as informações de parasitas);
- a publicação de imagens, ICOs, e descritores, com metadados visuais e textuais; e
- a centralização do processamento, encapsulamento, publicação e coleta de ICOs em somente uma infraestrutura.

Trabalhos futuros incluem a inclusão de outros processos (como anotação e estabelecimento de links externos), a combinação de descritores, o uso de OAI-ORE, e o uso da aplicação em outros domínios.

Capítulo 7

Cenários de Reutilização da Infraestrutura Proposta

Neste capítulo descrevemos a reutilização da infraestrutura apresentada no capítulo anterior em três contextos (Figura 7.1): (i) a integração de bibliotecas digitais de impressões digitais; (ii) a busca multimodal; e (iii) a busca visual na sumarização de estórias em vídeo.

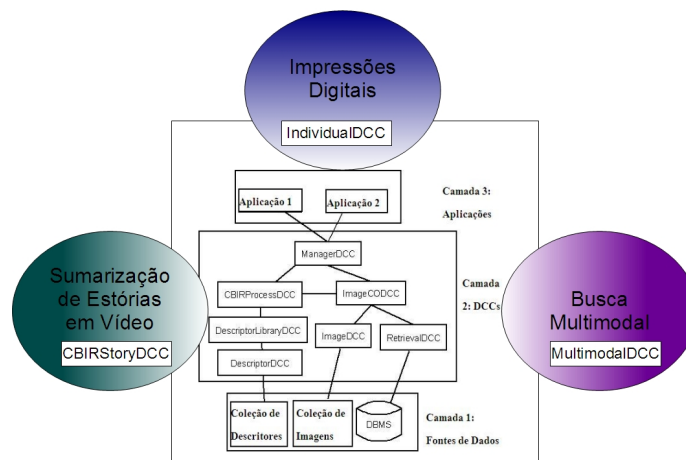


Figura 7.1: A reutilização da infraestrutura baseada em DCC e OAI-PMH.

7.1 A Integração de Bibliotecas Digitais de Impressões Digitais

Considere o estudo de caso descrito no Capítulo 3, sobre a integração de bibliotecas digitais de impressões digitais. Reutilizamos a infraestrutura apresentada no capítulo anterior, e criamos novos DCCs para representar as novas agregações necessárias neste contexto.

O primeiro tipo de Biblioteca Digital (DL1) está associada com o armazenamento de impressões digitais em bancos de dados. O segundo domínio está relacionado com a criação de materiais de treinamento para examinadores de impressões digitais (DL2). Um terceiro tipo de biblioteca digital se refere às provas e dados que descrevem uma cena de crime (DL3). O quarto e último tipo de DL está relacionado ao projeto NIJ (*National Institute of Justice*) de pesquisa da Virginia Tech - USA, implementando experimentos, análise, medidas de qualidade, e métodos de pareação em bibliotecas digitais (DL4). A combinação destas quatro bibliotecas em uma biblioteca única e integrada (como mostra a Fig. 5.1), utilizando objetos complexos como um meio de navegar e agregar estas informações, resulta em uma biblioteca digital de grande porte e interessante do ponto de vista da pesquisa.

Devido ao grande tamanho, aos tipos de variância de imagens de bases de dados de impressão digital, e ao custo computacional de algoritmos de verificação de impressões digitais, utilizamos uma etapa de pré-processamento para o protótipo, baseada em técnicas CBIR. Esta fase é responsável pela classificar imagens similares com base em um descritor de textura. O objetivo é reduzir o número de comparações um-para-um, buscando melhorias tanto em termos de precisão quanto em velocidade de recuperação. Neste sentido, estudamos a caracterização dos padrões de textura que podem ser encontrados em impressões digitais.

7.1.1 O Protótipo

A implementação do protótipo foi realizada em cinco fases: (1) a “descoberta” e definição de cada parte do CO; (2) a identificação das partes do CO; (3) o processo de CBIR; (4) o encapsulamento do imagem e metadados; e (5) a publicação do CO.

Fase 1. A descoberta e definição de cada parte do CO provenientes das 4 DLs (Capítulo 3) desempenha um papel fundamental para definir os tipos de dados e as diferentes DLs da integração de impressão digital. O objetivo do protótipo foi agregação dos dados, incluindo as imagens e metadados. Somente duas bibliotecas de impressão digital foram selecionadas para o protótipo: as imagens de impressões digitais da polícia e as impressões digitais da cena do crime.

Suponha por exemplo, um indivíduo chamado Joseph Murch, que tem suas impressões digitais registradas pela polícia (Figura 7.2), e por uma biblioteca digital que simula uma cena do crime (Figura 7.3). Todos os exemplos nas Figuras 7.2 e 7.3 pertencem ao mesmo indivíduo,

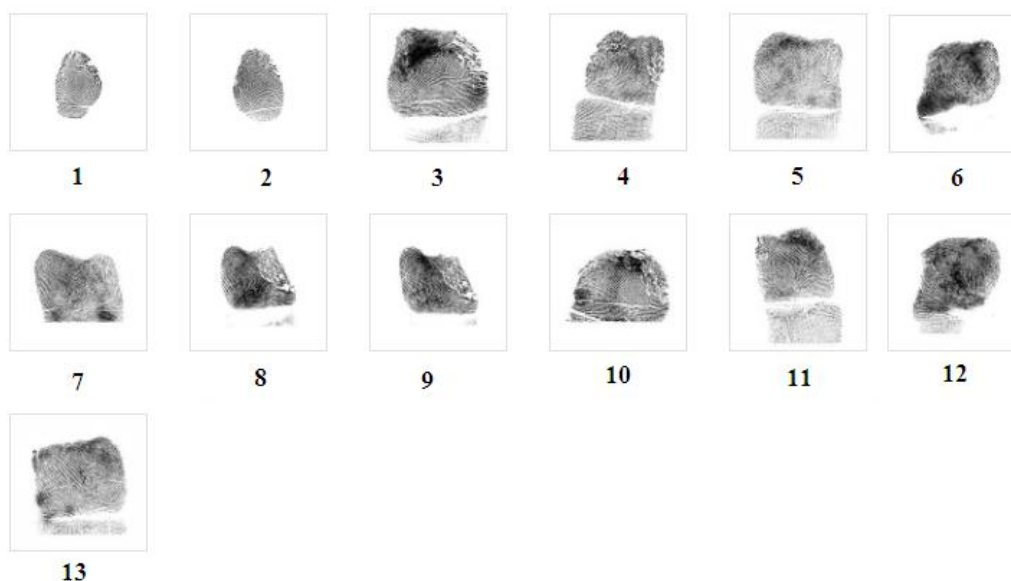


Figura 7.2: Exemplos de impressões digitais obtidas a partir do registro da polícia.

porém: (i) a Figura 7.2 - parte 4 tem mais detalhes da parte central e da parte inferior de um dedo; (ii) na Figura 7.2 - parte 7 falta uma pequena parte da parte superior da imagem; (iii) a Figura 7.3 - parte 1 tem quantidades diferentes de tinta; (iv) a Figura 7.3 - partes 2 até 6 tem regiões que estão faltando, devido à textura da superfície de qual foram retiradas; (v) a Figura 7.3 - partes 7 até 11 tem algum tipo de distorção (como rotação); (vi) a Figura 7.3 - parte 17 utiliza uma folha de alumínio como substrato e (vii) a Figura 7.3 - partes 18 a 21 tem somente parte da imagem. Outros fatores podem impactar a comparação e busca destas imagens, como a superfície da qual foram retiradas ou a composição das imagens (como ilustra a Figura 7.3 - parte 1).

Fase 2. Na fase dois definimos que a agregação seria realizada em nível de “indivíduo”. No exemplo, agregamos todas as informações e impressões digitais referentes a Joseph Murch. Para a identificação das partes do CO, utilizamos o banco de dados, relacionado as imagens às respectivas DLs, metadados, vetores de características e medidas de distância.

Fase 3. Na fase três foi realizada a integração do processo de CBIR, permitindo uma pré-categorização das imagens, utilizando o descritor de textura. Para isso o descritor *Statistical Analysis of Structural Information* (SASI) [16] foi utilizado. O processo CBIR da Figura 7.2 - parte 11, por exemplo, gera um vetor de características, e as medidas de distância para as outras imagens na coleção. A Figura 7.4 ilustra o resultado deste processo, utilizando o descritor de textura. As 10 imagens da região inferior são as mais similares se comparadas com a Figura 7.2 - parte 11.

O processo CBIR da segunda imagem (Figura 7.3 - parte 3) gera um segundo vetor de

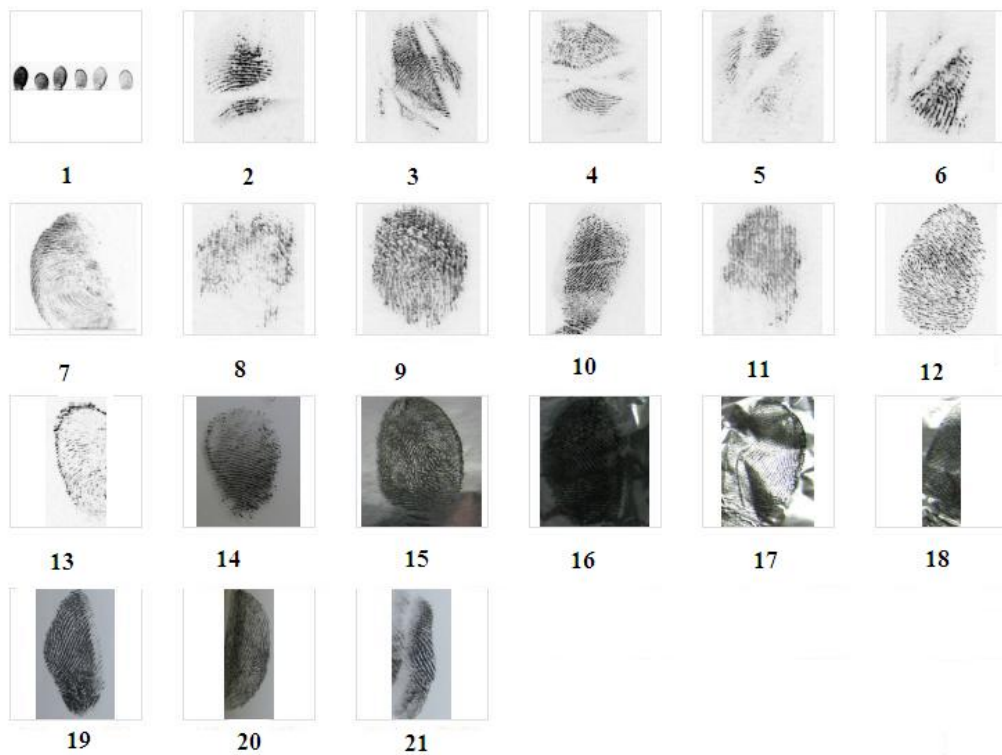


Figura 7.3: Exemplos de impressões digitais que simulam uma cena de crime.

características, e um outro conjunto de medidas de distância. A Figura 7.5 apresenta o resultado deste processo para a Figura 7.3 - parte 3.

Fase 4. O encapsulamento dos recursos foi realizado em um modelo de camadas de três níveis (ilustrado na Figura 7.6): (i) o ICO, agregando as informações provenientes do processo CBIR (encapsuladas no ImageCODCC, descrito na Seção 5.1.1); (ii) as agregações referentes ao conceito de “indivíduo” da DL da polícia (encapsulada no PoliceCODCC) e da DL de cena do crime (encapsulada no CrimeCODCC); e (iii) o CO indivíduo, agregando todas as imagens e metadados de uma mesma pessoa (encapsulado no IndividualDCC).

No exemplo mencionado, a Figura 7.2 - parte 11 e a Figura 7.3 - parte 3 foram agregadas em dois ICOs. Elas estão representadas pelo ImageCODCC, que centraliza o encapsulamento do CO, com sua respectiva imagem, arquivo XML (com os metadados e medida de distância), e o vetor de característica. Neste caso, os vetores de características eram arquivos binários.

A segunda camada contém a agregação referente ao conceito de indivíduo em cada uma das DLs. Neste caso, a Figura 7.2 - parte 11 pertence a um indivíduo da DL de polícia (encapsulada no PoliceCODCC). A Figura 7.3 - parte 3 pertence ao mesmo indivíduo, porém está localizada na DL de cena do crime (encapsulada no CrimeDCC).

A terceira camada corresponde ao CO indivíduo, utilizando DLs distintas. No exemplo

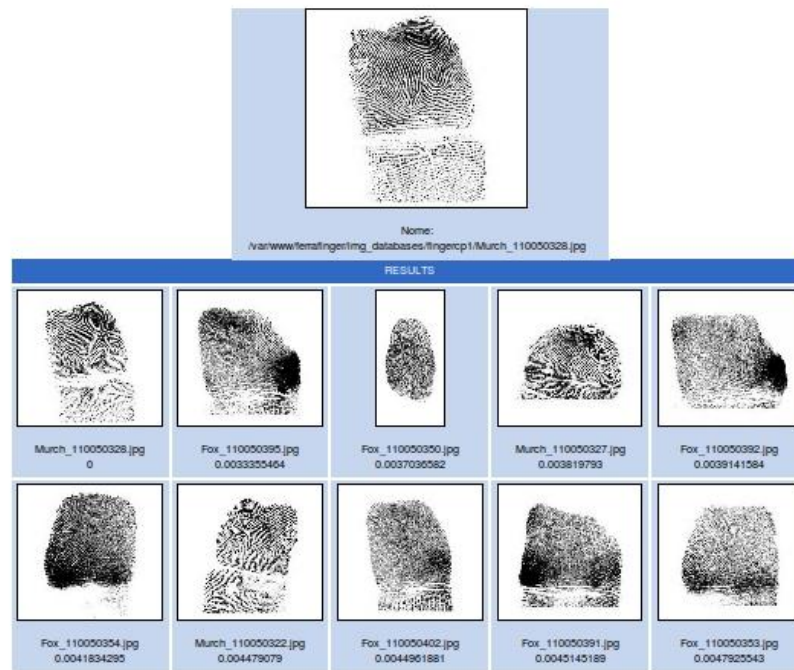


Figura 7.4: Resultado da consulta CBIR para a Figura 7.2 - parte 11.

mencionado, este CO contém todas as imagens presentes na Figura 7.3 e Figura 7.2 (já que pertencem ao mesmo indivíduo), juntamente com seus respectivos vetores de característica, medidas de distância e metadados. No exemplo mencionado, este CO é representado pelo IndividualDCC com o nome Joseph Murch. A Figura 7.7 apresenta o arquivo XML para a agregação indivíduo: o bloco inicial apresenta os metadados referentes ao indivíduo (nome, idade, sexo); o segundo bloco apresenta o XML para a DL de polícia, e o último bloco apresenta o XML para a DL de cena do crime. Note que o segundo e o terceiro bloco iniciam com a tag `<image>`.

Fase 5. Na fase cinco (publicação do CO), utilizamos a mesma infraestrutura descrita na Seção 6.1.2. Além da publicação dos objetos de imagem, ICO, e descritor (SASI [16]), também publicamos o CO indivíduo, resultante da fase quatro. A publicação do CO tem por objetivo facilitar a troca e intercâmbio de bibliotecas digitais de impressões digitais. Usuários de CBIR, podem por exemplo, reutilizar somente os objetos de imagem (CO imagem), enquanto usuários que manipulam bibliotecas de impressões digitais podem reutilizar todas as informações pertencentes ao indivíduo (CO indivíduo).

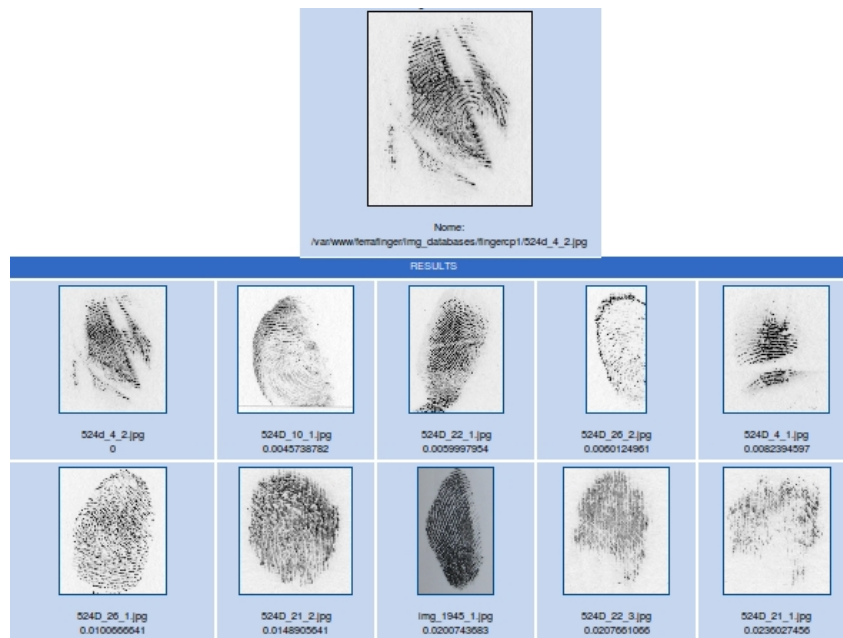


Figura 7.5: Resultado da consulta CBIR para a Figura 7.3 - parte 3.

7.1.2 A Implementação

O protótipo permite a instalação de descritores diferentes, porém para os testes apresentados, o descritor *Statistical Analysis of Structural Information* (SASI) [16] foi utilizado.

A biblioteca foi implementada na linguagem C, e os DCCs em Java. Os metadados, e as funções e os parâmetros disponíveis para cada DCC foram armazenados em um banco de dados PostgreSQL. Os objetos foram publicados com o software jOAI¹.

Como pode ser observado, foi possível utilizar toda a infraestrutura CBIR para a integração de bibliotecas digitais de imagens digitais, criando novos DCCs que as integrassem (PoliceCODCC e CrimeCODCC) e representassem a agregação das informações referentes ao indivíduo (IndividualDCC).

7.2 Busca Multimodal

Um dos principais desafios no uso de COs abrange a necessidade de não só realizar a busca individual dos componentes, mas a busca integrada (multimodal) do CO como uma unidade única [8, 104]. A busca multimodal através da composição de serviços permite estas duas modalidades: quando necessário, utiliza-se o serviço específico do componente (como a busca textual), ou combina-se vários (como a busca textual e visual) quando é necessária a busca

¹<http://www.dlese.org/ddservices/joai> (último acesso em 05/05/11).

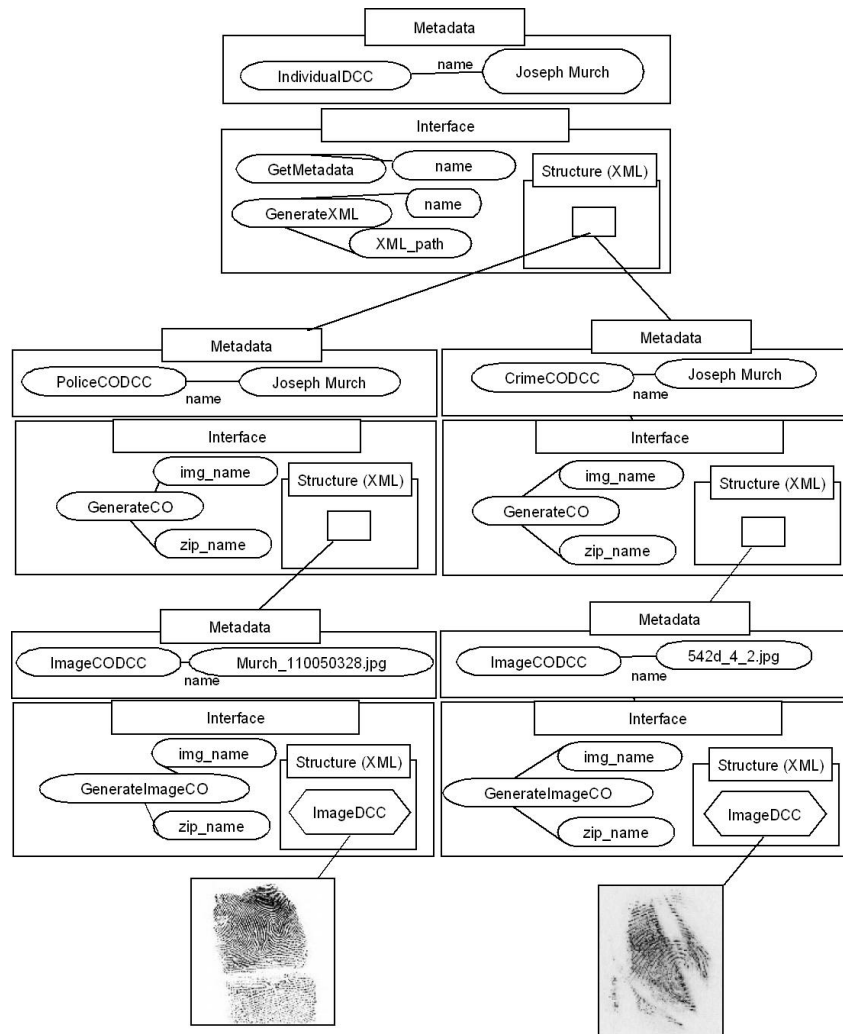


Figura 7.6: Estrutura do IndividualDCC.

integrada.

Neste segundo estudo de caso, para combinar a busca visual e textual em busca multimodal, adaptamos a infraestrutura apresentada no capítulo anterior, adicionando um novo componente: a recuperação multimodal, baseada em *re-rank* e *rank-aggregation*.

Basicamente, as abordagens de *rank aggregation* têm por objetivo combinar diferentes rankings, a fim de obter um mais preciso. Cada ranking a ser combinado pode ser calculado por um descritor diferente. Em suma, exploramos a relação entre imagens e informação contextual para combinar listas ordenadas de diferentes descritores.

A motivação principal do algoritmo de *re-rank* e *rank-aggregation* baseia-se na suposição de que *informações contextuais codificadas na semelhança entre ranked lists podem fornecer recursos para melhorar a eficácia de descritores*. Em geral, se duas imagens são semelhantes,

```

<?xml version="1.0" encoding="UTF8"?>
<individual>Joseph Murch
  <individual_name>Joseph Murch</individual_name>
  <individual_age>22</individual_age>
  <individual_sex> M</individual_sex>
  <image_DL_indiv>Joseph Murch
    <image_DL>Police Prints Digital Library</image_DL>
    <image> Murch_110050328.jpg
      <image_name>Murch_110050328.jpg</image_name>
      <image_feature_vector_name>/home/nadiapk/data/fv/Murch_110050328.jpg.txt</image_feature_vector_name>
      <image_descriptor>SASI
        <image_name> Murch_110050328.jpg <image_dist_value>0</image_dist_value></image_name>
        <image_name>Fox_110050395.jpg<image_dist_value>0.0033</image_dist_value></image_name>
        <image_name>Fox_110050350.jpg<image_dist_value>0.0037</image_dist_value></image_name>
        <image_name>Murch_110050327.jpg<image_dist_value>0.0038</image_dist_value></image_name>
        <image_name>Fox_110050392.jpg<image_dist_value>0.0039</image_dist_value></image_name>
      </image_descriptor></image></image_DL_indiv>
    <image_DL_indiv>Joseph Murch
      <image_DL>Crime Scene Digital Library</image_DL>
      <image> 524d_4_2.jpg
        <image_name>524d_4_2.jpg</image_name>
        <image_feature_vector_name>/home/nadiapk/data/fv/524d_4_2.txt</image_feature_vector_name>
        <image_descriptor>SASI
          <image_name> 524d_4_2.jpg <image_dist_value>0</image_dist_value></image_name>
          <image_name>524D_10_1.jpg<image_dist_value>0.0045</image_dist_value></image_name>
          <image_name>524D_22_1.jpg<image_dist_value>0.0059</image_dist_value></image_name>
          <image_name>524D_25_2.jpg<image_dist_value>0.0060</image_dist_value></image_name>
          <image_name>524D_4_1.jpg<image_dist_value>0.0082</image_dist_value></image_name>
        </image_descriptor>
      </image>
    </image_DL_indiv></individual>

```

Figura 7.7: XML para a agregação do conceito de indivíduo.

as suas *ranked lists* devem ser semelhantes também [96]. É de certa forma semelhante com a hipótese de cluster [125], que afirma que “*documentos estritamente associados tendem a ser relevantes para as mesmas requisições*”.

O objetivo é utilizar um algoritmo de *re-ranking* que utilize como entrada um conjunto de *ranked lists* e utilize uma função para computar uma matriz de distância resultante, e portanto, um conjunto de *ranked lists* mais efetivo. Mais detalhes sobre o algoritmo de re-ranking, escores de similaridade e o algoritmo de ranking podem ser obtidos em [95,97].

7.2.1 Utilizando DCCs para a Busca Multimodal

Considerando a infraestrutura apresentada no capítulo anterior, o objetivo era reutilizar os componentes, adicionando a busca multimodal. Esta abordagem favorece a integração de serviços aos desenvolvedores de aplicações (já que reutiliza os recursos já disponíveis) e favorece o uso de COs pela infraestrutura (já que permite o uso do serviço pelo componente individual, ou o uso integrado com outros serviços pelo CO). Dois novos DCCs foram criados (Figura 7.8):

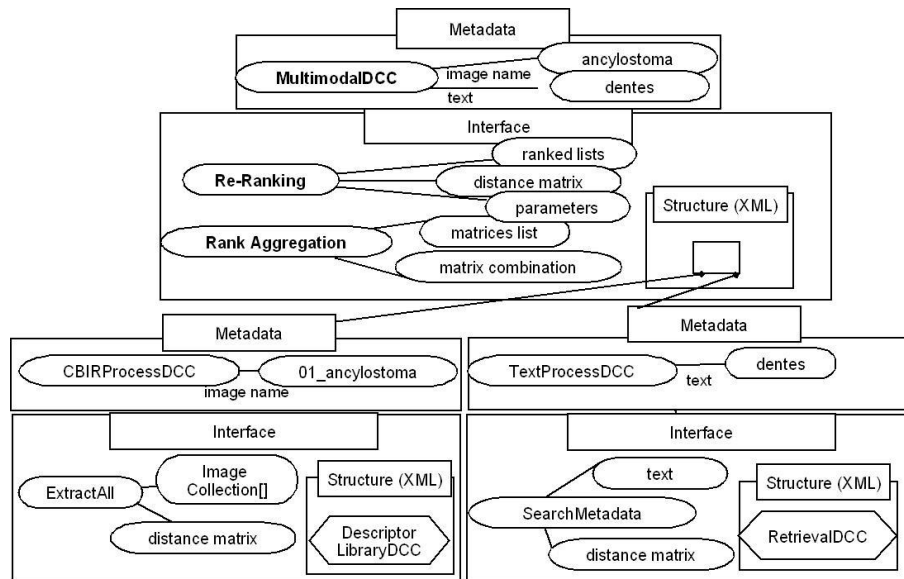


Figura 7.8: A estrutura do MultimodalDCC.

TextProcessDCC and MultimodalDCC.

O TextProcessDCC é um DCC de processo, encapsulando a biblioteca responsável pela busca textual. Basicamente, ele realiza a busca de texto (utilizando o RetrievalDCC), retornando uma *ranked list* com os metadados mais similares.

O MultimodalDCC é outro DCC de processo, encapsulando os algoritmos de *re-ranking* e *rank aggregation*, e utilizando como fonte de dados os *ranked lists* originários do TextProcessDCC e CBIRProcessDCC (apresentado no capítulo anterior). A Figura 7.8 ilustra a estrutura do MultimodalDCC, e como é reutilizada a estrutura CBIRProcessDCC. No exemplo, a busca está combinando os resultados da palavra “dentes” com o processo CBIR da imagem *01_ancylostoma.jpg*.

7.2.2 A Publicação dos COs

Na publicação do CO, utilizamos a mesma infraestrutura descrita na Seção 6.1.2.

7.2.3 Estudo de Caso

Considere o estudo de caso apresentado no Capítulo 3, sobre o projeto *Deployment and Assessment of an Image Annotation and Retrieval Tool, Including Biodiversity* [55, 56] e a Ferramenta SuperIDR. Nesta seção exploraremos a ferramenta no contexto da busca textual, CBIR e a combinação de ambas na busca multimodal, utilizando a infraestrutura apresentada no capítulo anterior.

A Implementação

A infraestrutura proposta permite a instalação de descritores distintos, mas para os testes apresentados, o descritor *Border/Interior pixel Classification* (BIC) [118] foi utilizado. A biblioteca foi implementada em linguagem C, e os DCCs em Java. As funções e parâmetros de cada DCC, além dos dados da aplicação foram armazenados em um banco de dados PostgreSQL.

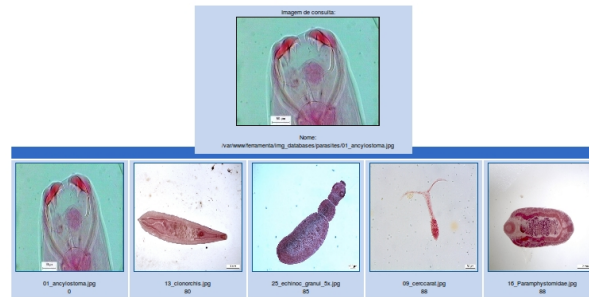


Figura 7.9: Resultado da consulta CBIR para a imagem *01_ancylostoma.jpg*.

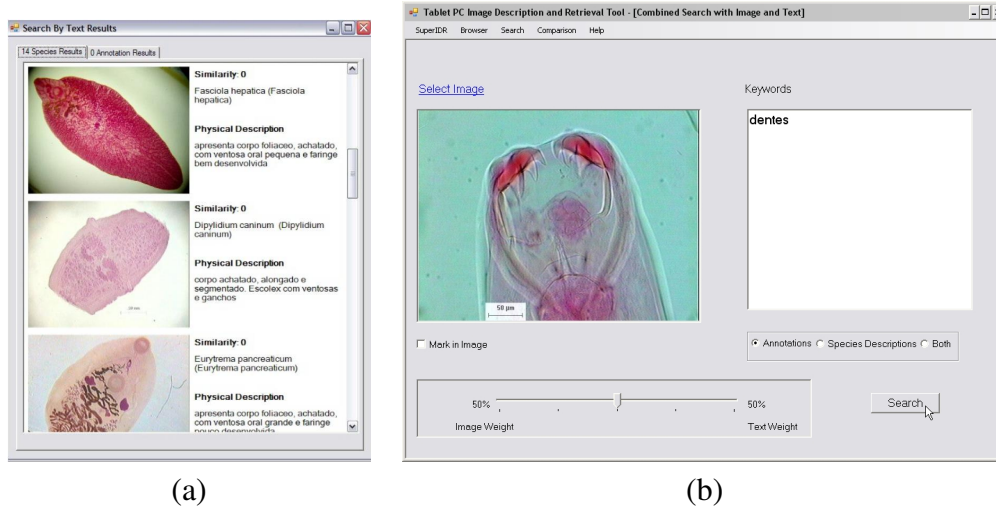


Figura 7.10: (a) Resultados da Busca Textual. (b) Interface para a Busca Multimodal.

O Protótipo

O protótipo teve por objetivo incorporar a busca textual e a busca multimodal (com os algoritmos de *re-rank* e *rank-aggregation*).

A busca textual (utilizando o TextProcessDCC) é ilustrada na Figura 7.10-a, com os resultados apresentados pela busca da palavra “corpo”. Note que a busca textual nesta aplicação pode ser delimitada às anotações, a descrição das espécies, ou ambos.

Para a busca multimodal, considere que o usuário inicialmente realiza a busca por conteúdo utilizando como imagem de consulta a imagem *01_ancylostoma.jpg*, como ilustra a Figura 7.9. Além disso, o usuário também pode combinar a busca por conteúdo com a busca textual utilizando a palavra “dentes” (realizada nos metadados disponíveis na base de dados). A busca textual retorna também uma *ranked list*, com as palavras mais semelhantes à palavra utilizada na busca. Note que na parte inferior da Figura 7.10-b (ilustrando a busca multimodal), o usuário também pode estabelecer o peso da relevância da busca visual (*Image Weight*) e da busca textual (*Text Weight*).

Posteriormente, os mesmos COs apresentados no capítulo anterior foram disponibilizados para a publicação. A busca multimodal não impactou nenhuma outra etapa da infraestrutura, já que ela depende especificadamente da palavra utilizada na busca.

Como pode ser observado, foi possível utilizar toda a infraestrutura CBIR para a busca multimodal, criando somente dois novos DCCs (TextProcessDCC e MultimodalDCC).

7.3 Busca em Estórias de Vídeo

Nesta seção será apresentado como reutilizamos a estrutura apresentada no capítulo anterior para a busca visual de estórias em vídeo. Os benefícios deste reuso serão demonstrados com um estudo de caso, com uma coleção de 50 vídeos.

7.3.1 A Geração da Sumarização de Estórias de Vídeos

Os seres humanos julgam mais rapidamente a relevância de vídeos quando existem segmentos interligados. No entanto, descobrir a relação ideal para esse tipo de interligação não é trivial. A abordagem mais simples consiste em agrupar quadros de vídeos [6] com conteúdo semelhante, de modo que um julgamento relevante para um frame de vídeo possa ser aplicado a todos os frames semelhantes.

Na perspectiva abordada, estórias são uma maneira gerenciável para pesquisar os segmentos de vídeo. Eles consistem de várias imagens e são representados por uma coleção de frames, como ilustrado na Figura 7.11.

7.3.2 Utilizando DCCs para a Busca de Estórias

Considerando a infraestrutura apresentada no capítulo anterior, o objetivo deste estudo de caso era reutilizar os componentes, adicionando as imagens e componentes para a sumarização de vídeo. Para isso, dois novos DCCs foram criados (Figura 7.12): StoryDCC e CBIRStoryDCC.

O StoryDCC é um DCC passivo, responsável pela agregação dos frames das várias estórias presentes em um vídeo. Basicamente, este DCC agrega os metadados e a coleção dos frames



Figura 7.11: Um exemplo de estórias produzidas para o vídeo *A New Horizon, segment 04*.

de vídeo resultantes da sumarização do vídeo. A Figura 7.12 apresenta o StoryDCC para as estórias apresentadas na Figura 7.11. Note que existem 15 ImageCODCCs, encapsulando as 15 imagens resultantes do método de sumarização.

O CBIRStoryDCC é um DCC de processo DCC, centralizando a busca CBIR para as estórias de vídeo, utilizando o StoryDCC e o CBIRProcessDCC. Note que o CBIRStoryDCC na Figura 7.11 reutiliza o CBIRProcessDCC.

7.3.3 Estudo de Caso

Para este estudo de caso, utilizamos o descritor *Border/Interior pixel Classification (BIC)* [118]. A biblioteca foi implementada em C, e os DCCs na linguagem Java. Os metadados, e as funções e parâmetros disponíveis para cada DCC foram armazenados em um banco de dados PostgreSQL.

A implementação deste estudo de caso teve as seguintes fases: (i) a “descoberta” e definição de cada parte do CO; (ii) a identificação de cada parte do CO; (iii) o processo de sumarização de vídeo; (iv) o processo CBIR; e (v) o encapsulamento da imagem e metadados relacionados.

Inicialmente, uma amostragem de 50 vídeos foram selecionadas randomicamente do *Open Video Project*², com seus respectivos metadados. Todos os vídeos estavam no formato MPEG-1 (com a resolução 352×240 e 29.97 frames por segundo), em cores e com som. Os vídeos selecionados estavam distribuídos entre vários gêneros (como por exemplo, documentário, educacional, histórico, etc.) e suas durações variavam entre hum e quatro minutos. Estes vídeos foram os mesmos que os utilizados em [6, 7, 30, 77].

Na fase dois, definimos que a agregação trataria o conceito das estórias para cada vídeo. Para a identificação das partes deste CO, utilizamos o banco de dados, responsável por relacionar as imagens aos seus respectivos vídeos, metadados e informações CBIR.

²<http://www.open-video.org/> (último acesso em 05/05/11).

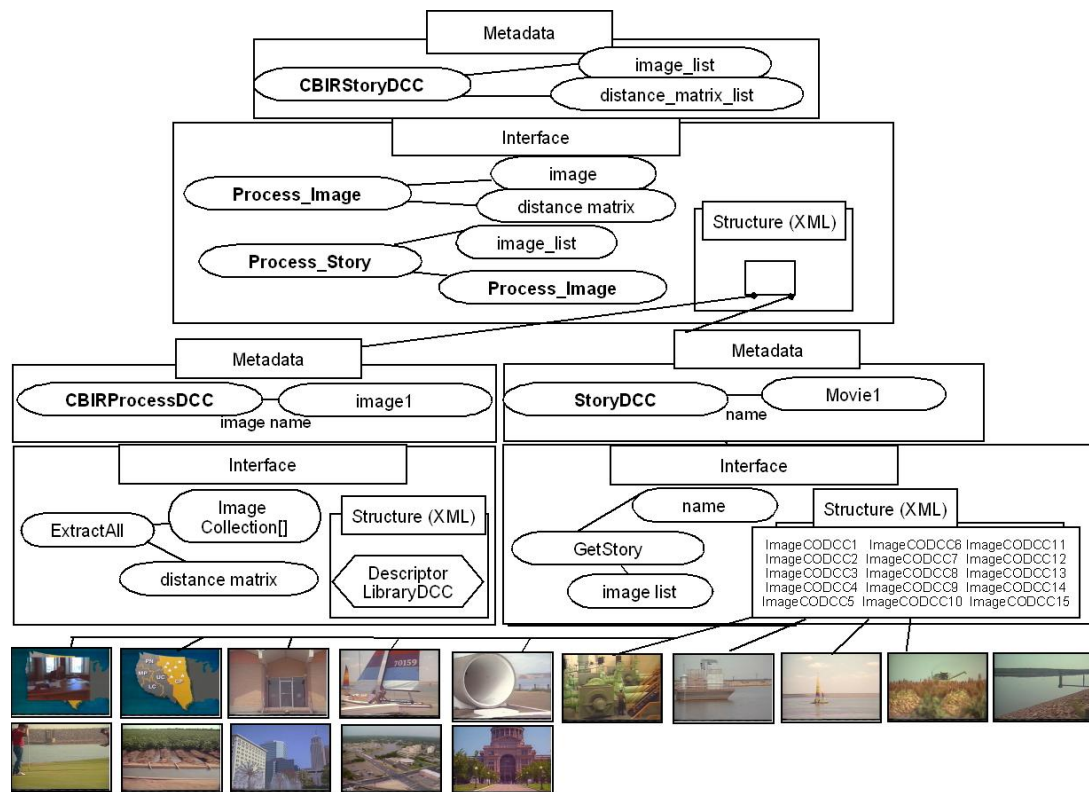


Figura 7.12: A estrutura do CBIRStoryDCC.

Na fase três, as coleções de histórias foram geradas para as 50 amostras de vídeos selecionadas, utilizando os algoritmos *Delaunay Triangulation* (DT) [77], *STill and MOving Video Storyboard* (STIMO) [30], *Video SUMMarization* (VSUMM) [7], e *Video Summarization for ONline applications* (VISON) [6]. As histórias produzidas pelos três primeiros algoritmos estão disponíveis online³. As histórias geradas pelo algoritmo VISON estão disponíveis em <http://www.liv.ic.unicamp.br/jurandy/summaries/> (último acesso em 05/05/11).

O processo CBIR compreende a execução de um algoritmo de extração em uma coleção de imagens, seguido pela comparação das medidas de distância. Nesta etapa outros componentes relacionados às imagens tornaram-se disponíveis, como os vetores de características e as medidas de distância.

Considere a quinta história de vídeo apresentada na Figura 7.11. O processo CBIR para esta imagem gera um vetor de características, e as medidas de distância relativas as outras imagens da coleção. A Figura 7.13 ilustra as imagens mais semelhantes a história mencionada, considerando o descritor de cor.

Se analisarmos as cinco imagens mais semelhantes, na região inferior da Figura 7.11, temos

³<http://sites.google.com/site/vsummsite/> (último acesso em 05/05/11).



Figura 7.13: Resultado da consulta CBIR para uma estória do vídeo *A New Horizon*, segment 04.

que: (i) as três primeiras imagens mais semelhantes são relativas à mesma estória, do mesmo vídeo; (ii) a quarta imagem é relativa a uma estória no vídeo *Drift Ice as a Geologic Agent*, segment 03; e (iii) a quinta imagem é relativa a uma estória no vídeo *Exotic Terrane*, segment 02.

Na última fase, os objetos imagem, ICO e descritores foram publicados através do protocolo OAI-PMH.

Em resumo este caso de estudo apresentou uma nova abordagem para reutilizar uma infraestrutura CBIR baseada em componentes para a busca de histórias de vídeo. Somente dois componentes foram criados (StoryDCC e CBIRStoryDCC), agregando os objetos de uma estória e combinando o processo CBIR para cada etapa, reutilizando para isso vários componentes apresentados no capítulo anterior.

7.4 Conclusões

Neste capítulo foram apresentadas três aplicações para exemplificar o reuso da infraestrutura baseada em DCC-OAI para objetos complexos (a integração de bibliotecas digitais de imagens digitais, a busca multimodal e a busca de histórias de vídeo).

A contribuição deste capítulo reside em exemplificar os diferentes domínios que carecem de um suporte específico e que podem reaproveitar componentes para o tratamento de objetos complexos de imagem. Este capítulo também verificou outro desafio relacionado a objetos complexos: a busca individual dos componentes e integrada no CO (através da busca multimodal).

Existem vários tópicos que podem ser explorados no reuso de nossa infraestrutura, como o

serviço de anotação, outros domínios (biodiversidade e GIS), outros tipos de busca multimodais (como anotações e áudio), a análise do impacto em coleções com grande número de imagens e possíveis integrações com softwares de bibliotecas digitais.

Capítulo 8

Conclusões

Esta tese desenvolveu pesquisa combinando aspectos teóricos e práticos de objetos complexos em bibliotecas digitais. Ao combinar conceitos de objetos complexos e bibliotecas digitais, apresentamos uma solução alternativa no suporte para criação, manutenção e integração de sistemas heterogêneos.

A partir dos conceitos básicos (como componentes, identidade, estrutura, granularidade, DLs, Arcabouço 5S e CBIR, entre outros) esta tese analisou objetos complexos perante três perspectivas: aplicações, formalização, e tecnologias.

Esta tese baseou-se em três aplicações (Capítulo 3). O sistema Nou-Rau exemplifica a integração de vários documentos relacionados à teses e dissertações. Seus principais serviços incluem a busca e a navegação. A aplicação SuperIDR trata de imagens, metadados, marcas e anotações, relacionados ao domínio de parasitologia. Seus principais serviços incluem a navegação, a busca textual, a busca visual e a busca multimodal. A integração de bibliotecas digitais de Impressões Digitais utiliza imagens e metadados relacionados a quatro bibliotecas digitais distintas.

Da perspectiva da formalização (Capítulo 4), o Arcabouço 5S foi estendido para o tratamento de busca de imagens por conteúdo e objetos complexos. Este formalismo pode ser utilizado para descrever agregações e vários conceitos disponíveis no Arcabouço 5S. O conjunto de definições também pode impactar aplicações futuras (como as apresentadas em [33, 76, 112, 115, 135]), já que servem de suporte para aplicações baseadas em objetos complexos e conteúdo de imagem. Os conceitos propostos foram ilustrados a partir de um estudo de caso com a ferramenta SuperIDR. Há várias possibilidades de extensões para trabalhos futuros nesta área, como a formalização de aspectos relacionados a qualidade de dados, granularidade, persistência, versões. Esta formalização poderia facilitar o uso da abordagem em outros domínios, como o uso de diferentes versões de dados e softwares.

Da perspectiva de tecnologias (Capítulo 5), analisamos como objetos complexos podem ser combinados com o Arcabouço 5S e as tecnologias de CO. Uma vez que existem várias di-

mensões de informações relacionadas, nós nos concentramos em quatro aspectos das COs: identidade, estrutura, componentes, e delimitação do objeto. Nossa contribuição baseia-se na análise das tecnologias DCC, Buckets e OAI-ORE, investigando suas semelhanças, além de explorá-los sob a perspectiva do Arcabouço 5S. Finalmente, apresentamos um estudo de caso utilizando uma biblioteca digital de impressões digitais para ilustrar os benefícios da nossa abordagem de modelagem. Existem vários tópicos de pesquisa que podem ser exploradas para ampliar ainda mais nosso trabalho atual. Eles incluem o estudo do impacto das COs em outros aspectos do Arcabouço 5S, e o uso de COs em outros domínios (como os sistemas de informação de biodiversidade) e os serviços específicos (como a recuperação de imagens baseado em conteúdo e anotação).

O Capítulo 6 apresenta uma infraestrutura baseada no conceito de CO, gerenciando o processo de busca de imagens baseado em conteúdo, além do encapsulamento, e publicação de objetos complexos de imagem. A principal novidade reside na utilização do modelo de componentes *Digital Content Component* (DCC) para encapsular o processo de busca de imagens por conteúdo e o objeto complexo de imagem. Posteriormente, os dados são publicados como itens de um repositório de dados, usando o protocolo *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH). Trabalhos futuros nesta área incluem a formalização de serviços que utilizam estas estruturas. Exemplos incluem serviços de busca multimodais, serviços de recomendação para objetos complexos, e o serviço de navegação de imagens baseado em conteúdo. Além disso podemos citar o mapeamento dos diferentes componentes de um CO através do OAI-ORE, e a utilização/composição de outros descritores (já que para facilitar a implementação, somente um algoritmo de cor, outro de textura e outro de forma foram utilizados).

O Capítulo 7 apresenta exemplos de diferentes domínios que carecem de um suporte específico e que podem reaproveitar componentes para o tratamento de objetos complexos de imagem. As aplicações propostas exploram tanto os componentes individuais quanto as agregações baseadas no CO. Este capítulo também verificou outro desafio relacionado a objetos complexos: a busca individual dos componentes e integrada no CO (através da busca multimodal). Existem vários tópicos que podem ser explorados no reuso de nossa infraestrutura, como o serviço de anotação, outros domínios (biodiversidade e GIS), outros tipos de busca multimodais (como anotações e áudio), e possíveis integrações com softwares de bibliotecas digitais.

Em resumo, as principais contribuições deste trabalho são:

1. formalização de conceitos relacionados a objetos complexos e busca de imagens por conteúdo, ilustrando o objeto complexo baseado em imagens da aplicação SuperIDR;
2. comparação e análise de tecnologias de objetos complexos (DCC, Buckets e OAI-ORE), da perspectiva do framework 5S;
3. proposta e desenvolvimento de um protótipo que provê:

- o acesso e gestão homogênea ao processo de busca de imagens por similaridade, e fontes de dados (imagens, metadados e descritores);
 - a fácil integração do processo da busca de imagens por similaridade através da componentização e encapsulamento;
 - a distribuição e coleta de metadados de objetos complexos de imagens, juntamente com as medidas de similaridade e vetores de características;
 - o acesso aos componentes individuais e ao CO como entidade única;
 - a centralização do processamento, encapsulamento, publicação e coleta do objeto complexo de imagem.
4. o uso da infraestrutura desenvolvida para buscas multimodais;
 5. a integração desta infraestrutura para a busca em sumarização de vídeos;
 6. exemplos de cenários que utilizam integração e agregação de informações, ilustrando serviços e conceitos relacionados a objetos complexos;
 7. exemplo de integração de metadados, imagem, vetores de características e processamento na aplicação SuperIDR, utilizando imagens de parasitas;
 8. exemplo de integração de quatro bibliotecas digitais de impressões digitais, utilizando a perspectiva de objeto complexo.

Dentre as publicações no doutorado, as seguintes fazem parte do corpo da tese:

Autoria:

1. *Complex Objects in Digital Libraries* [51]: artigo apresentado no *Doctoral Consortium - ACM-IEEE Joint Conference on Digital Libraries* (2009). Este artigo descreve a motivação, trabalhos relacionados e cenários referentes ao uso de objetos complexos em bibliotecas digitais;
2. *SuperIDR: A tablet PC tool for Image Description and Retrieval* [83], artigo apresentado no *Workshop on the Impact of Pen-Based Technology on Education* (Wipte 2010). Este artigo apresenta a ferramenta SuperIDR e os experimentos realizados no domínio da Parasitologia e Ictiologia;
3. *Rethinking Fingerprint evidence through integration of very large digital libraries* [59]: artigo apresentado no *Very Large Digital Libraries Workshop at 14th European Conference on Research and Advanced Technology for Digital Libraries* (ECDL2010). Este artigo descreve a integração de bibliotecas digitais de impressões digitais;
4. *Reusing a Compound-Based Infrastructure for Searching Video Stories* [52], artigo apresentado no *12th IEEE International Conference on Information Reuse and Integration* (IRI 2011). Este artigo apresenta o reuso da infraestrutura proposta nesta tese para a busca de histórias sumarizadas de vídeo;

5. *Describing OAI-ORE from the 5S Framework Perspective* [54]: pôster apresentado no *ACM-IEEE Joint Conference on Digital Libraries* (2010). Este artigo escreve o Open Archives Initiative Object Reuse and Exchange (OAI-ORE) utilizando a perspectiva 5S;
6. *A Teaching Tool for Parasitology: Enhancing Learning with Annotation and Image Retrieval* [55], pôster apresentado no *European Conference on Research and Advanced Technology for Digital Libraries* (ECDL 2010). Este artigo analisa os experimentos da ferramenta SuperIDR no domínio da Parasitologia;
7. *A Database for Fingerprint Experiments* [58], pôster apresentado no CESCA (*Center for Embedded Systems for Critical Applications*) Day (2010). Este artigo descreve o banco de dados utilizado em experimentos de imagens de impressões digitais;
8. *A .Net Implementation of a Content-Based Image Search Component* [60]: pôster apresentado na Sessão de Demos, XXIII Simpósio Brasileiro de Banco de Dados (2008). Ele descreve uma implementação baseada na linguagem .Net de um componente CBIR, utilizado para a busca e anotações de imagens;
9. *A Geographic Annotation Service in SuperIDR* [57]: pôster apresentado no VT-GIS 2010 symposium. Este artigo descreve a integração de anotações geográficas na ferramenta SuperIDR;
10. *Evaluation of a Tablet PC image annotation and retrieval tool in the parasitology domain*(2009) [56]. Este relatório técnico apresenta as atividades relacionadas ao uso da aplicação tablet PC no domínio da parasitologia na Unicamp;
11. *Extending the 5S Framework of Digital Libraries to support Complex Objects* (2010) [84]. Este relatório técnico apresenta a definição formal de objetos complexos utilizando conceitos do Arcabouço 5S ;
12. *Exploring CBIR concepts in the CTRnet Project*(2010) [53]. Este relatório técnico apresenta as atividades durante o primeiro semestre na Virginia Tech - USA, voltadas ao uso de técnicas de recuperação de imagens baseadas em conteúdo relacionadas ao Terremoto no Haiti, no projeto CTRnet;
13. *Analyzing Compound Object Technologies from the 5S Perspective*(2011) [61]. Este relatório técnico apresenta a análise de tecnologias de objetos complexos da perspectiva do Arcabouço 5S.
14. *A Complex-Based Infrastructure for Searching and Harvesting Complex Image Objects* (2011). Este artigo de periódico aceito no *Information-Interaction-Intelligence (I3) Journal* apresenta uma infraestrutura baseada em objetos complexos para a busca e publicação de objetos complexos de imagem.

Co-autoria:

1. *A Geographic Annotation Service for Biodiversity Systems* [31], artigo apresentado no XI Simpósio Brasileiro de Informática (Geoinfo 2010). Este artigo apresenta um serviço para anotações geográficas em Sistemas de Biodiversidade;
2. *Model-based fingerprint image quality Analysis* [92], pôster apresentado no CESSA (*Center for Embedded Systems for Critical Applications*) Day (2010). Este artigo analisa modelos de distorções de imagens de impressões digitais;
3. *Fingerprint Sufficiency* [74], pôster apresentado no CESSA (*Center for Embedded Systems for Critical Applications*) Day (2010). Este artigo descreve a detecção e remoção de falsas minúcias em imagens de impressões digitais;
4. *Latent Fingerprint Matching* [116], pôster apresentado no CESSA (*Center for Embedded Systems for Critical Applications*) Day (2010). Este artigo descreve a identificação de imagens de impressões digitais com níveis diferenciados de atrito;
5. *CTRnet DL for Disaster Information Services* [134], pôster apresentado no *ACM-IEEE Joint Conference on Digital Libraries* (2011). Este artigo descreve os serviços disponibilizados na ferramenta CTRnet;
6. *Toward a quantitative basis for sufficiency of friction ridge pattern detail* [44], artigo apresentado no *Impression Pattern Evidence Symposium* (2010). Este artigo descreve experimentos relacionados a análise de detalhes em impressões digitais.

Referências Bibliográficas

- [1] P. Achananuparp, K. W. McCain, and R. B. Allen. Supporting student collaboration for image indexing. In *Proceedings of the 10th international conference on Asian digital libraries*, ICADL'07, pages 24–34, Berlin, Heidelberg, 2007. Springer-Verlag.
- [2] Advanced Distributed Learning (ADS). Sharable Content Object Reference Model (SCORM), disponível em <http://www.adlnet.org/index.cfm?fuseaction=SCORDown> - último acesso em 05/05/11, 2004.
- [3] A. Gerber and J. Hunter. Authoring, editing and visualizing compound objects for literary scholarship. *Journal of Digital Information (JoDI)*, 11(1), 2010.
- [4] M. Agosti and N. Ferro. A formal model of annotations of digital content. *Transactions on Information Systems (TOIS)*, 26(1):1–55, 2008.
- [5] J. Almeida, A. Rocha, R. S. Torres, and S. Goldenstein. Making colors worth more than a thousand words. In *Proceedings of the 2008 ACM symposium on Applied computing*, SAC '08, pages 1180–1186, New York, NY, USA, 2008. ACM.
- [6] J. Almeida, R. S. Torres, and N. J. Leite. Rapid video summarization on compressed video. In *Proceedings of the IEEE International Symposium on Multimedia*, pages 113–120, Taichung, Taiwan, Dec. 13–15 2010. IEEE Computer Society.
- [7] S. E. F. Avila, A. P. B. Lopes, A. Luz Jr., and A. A. Araújo. VSUMM: A mechanism designed to produce static video summaries and a novel evaluation method. *Pattern Recognition Letters*, 32(1):56–68, Jan. 2011.
- [8] C. Awre. Legal issues of compound ETDs. In *Knowledge Exchange Group - Research paper - Enhanced E-theses Project Deliverable 9.1*, disponível em <http://igitur-archive.library.uu.nl/DARLIN/2010-0526-200238/UUindex.html> - último acesso em 05/05/11, 2009.
- [9] C. Awre. Managing compound objects within Fedora, Knowledge Exchange Group - Enhanced E-theses Project Deliverable 9, disponível em <http://igitur-archive.library.uu.nl/DARLIN/2010-0526-200241/UUindex.html> - último acesso em 09/20/10. 2009.

- [10] D. Bainbridge and I. H. Witten. A fedora librarian interface. In *JCDL '08: Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries*, pages 407–416, New York, NY, USA, 2008. ACM.
- [11] I. S. Burnett, F. Pereira, R. V. de Walle, and R. Koenen. *The MPEG-21 Book*. John Wiley & Sons, 2006.
- [12] A. J. C. Smythe. IMS Content Packaging Information Model, specification, IMS Global Learning Consortium, Inc., Oct, 2009.
- [13] R. T. Calumby. Multimodal image retrieval with relevance feedback based on genetic programming. *Master's Thesis, Universidade Estadual de Campinas*, 2010.
- [14] L. Candela, D. Castelli, N. Ferro, Y. Ioannidis, G. Koutrika, C. Meghini, P. Pagano, S. Ross, D. Soergel, M. Agosti, M. and Dobрева, V. Katifori, and H. Schuldt. The DELOS Digital Library Reference Model - Foundations for Digital Libraries. Version 0.98, Delos Digital Library, disponível em http://www.delos.info/files/pdf/ReferenceModel/DELOS_DLReferenceModel_0.98.pdf - último acesso em 04/24/11 , 2008.
- [15] L. Candela, D. Castelli, Y. Ioannidis, G. Koutrika, P. Pagano, S. Ross, H. J. Schek, and H. Schuldt. Deliverable d1.4.2: A reference model for digital library management systems interim report, delos digital library, disponível em http://www.delos.info/index.php?option=com_contenttask=view&id=345 - último acesso em 04/24/11, 2006.
- [16] A. Carkacioglu and F. Yarman-vural. Sasi: A new texture descriptor for content based image retrieval. *IEEE International Conference on Image Processing*, 2:137–140, 2001.
- [17] K. Cheung, A. Lashtabeg, and D. John. Scope: A scientific compound object publishing and editing system. *The International Journal of Digital Curation*, 2(2):4–18, 2008.
- [18] M. V. Cundiff. An introduction to the Metadata Encoding and Transmission Standard (METS). *Library Hi Tech*, 22(1):52–64, 2004.
- [19] R. da S. Torres and A. X. Falcão. Content-based image retrieval: Theory and applications. *Revista de Informática Teórica e Aplicada*, 13(2):161–185, 2006.
- [20] R. da S. Torres, C. Medeiros, M. Gonçalves, and E. A. Fox. A digital library framework for biodiversity information systems. *International Journal on Digital Libraries*, 6(1):3–17, 2006.
- [21] R. Datta, D. Joshi, J. Li, James, and Z. Wang. Image retrieval: Ideas, influences, and trends of the new age. *ACM Transactions on Computing Surveys*, 40(2), April 2008.
- [22] D. Davis. Fedora repository 3 documentation, available at <http://www.fedora-commons.org/confluence/display/fcr30> - último acesso em 03/03/11, 2009.

- [23] Digital Library Federation. Metadata Encoding and Transmission Standard (METS), METS Schema 1.9 and Documentation disponível em <http://www.loc.gov/standards/mets/> - último acesso em 04/04/11, 2009.
- [24] R. Dodd. 20 years on: the Dexter Model of Hypertext and its impact on web accessibility. *ACM SIGACCESS Accessibility and Computing*, (90):3–15, 2008.
- [25] P. Dourado, P. Ferreira, and A. Santanchè. Representação Unificada de Objetos Digitais Complexos: Confrontando o RAS com o IMS CP. In *III Workshop de Bibliotecas Digitais, Webmedia 2007*, 2006.
- [26] L. K. El Demerdash and S. Bergler. Semantic inter-media image retrieval in photographic collections. In *Proceeding of IEEE Multimedia Signal Processing (MMSP) 2009*, MMSP' 2009, 2009.
- [27] P. Eyler. Koha: a gift to libraries from new zealand. *Linux Journal*, 2003(106):1, 2003.
- [28] F. Fonseca, M. Egenhofer, C. Davis, and G. Câmara. Semantic granularity in ontology-driven geographic information systems. *Annals of Mathematics and Artificial Intelligence*, 36(1-2):121–151, 2002.
- [29] N. Fuhr, P. Hansen, M. Mabe, A. Micsik, and I. Solvberg. Digital libraries: A generic classification and evaluation scheme. In *Proceedings European Conference on Digital Libraries*, pages 187–199, Heidelberg et al., 2001. Springer.
- [30] M. Furini, F. Geraci, M. Montangelo, and M. Pellegrini. STIMO: STill and MOving video storyboard for the web scenario. *Multimedia Tools and Applications*, 46(1):47–69, Jan. 2010.
- [31] F. B. Gil, N. P. Kozievitch, and R. da S. Torres. A geographic annotation service for biodiversity systems. In *Proceedings of XV GeoInfo*, pages 33–44, 2010.
- [32] M. A. Gonçalves. *Streams, Structures, Spaces, Scenarios, and Societies (5S): A Formal Digital Library Framework and Its Applications*. PhD thesis, Virginia Polytechnic Institute & State University, Blacksburg, VA, 2004.
- [33] M. A. Gonçalves and E. A. Fox. 5SL: a language for declarative specification and generation of digital libraries. In *JCDL'02: Proceedings of the 2002 Conference on Digital libraries*, pages 263–272, Portland, Oregon, USA, 2002.
- [34] M. A. Gonçalves, E. A. Fox, L. T. Watson, and N. A. Kipp. Streams, Structures, Spaces, Scenarios, Societies (5S): A Formal Model for Digital Libraries. *ACM Transactions on Information Systems (TOIS)*, 22(2):1–43, 2004.
- [35] M. A. Gonçalves, B. L. Moreira, E. A. Fox, and L. T. Watson. “What is a good digital library?” - A quality model for digital libraries. *Information Processing and Management: an International Journal*, 43(5):1416–1437, 2007.

- [36] D. Gorton. Practical Digital Library Generation into DSpace with the 5S Framework. *Master's Thesis, Virginia Tech*, 2010.
- [37] D. Grigori, J. C. Corrales, M. Bouzeghoub, and A. Gater. Ranking BPEL Processes for Service Discovery. *IEEE Transactions on Services Computing*, 3:178–192, July 2010.
- [38] K. Grønby. Composites in a Dexter-based hypermedia framework. In *ECHT '94: Proceedings of the 1994 ACM European conference on Hypermedia technology*, pages 59–69, New York, NY, USA, 1994. ACM.
- [39] R. Grossman, X. Qin, W. Xu, H. Hulen, and T. Tyler. An architecture for a scalable, high-performance digital library. In *MSS '95: Proceedings of the 14th IEEE Symposium on Mass Storage Systems*, page 89, Washington, DC, USA, 1995. IEEE Computer Society.
- [40] V. A. Gruzman and V. I. Senichkin. Hypermedia models. *Automation and Remote Control*, 62(5):677–694, 2001.
- [41] D. R. Hardy, M. F. Schwartz, and D. Wessels. Harvest 1.4 user's manual. In *Technical Report CU-CS-743-94, University of Colorado. Revised by K.-J. Lee in 2002 for Harvest version 1.8.*, 1996.
- [42] D. R. Hardy, M. F. Schwartz, and D. Wessels. Integrating, navigating, and analysing open Eprint archives through open citation linking (the OpCit project). *Current Science*, 79(5):629–638, 2000.
- [43] S. Hitchcock, L. Carr, Z. Jiao, D. Bergmark, W. Hall, C. Lagoze, and S. Harnad. Developing services for open eprint archives: globalisation, integration and the impact of links. In *DL '00: Proceedings of the fifth ACM conference on Digital libraries*, pages 143–151, New York, NY, USA, 2000. ACM.
- [44] M. Hsiao, N. P. Kozievitch, S. H. Park, E. A. Fox, N. Short, A. L. Abbott, S. Misra, R. Murch, and B. Budowle. Toward a quantitative basis for sufficiency of friction ridge pattern detail. *Impression Pattern Evidence Symposium, USA*, 8 2010.
- [45] P. J. J. Allinson and A. Powell. The Eprints Application Profile: a FRBR approach to modelling repository metadata. *OR '07: Second International Conference on Open Repositories*, january 2007.
- [46] P. H. J. Bekaert and H. V. de Sompel. Using MPEG-21 DIDL to Represent Complex Digital Objects in the Los Alamos National Laboratory Digital Library, disponível em <http://www.dlib.org/dlib/november03/bekaert/11bekaert.html>. *D-Lib Magazine*, 9(11), November 2003.
- [47] A. K. Jain and D. Maltoni. *Handbook of Fingerprint Recognition*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2003.

- [48] F. Jayakanth, K. Maly, M. Zubair, and L. Aswath. Approaches to make CDS/ISIS databases interoperable with OAI-compliant digital libraries. *Program: electronic library & information systems*, 39(3):269–278, Mar. 2005.
- [49] G. Z. P. Junior. *Managing the lifecycle of sensor data: from production to consumption*. PhD thesis, Instituto de Computação - Universidade Estadual de Campinas, December 2008.
- [50] J. F. Karpovich, A. S. Grimshaw, and J. C. French. Extensible file system (ELFS): an object-oriented approach to high performance file I/O. *ACM SIGPLAN Notices*, 29(10):191–204, 1994.
- [51] N. P. Kozievitch. Complex objects in digital libraries. In *JCDL '09: Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries, Doctoral Consortium*, disponível em <http://www.ieee-tcdl.org/Bulletin/v5n3/Kozievitch/kozievitch.html> - último acesso em 05/05/11, 2009.
- [52] N. P. Kozievitch, J. Almeida, R. da S. Torres, A. Santanchè, and N. Leite. Reusing a Compound-Based Infrastructure for Searching Video Stories. In *12th IEEE International Conference on Information Reuse and Integration (IRI)*. aug 2011.
- [53] N. P. Kozievitch, S. Codio, J. A. Francois, E. Fox, and R. da S. Torres. Exploring CBIR concepts in the CTRnet Project. Technical Report IC-10-32, Institute of Computing, University of Campinas, November 2010. In English, 20 pages.
- [54] N. P. Kozievitch and R. da S. Torres. Describing oai-ore from the 5s framework perspective. In *Proceedings of the role of digital libraries in a time of global change, and 12th international conference on Asia-Pacific digital libraries*, ICADL'10, pages 260–261, Berlin, Heidelberg, 2010. Springer-Verlag.
- [55] N. P. Kozievitch, R. da S. Torres, F. Andrade, U. Murthy, E. Fox, and E. Hallerman. A teaching tool for parasitology: enhancing learning with annotation and image retrieval. In *Proceedings of the 14th European conference on Research and advanced technology for digital libraries*, ECDL'10, pages 466–469, Berlin, Heidelberg, 2010. Springer-Verlag.
- [56] N. P. Kozievitch, R. da S. Torres, T. Falcão, E. Ramos, F. Andrade, S. M. Allegretti, M. T. Ueta, R. R. Madi, U. Murthy, E. A. Fox, Y. Chen, and E. Hallerman. Evaluation of a Tablet PC image annotation and retrieval tool in the parasitology domain. Technical Report IC-09-23, Institute of Computing, University of Campinas, July 2009. In English, 16 pages.
- [57] N. P. Kozievitch, R. da S. Torres, T. Falcão, E. Ramos, F. Andrade, S. M. Allegretti, M. T. Ueta, R. R. Madi, U. Murthy, E. A. Fox, and E. Hallerman. A Geographic Annotation Service in SuperIDR. *VT-GIS, Virginia Tech, Blacksburg, VA, USA*, 2010.

- [58] N. P. Kozievitch, R. da S. Torres, S. H. Park, E. A. Fox, N. Short, A. L. Abbott, S. Misra, and M. Hsiao. Database for fingerprint experiments. *Poster for CESSA (Center for Embedded Systems for Critical Applications) Day, Virginia Tech, Blacksburg, VA, USA*, 5 2010.
- [59] N. P. Kozievitch, R. da S. Torres, S. H. Park, E. A. Fox, N. Short, A. L. Abbott, S. Misra, and M. Hsiao. Rethinking fingerprint evidence through integration of very large digital libraries. *VLDL Workshop at 14th European Conference on Research and Advanced Technology for Digital Libraries (ECDL2010), Glasgow, Sept. 6-10, 8 pages*, 2010.
- [60] N. P. Kozievitch, T. R. C. Falcão, and R. da S. Torres. A .Net Implementation of a Content-Based Image Search Component. In *Sessão de Demos, XXIII Simpósio Brasileiro de Banco de Dados*, Campinas, Brasil, 2008.
- [61] N. P. Kozievitch, E. Fox, and R. da S. Torres. Analyzing Compound Object Technologies from the 5S Perspective. Technical Report IC-11-01, Institute of Computing, University of Campinas, January 2011.
- [62] D. B. Krafft, A. Birkland, and E. J. Cramer. Ncore: architecture and implementation of a flexible, collaborative digital library. In *JCDL '08: Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries*, pages 313–322, New York, NY, USA, 2008. ACM.
- [63] C. Lagoze and H. V. de Sompel. The open archives initiative: Building a low-barrier interoperability framework. In *JCDL '01: In Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, pages 54–62. ACM Press, 2001.
- [64] C. Lagoze, S. Payette, E. Shin, and C. Wilper. Fedora: an architecture for complex objects and their relationships. *International Journal of Digital Libraries*, 6(2):124–138, 2006.
- [65] C. Lagoze and H. V. Sompel. Compound Information Objects: the OAI-ORE Perspective. *Open Archives Initiative Object Reuse and Exchange, White Paper*, disponível em <http://www.openarchives.org/ore/documents> - último acesso em 05/02/11, 2007.
- [66] X. Li. Watermarking in secure image retrieval. *Pattern Recognition Letters*, 24(14):2431–2434, October 2003.
- [67] C. Lynch, S. Parastatidis, N. Jacobs, H. Van de Sompel, and C. Lagoze. The OAI-ORE effort: progress, challenges, synergies. In *JCDL '07: Proceedings of the 2007 conference on Digital libraries*, page 80. ACM, 2007.
- [68] Y. Ma, E. A. Fox, and M. A. Gonçalves. Pim through a 5s perspective. In *JCDL '07: Proceedings of the 7th ACM/IEEE-CS joint conference on Digital libraries*, pages 491–491, New York, NY, USA, 2007. ACM.

- [69] K. J. Maly, S. N. Shen, M. Zubair, and M. L. Nelson. Buckets: Aggregative, intelligent agents for publishing in digital libraries. *WebNet Journal: Internet Technologies, Applications Issues*, 1(1):58–66, 1999.
- [70] P. Manghi, M. Mikulicic, L. Candela, D. Castelli, and P. Pagano. Realizing and Maintaining Aggregative Digital Library Systems: D-NET Software Toolkit and OAster System. *D-Lib Magazine*, 16(3/4), 2010.
- [71] A. Maslov, A. Mikeal, S. Phillips, J. Leggett, and M. McFarland. Adding OAI-ORE Support to Repository Platforms. *Journal of Digital Information (JoDI)*, 11(1), 2010.
- [72] J. Melton and A. Eisenberg. SQL multimedia and application packages (SQL/MM). *ACM SIGMOD Record*, 30(4):97–102, 2001.
- [73] M. Minsky. A framework for representing knowledge. Technical report, Cambridge, MA, USA, 1974.
- [74] S. Misra, N. Short, A. L. Abbott, M. Hsiao, N. P. Kozievitch, and E. A. Fox. Fingerprint sufficiency. *Poster for CESCA (Center for Embedded Systems for Critical Applications) Day, Virginia Tech, Blacksburg, VA, USA*, 5 2010.
- [75] L. Moreau, B. Clifford, J. Freire, J. Futrelle, Y. Gil, P. Groth, N. Kwasnikowska, S. Miles, P. Missier, J. Myers, B. Plale, Y. Simmhan, E. Stephan, and J. V. den Bussche. The open provenance model core specification (v1.1). *Future Generation Computer Systems*, 27:743–756, June 2011.
- [76] B. L. Moreira, M. A. Gonçalves, A. H. F. Laender, and E. A. Fox. 5SQual: a quality assessment tool for digital libraries. In *JCDL '07: Proceedings of the 7th ACM/IEEE-CS joint conference on Digital libraries*, pages 513–513, New York, NY, USA, 2007. ACM.
- [77] P. Mundur, Y. Rao, and Y. Yesha. Keyframe-based video summarization using Delaunay clustering. *International Journal on Digital Libraries*, 6(2):219–232, Apr. 2006.
- [78] U. Murthy. *Digital Libraries with Superimposed Information: Supporting Scholarly Tasks that Involve Fine Grain Information*. PhD thesis, Virginia Polytechnic Institute & State University, Blacksburg, VA, 2011.
- [79] U. Murthy, R. da S. Torres, and E. A. Fox. Sierra – A Superimposed Application for Enhanced Image Description and Retrieval. In *European Conference on Digital Libraries*, volume LNCS 4172, pages 540–543, Alicante, Espanha, 2006.
- [80] U. Murthy, E. A. Fox, Y. Chen, E. Hallerman, R. Da Silva Torres, E. J. Ramos, and T. R. C. Falcão. Superimposed image description and retrieval for fish species identification. In *Proceedings of the 13th European conference on Research and advanced technology for digital libraries*, ECDL'09, pages 285–296, Berlin, Heidelberg, 2009. Springer-Verlag.

- [81] U. Murthy, E. A. Fox, Y. Chen, E. Hallerman, R. Torres, E. J. Ramos, and T. R. Falcão. Species identification: fish images with cbir and annotations. In *JCDL '09: Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries*, pages 435–436, New York, NY, USA, 2009. ACM.
- [82] U. Murthy, D. Gorton, R. da S. Torres, M. A. Gonçalves, E. A. Fox, and L. Delcambre. Extending the 5S Digital Library (DL) Framework: From a Minimal DL towards a DL Reference Model. In *1st Workshop on Digital Library Foundations, ACM IEEE Joint Conference on Digital Libraries*, Vancouver, Canada, June 2007.
- [83] U. Murthy, N. P. Kozievitch, E. A. Fox, R. da S. Torres, and E. Hallerman. SuperIDR: A tablet PC tool for Image Description and Retrieval, Workshop on the Impact of Pen-Based Technology on Education (Wipte), Virginia Tech, October, 2010.
- [84] U. Murthy, N. P. Kozievitch, J. Leidig, R. da S. Torres, S. Yang, M. Gonçalves, L. Delcambre, D. Archer, and E. A. Fox. Extending the 5s framework of digital libraries to support complex objects, superimposed information, and content-based image retrieval services. Technical Report TR-10-05, Virginia Tech, Department of Computer Science, April 2010.
- [85] L. Nelson and H. V. de Sompel. IJDL special issue on complex digital objects: Guest editors' introduction. *International Journal of Digital Libraries*, 6(2):113–114, 2006.
- [86] M. L. Nelson, B. Argue, M. Efron, S. Denn, and M. C. Pattuelli. A survey of complex object technologies for digital libraries. Technical report, NASA/TM-2001-211426, 2001.
- [87] M. L. Nelson and K. Maly. Buckets: smart objects for digital libraries. *Communications of the ACM*, 44(5):60–62, May 2001.
- [88] M. L. Nelson, G. Marchionini, G. Geisler, and M. Yang. A bucket architecture for the open video project. In *Proceedings of the 1st ACM/IEEE-CS Joint Conference on Digital libraries*, JCDL '01, pages 310–311. ACM, 2001.
- [89] D. Nichols, I. Witten, T. T. Keegan, D. Bainbridge, and M. Dewsnip. Digital libraries and minority languages. *New Review in Hypermedia and Multimedia*, 11(2):139–155, December 2005.
- [90] P. Nieuwenhuysen. Computerised storage and retrieval of structured text information: Cds/isis version 2.3. *Program: Automated Library and Information Systems*, 25(1):1–18, 1991.
- [91] Object Management Group (OMG). Reusable Asset Specification (RAS) - version 2.2, disponível em <http://www.omg.org/docs/formal/05-11-02.pdf> - último acesso em 05/05/11, url = <http://www.omg.org/docs/formal/05-11-02.pdf>, year = 2002.

- [92] S. H. Park, N. P. Kozievitch, E. A. Fox, N. Short, A. L. Abbott, S. Misra, and M. Hsiao. Model-based fingerprint image quality analysis. *Poster for CESCA (Center for Embedded Systems for Critical Applications) Day, Virginia Tech, Blacksburg, VA, USA*, 5 2010.
- [93] S. H. Park, N. Lynberg, J. Racer, P. McElmurray, and E. A. Fox. HTML5 ETDs. *13th International Symposium on Electronic Thesis and Dissertations (ETD 2010), UT Austin Libraries, Austin*, 06 2010.
- [94] Y. A. Park, U. Karadkar, and R. Furuta. Component-based authoring of complex, Petri net-based digital library infrastructure. In *Proceedings of the 14th European conference on Research and advanced technology for digital libraries, ECDL'10*, pages 22–29, Berlin, Heidelberg, 2010. Springer-Verlag.
- [95] D. C. G. Pedronette and R. da S. Torres. Exploiting contextual information for image re-ranking. In *CIARP'10*, pages 541–548, 2010.
- [96] D. C. G. Pedronette and R. da S. Torres. Shape retrieval using contour features and distance optimization. In *VISAPP*, volume 1, pages 197 – 202, 2010.
- [97] D. C. G. Pedronette and R. d. S. Torres. Distances correlation for re-ranking in content-based image retrieval. *23rd Conference on Graphics, Patterns and Images, 2010, Gramado*, pages 1–8, 2010.
- [98] O. A. B. Penatti and R. S. Torres. Eva: an evaluation tool for comparing descriptors in content-based image retrieval tasks. In *Multimedia Information Retrieval*, pages 413–416, 2010.
- [99] F. Pereira. Mpeg multimedia standards: evolution and future developments. In *MULTIMEDIA '07: Proceedings of the 15th international conference on Multimedia*, pages 8–9, New York, NY, USA, 2007. ACM.
- [100] E. Persoon and K. Fu. Shape Discrimination Using Fourier Descriptors. *IEEE Transactions on Systems, Man, and Cybernetics*, 7(3):170– 178, 1977.
- [101] R. R. Tansley, M. Bass, M. Branschovsky, G. Carpenter, G. McClellan, and D. Stuve. Dspace system documentation, disponível em <http://www.dspace.org/index.php?/architecture/technology/system-docs/index.html> - último acesso em 03/04/2011, 2009.
- [102] D. Rehberger, M. Fegan, and M. Kornbluh. Reevaluating access and preservation through secondary repositories: Needs, promises, and challenges. In *Research and Advanced Technology for Digital Libraries*, volume 4172 of *Lecture Notes in Computer Science*, pages 39–50. Springer Berlin / Heidelberg, 2006.
- [103] W. Reilly, R. Wolfe, and M. Smith. Mit's cwspace project: packaging metadata for archiving educational content in dspace. *International Journal of Digital Libraries*, 6(2):139–147, 2006.

- [104] P. Ruijgrok and M. Slabbertje. Requirements for Management & Storage to support complex objects & ORE in DSpace. In *Knowledge Exchange Group - Research paper - Enhanced E-theses Project Deliverable 9.1*, disponível em <http://igitur-archive.library.uu.nl/DARLIN/2010-0526-200239/UUindex.html>, último acesso em 05/05/11, 2009.
- [105] S. Rumsey and B. O'Steen. OAI-ORE, PRESERV2 and digital preservation, Ariadne, 6, available at <http://www.ariadne.ac.uk/issue57/rumsey-osteen/> - último acesso em 03/03/2011, 2008.
- [106] K. Saidis and A. Delis. Integrating multi-dimensional information spaces. *Second Workshop on Very Large Digital Libraries, in conjunction with the 13th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2009)*, 2009.
- [107] A. Santanchè. *Self Describing Components: Searching for Digital Artifacts on the Web*. PhD thesis, Instituto de Computação, Unicamp, Campinas, SP, 2006.
- [108] A. Santanchè and C. B. Medeiros. Fluid Web and Digital Content Components: from a document-centric perspective to a content-centric perspective. In *Proceedings of the XX Brazilian Symposium on Databases*, pages 10–24, 2005.
- [109] A. Santanchè and C. B. Medeiros. A Component Model and Infrastructure for a Fluid Web. *IEEE Transactions on Knowledge and Data Engineering*, 19(2):324–341, February 2007.
- [110] A. Santanchè, C. B. Medeiros, and G. Z. Pastorello Jr. User-author centered multimedia building blocks. *Multimedia Systems*, 12(4):403–421, March 2007.
- [111] A. Santanchè, M. Mota, D. Costa, N. Oliveira, and C. O. Dalforno. Componere: component-based in web authoring. In *Proceedings of the XV Brazilian Symposium on Multimedia and the Web, WebMedia '09*, pages 12:1–12:8, 2009.
- [112] R. L. T. Santos, P. A. Roberto, M. A. Gonçalves, and A. H. F. Laender. A Web services-based framework for building componentized digital libraries. *Journal of Systems and Software*, 81(5):809–822, May 2008.
- [113] H. Sasaki and Y. Kiyoki. Theory and implementation on automatic adaptive metadata generation for image retrieval. In *Proceeding of the 2006 conference on Information Modelling and Knowledge Bases XVII*, pages 68–82, Amsterdam, The Netherlands, The Netherlands, 2006. IOS Press.
- [114] R. C. Schank. *Tell Me a Story: Narrative and Intelligence*. Northwestern University Press, 1995. Introduction-Morson, Gary Saul.
- [115] R. Shen, N. S. Vemuri, W. Fan, and E. A. Fox. Integration of complex archeology digital libraries: An etana-dl experience. *Information Systems*, 33(7-8):699–723, 2008.

- [116] N. Short, S. Misra, A. L. Abbott, M. Hsiao, N. P. Kozievitch, and E. A. Fox. Latent fingerprint matching. *Poster for CESSA (Center for Embedded Systems for Critical Applications) Day, Virginia Tech, Blacksburg, VA, USA*, 5 2010.
- [117] C. Sperberg-McQueen and L. Burnard. TEI-P5: Guidelines for Electronic Text Encoding and Interchange, by the TEI Consortium, disponível em <http://www.tei-c.org/release/doc/tei-p5-doc/en/Guidelines.pdf> - último acesso em 03/03/11, 2010.
- [118] R. O. Stehling, M. A. Nascimento, and A. X. Falcão. A compact and efficient image retrieval approach based on border/interior pixel classification. In *Proceedings of the eleventh international conference on Information and knowledge management, CIKM '02*, pages 102–109, New York, NY, USA, 2002. ACM.
- [119] H. Suleman and E. A. Fox. Beyond Harvesting: Digital Library Components as OAI Extensions. Technical report, Virginia Tech, 2002.
- [120] M. Swain and D. Ballard. Color Indexing. *International Journal of Computer Vision*, 7(1):11–32, 1991.
- [121] S. P. T. Staples, R. Wayland. The fedora project - an open-source digital object repository management system. *D-lib Magazine*, 9(4), april 2003.
- [122] R. Tansley, M. Bass, D. Stuve, M. Branschofsky, D. Chudnov, G. McClellan, and M. Smith. The dspace institutional digital repository system: current functionality. In *JCDL '03: Proceedings of the 3rd ACM/IEEE-CS joint conference on Digital libraries*, pages 87–97, Washington, DC, USA, 2003. IEEE Computer Society.
- [123] D. Tarrant, B. O'Steen, T. Brody, S. Hitchcock, N. Jefferies, and L. Carr. Using OAI-ORE to Transform Digital Repositories into Interoperable Storage and Services Applications. *Code4Lib Journal*, (6), 2009.
- [124] R. S. Torres, A. X. Falcão, and L. da F. Costa. A Graph-based Approach for Multiscale Shape Analysis. *Pattern Recognition*, 37(6):1163–1174, June 2004.
- [125] C. J. van Rijsbergen. *Information Retrieval*. Butterworth, 1979.
- [126] A. Veerasamy and N. J. Belkin. Evaluation of a tool for visualization of information retrieval results. In *Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval, SIGIR '96*, pages 85–92, New York, NY, USA, 1996. ACM.
- [127] C. C. Vogt and G. W. Cottrell. Predicting the performance of linearly combined ir systems. In *Proc. of the 21st annual international ACM SIGIR conference on Research and development in information retrieval, SIGIR '98*, pages 190–196, New York, NY, USA, 1998. ACM.
- [128] S. L. Weibel and T. Koch. The Dublin Core Metadata Initiative: Mission, Current Activities, and Future Directions. *D-Lib Magazine*, 6(12), December 2000.

- [129] K. Williams and H. Suleman. Translating handwritten bushman texts. In *Proceedings of the 10th annual joint conference on Digital libraries*, JCDL '10, pages 109–118, New York, NY, USA, 2010. ACM.
- [130] I. H. Witten and D. Bainbridge. A retrospective look at Greenstone: Lessons from the first decade, Research Commons at The University of Waikato. 2007.
- [131] I. H. Witten, D. Bainbridge, G. Paynter, and S. Boddie. The greenstone plugin architecture. In *JCDL '02: Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries*, pages 285–286, New York, NY, USA, 2002. ACM.
- [132] R. Yan and A. G. Hauptmann. A review of text and image retrieval approaches for broadcast news video. *Information Retrieval*, 10(4-5):445–484, 2007.
- [133] H.-C. Yang and C.-H. Lee. Image semantics discovery from web pages for semantic-based image retrieval using self-organizing maps. *Expert Systems with Applications: An International Journal*, 34(1):266–279, January 2008.
- [134] S. Yang, A. L. Kavanaugh, N. P. Kozievitch, L. T. Li, V. Srinivasan, S. D. Sheetz, T. Whalen, D. Shoemaker, R. da Silva Torres, and E. A. Fox. CTRnet DL for disaster information services. In *JCDL*, pages 437–438, 2011.
- [135] Q. Zhu, M. A. Gonçalves, R. Shen, L. N. Cassel, and E. A. Fox. Visual semantic modeling of digital libraries. In *ECDL'03: Proceedings of European Conference on Digital Libraries*, pages 325–337, 2003.