

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE INFORMÁTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

**Conversão de Imagens do Formato  
DICOM Visando a Inter-  
Operacionalidade de Sistemas Através  
da *WEB***

por

RENATO RANGEL GUIMARÃES

Dissertação submetida à avaliação,  
como requisito parcial para a obtenção do grau de Mestre  
em Ciência da Computação

Prof. Dr. Jacob Scharcanski  
Orientador

Porto Alegre, outubro de 2002.

## CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Guimarães, Renato Rangel

Conversão de Imagens do Formato DICOM Visando a Inter-Operacionalidade de Sistemas Através da *Web* / por Renato Rangel Guimarães. - Porto Alegre : PPGC da UFRGS, 2002.

113 f.:il.

Dissertação (Mestrado) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, RS – BR, 2002. Orientador: Scharcanski, Jacob.

1.Informática Médica. 2.Sistema de Recuperação da Informação-Medicina 3.Arquivos Médicos 4.Imagens Médicas. 5. XML. I. Jacob Scharcanski. II. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitora: Profa. Dra. Wrana Panizzi

Pró-Reitor de Ensino : Prof. José Carlos Ferraz Hennemann

Pró-Reitor Adjunto de Pós-Graduação: Prof. Jaime Evaldo Fensterseifer

Diretor do Instituto de Informática: Prof. Dr. Philippe Olivier Alexandre Navaux

Coordenador do PPGC: Prof. Carlos Alberto Heuser

Bibliotecária-Chefe do Instituto de Informática: Beatriz Regina Bastos Haro

Se eu não acreditar naquilo que faço, jamais poderei  
realizar um sonho.

## **Agradecimentos**

Gostaria de fazer aqui um agradecimento a todos aqueles que de certa forma tiveram uma contribuição nesta dissertação.

Ao professor Jacob Scharcanski que me orientou, neste trabalho, pela sua insistência, sabedoria, dedicação e cooperação.

À Universidade Católica de Pelotas, pelo incentivo e apoio à pesquisa, tornando possível a realização desta dissertação.

A amiga e incentivadora Paula Prusk Yamim, pela liberdade e entendimento das horas necessárias à pesquisa.

Ao amigo Victor Chim Iser pela sua pouca idade, foi um grande inspirador.

À minha companheira, Cristiane de Freitas Chim, um grande obrigado pela paciência, pelo carinho, enfim, por estar sempre comigo!

A todos os colegas de trabalho, pelo entendimento durante esses anos de intensos estudos. Aos meus amigos e amigas, pelo carinho e incentivo durante esse período.

E obrigado àqueles que não mencionei, mas que fizeram parte deste sonho!

## Sumário

<b>Lista de Figuras.....</b>	<b>8</b>
<b>Lista de Tabelas .....</b>	<b>9</b>
<b>Lista de Quadros.....</b>	<b>10</b>
<b>Lista de Abreviaturas .....</b>	<b>11</b>
<b>Resumo .....</b>	<b>14</b>
<b>1 A Internet como Mecanismo de Integração de Sistemas de Informação .....</b>	<b>16</b>
<b>1.1 A Internet.....</b>	<b>16</b>
<b>1.2 Publicação de Imagens Médicas Através da Internet .....</b>	<b>16</b>
<b>1.3 E-Health .....</b>	<b>16</b>
1.3.1 Fatores Facilitadores para a Integração de Sistemas Providos pela Internet....	17
1.3.2 Perspectivas para a Integração de Sistemas de Saúde Através da Internet .....	18
1.3.3 Prontuário Eletrônico do Paciente na Internet.....	19
<b>1.4 PACS - Sistema de Armazenamento e Comunicação de Imagens .....</b>	<b>24</b>
1.4.1 Componentes .....	25
1.4.2 Estrutura .....	26
1.4.3 Crescimento.....	26
1.4.4 Tecnologias padronizadas e disponíveis .....	27
1.4.5 Considerações sobre os PACS.....	27
<b>1.5 Considerações finais .....</b>	<b>28</b>
<b>2 Tecnologias de Desenvolvimento para a Web.....</b>	<b>29</b>
<b>2.1 Aplicação de Ambiente Web .....</b>	<b>29</b>
<b>2.2 Tecnologias para a Interface .....</b>	<b>30</b>
2.2.1 HTML .....	30
2.2.2 JAVA .....	31
2.2.3 BROWSERS.....	31
2.2.4 CSS .....	31
2.2.5 XML.....	31
<b>2.3 Tecnologias para o Processamento no Servidor .....</b>	<b>32</b>
2.3.1 CGI.....	33
2.3.2 ASP .....	33
2.3.3 PHP .....	33
2.3.4 ISAPI/NSAPI.....	33
2.3.5 Servlets .....	33
2.3.6 JSP.....	34
2.3.7 ColdFusion .....	34
2.3.8 Banco de Dados .....	34
<b>2.4 Tecnologia de Protocolo .....</b>	<b>34</b>
<b>2.5 Tecnologia Utilizada em Sistemas de Imagens Médicas.....</b>	<b>35</b>

2.5.1	PACS.....	35
2.5.2	DICOM.....	35
2.5.3	HL7 .....	36
2.5.4	CORBA .....	36
2.5.5	SGML.....	37
<b>2.6</b>	<b>Considerações Finais .....</b>	<b>37</b>
<b>3</b>	<b>Padrões e Arquiteturas de Metadados.....</b>	<b>38</b>
<b>3.1</b>	<b>Padrões de Metadados.....</b>	<b>38</b>
3.1.1	Dublin Core .....	38
3.1.2	SAIF .....	40
<b>3.2</b>	<b>Arquiteturas de Metadados.....</b>	<b>41</b>
3.2.1	Arquitetura Warwick.....	41
3.2.2	Arquitetura de Descrição de Conteúdo (MCF).....	42
3.2.3	Arquitetura de Descrição de Recurso (RDF).....	43
3.2.4	Arquitetura Segundo Chang e Zhang.....	45
<b>3.3</b>	<b>Arquitetura de Modelagem em 4 Camadas Segundo Kerhervé .....</b>	<b>46</b>
<b>3.4</b>	<b>Considerações finais .....</b>	<b>48</b>
<b>4</b>	<b>Proposta de um Modelo para Representar Imagens DICOM... 49</b>	
<b>4.1</b>	<b>Modelo para Geração de Interfaces para Imagens .....</b>	<b>49</b>
<b>4.2</b>	<b>Metadados.....</b>	<b>50</b>
<b>4.3</b>	<b>Padrão Normalizado.....</b>	<b>51</b>
<b>4.4</b>	<b>Dublin Core.....</b>	<b>52</b>
4.4.1	Análise do Padrão .....	53
4.4.2	Análise do Formato.....	54
4.4.3	Formato de Metadado da Ferramenta Proposta.....	54
<b>4.5</b>	<b>A Recuperação de Informações Usando o Modelo Proposto .....</b>	<b>56</b>
4.5.1	Modelo .....	56
4.5.2	Resultados .....	56
4.5.3	Discussão.....	58
<b>4.6</b>	<b>Considerações Finais .....</b>	<b>58</b>
<b>5</b>	<b>Protótipo do Sistema Proposto .....</b>	<b>60</b>
<b>5.1</b>	<b>Ambiente dinâmico.....</b>	<b>60</b>
5.1.1	Sistema Operacional .....	60
5.1.2	Java .....	61
5.1.3	PHP .....	61
5.1.4	XML.....	61
<b>5.2</b>	<b>Características do Protótipo.....</b>	<b>62</b>
5.2.1	Arquivos Gerados .....	62
5.2.2	Diretórios da Ferramenta.....	63
5.2.3	Estruturação Interna (variáveis).....	64
<b>5.3</b>	<b>Estruturação dos módulos.....</b>	<b>65</b>
5.3.1	Características de Implementação .....	65
5.3.2	Módulo de Controle da Interface do Usuário.....	67
5.3.3	Módulo de Controle dos Metadados .....	67
5.3.4	Módulo de Controle do Repositório de Imagens.....	67
<b>5.4</b>	<b>Diagrama do Protótipo.....</b>	<b>69</b>
5.4.1	Repositório de dados do Protótipo.....	69

<b>5.5</b>	<b>Definição .....</b>	<b>70</b>
5.5.1	Avaliação do Desenvolvimento.....	71
<b>5.6</b>	<b>Problemas Encontrados e Soluções Propostas.....</b>	<b>71</b>
<b>5.7</b>	<b>Funcionalidade e Especificação da Ferramenta e Apoio ao Modelo Proposto - DICOM DATABASE - .....</b>	<b>72</b>
<b>5.8</b>	<b>Organização da Integração do Protótipo.....</b>	<b>73</b>
5.8.1	Execução da tarefa para Extrair Cabeçalho DICOM.....	76
5.8.2	Execução da tarefa para Gerar RDF .....	78
<b>5.9</b>	<b>Utilização do Esquema em XML.....</b>	<b>80</b>
<b>5.10</b>	<b>Considerações Finais .....</b>	<b>82</b>
<b>6</b>	<b>Conclusões .....</b>	<b>84</b>
6.1	Confiabilidade.....	84
6.2	Contribuições.....	85
6.3	Trabalhos Futuros .....	85
<b>Anexo 1 Formatos para Imagens Médicas.....</b>		<b>89</b>
<b>Anexo 2 Modelagem Conceitual de Metadados para Imagem .....</b>		<b>97</b>
<b>Bibliografia.....</b>		<b>107</b>

## Lista de Figuras

FIGURA 1-1 - Componentes de PACS em ambiente Cooperativo Cliente/Servidor ....	25
FIGURA 2-1 - Fluxo de uma aplicação na Web. ....	30
FIGURA 3-1 - Quadro DLG representando dados MCF.....	43
FIGURA 3-2 - Exemplos RDF.....	44
FIGURA 3-3 - Arquitetura de Metadados.....	46
FIGURA 3-4 - Arquitetura de modelagem segundo Kerhervé em 4 camadas.....	47
FIGURA 4-1 - Categorias de metadados em imagens digitais em documentos. ....	54
FIGURA 5-1 - Estrutura do sistema .....	68
FIGURA 5-2 - Diagrama interno do Protótipo.....	69
FIGURA 5-3 - Tela da aplicação via Browser. ....	74
FIGURA 5-4 - Botões de execução .....	75
FIGURA 5-5 - Tela de envio das imagens ao repositório.....	76
FIGURA 5-6 - Tela de resultado da extração de conteúdo da imagem .....	77
FIGURA 5-7 - Tela de visualização da imagem. ....	78
FIGURA 5-8 - Tela de opção de atributos DICOM para RDF –Dublin Core. ....	79
FIGURA 5-9 - Tela de visualização de DICOM para RDF- Dublin Core.....	79
FIGURA 6-1 - Esquema Conceitual de metadados para Imagem.....	99
FIGURA 6-2 - Exemplo de Imagem Original.....	105
FIGURA 6-3 - Diagrama de Instâncias da figura exemplo FIGURA 6-2 .....	106

## **Lista de Tabelas**

TABELA 3.1 - Elementos do Padrão Dublin Core.....	40
TABELA 4.1 - Descritores do padrão Dublin Core .....	52
TABELA 4.2 - Mapeamento DICOM para Dublin Core.....	55
TABELA 4.3 - Resumo das Características dos Padrões.....	57
TABELA 6.1 - Características técnicas dos formatos de arquivo. ....	92
TABELA 6.2 - Descrição da Classe .....	103

## **Lista de Quadros**

QUADRO 1.1- Vantagens do Prontuário Eletrônico do Paciente baseado na Web.....	20
QUADRO 6.1 - Valores de referência de tags DICOM.....	92

## Lista de Abreviaturas

ACR	American College of Radiology
API	Application Program Interface
ASP <sup>1</sup>	Application Service Provider
ASP <sup>2</sup>	Active Server Pages
BMP	Microsoft Windows Bitmap
CGI	Common Gateway Interface
CMYK	Cyan, Magenta, Yellow, Black
CR	Radiografia Computadorizada
CT	Tomografia Computadorizada
DAS	Angiografia de Subextração Digital
DBMS	Database Management Systems
DC	Dublin Core
DICOM	Digital Imaging and Communications in Medicine
GIF	Graphics Interchange Format
HIS	Hospital Information System
HL7	Health Level Seven
HTML	HyperText Markup Language
IDL	Interface Definition Language
ISAPI	Information Server Application Programming Interface
JPEG	Joint Photographic Experts Group,
LZW	Limpel, Ziv ,Welch
MARC	Machine Readable Catalogue
MCF	Meta Content Framework

MR	Ressonância Magnética
MRI	Magnetic Resonance Imaging
NEMA	National Electrical Manufacturers Association
NIU	Network Interface Unit
NM	Medicina Nuclear
ODMG	Object Data Management Group
OMA	Object Management Architecture
OMG	Object Management Group
OSI	Open Systems Interconnection
PACS	Picture Archiving and Communication Systems
PEP	Prontuário Eletrônico do Paciente
PHP	Pré-processador de Hipertexto
PNG	Portable Networks Graphics
RDF	Resource Description Framework
RGB	Red, Green, Blue
RIS	Radiology Information System
SGML	Standard Generalized Markup Language
SIH	Sistema de Informação Hospitalar
SPIFF	Still Picture Interchange File Format
SQL	Structured Query Language
SSL	Secure Sockets Layer
SSP	Server Side Processing
Tag	Marcação
TCP/IP	Transmission Control Protocol/Internet Protocol
TIFF	Tagged Image File Format

UML	Unified Modeling Language
URI	Universal Resource Identifier
URL	Universal Resource Location
US	Ultra Sonografia
VIR	Visual Information Retrieval
W3C	World Wide Web Consortium
WWW	World Wide Web
XML	Extensible Markup Language

## Resumo

Este trabalho apresenta uma proposta para a publicação eletrônica de imagens e metadados do formato DICOM, visando à interoperabilidade de sistemas heterogêneos, que têm a capacidade de se adaptarem às particularidades do ambiente médico, através da *Web*, proporcionando assim, ambientes propícios à área da saúde.

Atualmente, para desenvolver ambientes de visualização de imagens na *Web*, são utilizados padrões de imagens, conceitos de metadados e adaptação de tecnologias para este fim.

Utilizando o formato padrão de imagens médicas DICOM, com este trabalho se propõe construir um *Framework* para visualizar e trocar informações através da *Web*. Entre as suas aplicações estão a telemedicina e a integração de imagens aos sistemas de informação na área da saúde.

**Palavras-chave:** Metadado, DICOM, Dublin Core, Imagens Digitais, Sistemas de Informação.

**TITLE:** “CONVERSION OF IMAGES OF THE DICOM FORMAT SEEKING INTERACTION OF SYSTEMS FROM WEB”

### **Abstract**

The present work presents a proposal for electronic publishing of images and metadata of the DICOM format. It aims interoperation of heterogeneous systems, which have the capacity to adapt themselves to the specificities of medical environment, via *Web*, providing this way, proper environments to the health area.

Nowadays, in order to develop environments for visualizing images on the Web, it is necessary to use patterns of images, metadata concepts and adaptations of technologies for this purpose.

Throughout the use of standard format of medical images DICOM, this work proposes the construction of a *Framework* to visualize and exchange information via Web. Among its applications are the telemedicine and the image integration to the information systems in the health area.

**Key-words:** Metadata, DICOM, Dublin Core, Digital Images, Information system.

# 1 A Internet como Mecanismo de Integração de Sistemas de Informação

## 1.1 A Internet

A *Web* surgiu em 1969, com o projeto de Pesquisa Avançada dos Estados Unidos, cuja necessidade de criação nasceu devido à descoberta da importância do uso de computadores na pesquisa, na educação e, principalmente, na indústria militar. Um momento posterior da grande importância na história da *Web* foi a criação da NFSNET, em 1986, que conectava redes acadêmicas federais e redes de nível intermediário, que ligavam universidades e centros de pesquisas. Na época em que a NFSNET foi criada, a *Web* começou a crescer. Além do crescimento exponencial, outro fator que certamente contribuiu muito para o crescimento e aceitação da *Web*, foi seu sucesso na obtenção de inter-operacionalidade [MAR 97].

A presença da *Web* mudou o modo de fazer negócios: antes se dizia que a informação era essencial para a empresa, hoje ela é o "negócio". A saúde tem um mercado distribuído e heterogêneo, que precisa compartilhar informações e isso é feito, atualmente, na maioria das vezes, de forma manual, principalmente no Brasil, e na maioria dos casos isso não ocorre bem. Além dessa necessidade de troca de informação, a prática médica diária necessita, invariavelmente, de orientação para funcionar, sem a qual é impossível um adequado atendimento ao paciente. Isso torna a saúde uma indústria altamente dependente de informação, tornando a *Web* uma solução para muitas dessas questões.

## 1.2 Publicação de Imagens Médicas Através da Internet

Na atualidade, apresentar a alternativa de utilização do WWW, como meio de integração entre repositórios de imagens médicas, é um trabalho para diminuir distâncias e custos. Diante disso, os expressivos progressos, associados ao aparecimento e aperfeiçoamento de novas tecnologias, fizeram, do diagnóstico por imagem uma das mais excitantes áreas da medicina. Com um novo modelo, WWW, está iniciando o desenvolvimento de novas ferramentas, para disponibilizar esta informações. Esta integração só poderá ser realizada através de novas idéias, buscando facilitar cada vez mais a tarefa de desenvolvimento e de acesso aos inúmeros serviços que surgirão, a partir de uma solução [SAB 99,TAC 96].

## 1.3 E-Health

Nesse cenário de uma indústria que precisa compartilhar a informação, a *Web* surgiu como a possibilidade de dar suporte a todas essas necessidades, além de oferecer um novo conjunto de aplicações que, antes, eram de difícil realização ou mesmo, simplesmente, não existiam. Assim, a *Web* na área de saúde vem propiciando o

surgimento de novas tecnologias, como é o caso da telemedicina, possibilitando oferecer uma grande quantidade de benefícios e de serviços agregados, tais como:

- melhorar o conhecimento dos pacientes;
- melhorar o intercâmbio eletrônico de dados;
- facilitar os relacionamentos entre pacientes, médicos e prestadores de serviço;
- melhorar o acesso às informações vitais do paciente;
- reduzir custos e melhora da qualidade.

O conceito de e-health engloba as aplicações da *Web* na área da saúde, sejam processos administrativos-financeiros, assistências ou informativos. E-health é definida como "um campo emergente na interseção da informática médica, saúde pública e negócios, relativo a serviços de saúde e informações transmitidas através da *Web* e tecnologias relacionadas" [EYS 2001, MAT 2001].

### **1.3.1 Fatores Facilitadores para a Integração de Sistemas Providos pela Internet**

A *Web* oferece alguns fatores facilitadores que, segundo [KIL 2000], podem ser divididos em três categorias principais:

- **Integração com sistemas legados:** no ambiente de saúde, há um grande número de sistemas de informação de diferentes aplicações, com diferentes interfaces, vocabulários e modelo de dados, adaptados em plataformas de hardware e software diferentes. O ambiente *Web* permite a integração entre estes sistemas, através de uma interface comum, utilizando os navegadores da *Web* (*browsers*), tais como o Netscape e o Internet Explorer, que acessam os dados legados, de forma distribuída e também unificam os modelos de dados. Todo esse processo necessita padronização de interfaces e mapas entre os diversos vocabulários de modo que a tecnologia *Web* seja capaz de oferecer integração rápida e simples.
- **O aumento do consumo:** nos últimos anos, com o aumento do poder de compra dos consumidores, um grande mercado surgiu em resposta a essa demanda. Esse crescente aumento de serviços de comércio eletrônico em outras áreas faz crescer as expectativas por serviços dessa natureza no mercado de saúde, sendo a *Web* o meio pelos quais esses podem chegar facilmente ao alcance dos novos consumidores.
- **Características próprias da Tecnologia *Web*:** a tecnologia *Web* oferece muitas características que favorecem a rápida penetração em todos os aspectos do negócio, tendo a área da saúde a oportunidade de aproveitá-las. Características como acessibilidade via *Web browsers* com rápida ou nenhuma curva de aprendizado, rapidez de distribuição visto que não se têm preocupações com questões relacionadas ao sistema operacional das máquinas clientes e, sendo os sistemas baseados na *Web* mais fáceis e mais baratos para desenvolver e manter do que sistemas cliente/servidor, e *mainframes* tornam a tecnologia *Web* extremamente adequada para qualquer tipo de empresa.

Alguns pontos foram destacados na literatura, afim de que a tecnologia *Web* seja adotada com sucesso como integradora de informações na área de saúde [GRU 99, KIL 2000]:

- desenvolvimento de uma estratégia de e-health;
- estabelecimento de um servidor seguro de dados seguro;
- planejamento do desenvolvimento de um *website* com recursos, no qual concentre todas as aplicações;
- desenvolvimento de uma estratégia de conteúdo para a comunidade (pacientes, médicos etc.);
- estabelecimento de uma estratégia específica para os médicos;
- implementação de uma estratégia de conectividade das atividades vinculadas ao atendimento da população.

A existência de um sistema de publicação de imagens e metadados de pacientes via *Web* requer um grande esforço de padronização de formatos de dados. O sucesso desta aplicação está diretamente relacionado com a integração de dados com os *browsers*, proporcionando uma velocidade maior de visualização das imagens através da *Web*, e troca de dados de paciente de maneira eficiente entre sistemas de informação, que podem ter modelos de dados diferentes [MAT 2000].

### 1.3.2 Perspectivas para a Integração de Sistemas de Saúde Através da Internet

A expansão da *Web* na área de saúde deverá ser contínua. Entretanto, para que esse crescimento ocorra, algumas barreiras precisam ser vencidas nos próximos anos [CAI 99]: segurança, qualidade duvidosa das informações em saúde, em determinados *sites*, cultura médica conservadora e resistente a novas tecnologias, demora para adequação dos sistemas legados às mesmas. A expansão da *Web* associada à concorrência do capital do conhecimento não foi suficiente para o crescimento destes recursos, mas causou um excesso de padrões.

Algumas tendências atuais devem continuar a progredir, invadindo os consultórios médicos e hospitais nos próximos anos. Também se destacam algumas aplicações *Web*, que serão largamente utilizadas nos próximos anos (algumas já em uso atualmente): Prontuário Eletrônico do Paciente na *Web* (inteligentes, disponíveis globalmente, 24h por dia, 7 dias por semana), Aplicações Wireless (PDAs etc.) e, por fim, a Convergência Digital (rádio, TV, fax, telefone e *Web* juntos numa só mídia) [SAB 98].

Com base nestas tendências, pode-se antecipar algumas aplicações que liderarão o segmento e-health nos próximos anos, visando um melhor entendimento de doenças e aumentar a eficiência e a qualidade dos sistemas de informações médicas, como [CAI 99, MCM 99]:

- serviços de informação em saúde para pacientes;
- grupos de apoio on-line para pacientes;
- serviços de educação médica continuada;
- comunicação via e-mail entre médicos e pacientes;
- serviços de transação e infraestrutura de comunicação (conectividade, servidores de aplicativos);
- prontuário Eletrônico do Paciente na *Web*.

### 1.3.3 Prontuário Eletrônico do Paciente na Internet

Com a evolução dos equipamentos para obter imagem e sinais em medicina, cada vez mais documentos de mídias diferentes têm sido agregados ao prontuário e, com isso, surgiu o chamado “registro médico multimídia”. Com a *Web* e o advento da World Wide Web (WWW), ou simplesmente *Web*, a existência de um registro médico multimídia tornou-se mais viável, criando-se a figura do “Web-based Electronic Medical Record” ou, em português, Prontuário Eletrônico do Paciente (PEP) baseado na “*Web*”.

O PEP na *Web* torn-se possível, devido ao da *Web* e dos *browsers* (Internet Explorer, Netscape Navigator, etc.) possibilitarem uma nova forma de acesso interativo à informação, com a capacidade de combinar várias mídias num único ambiente. Além disso, a tecnologia *Web* permite mais facilmente a integração entre sistemas heterogêneos, unificando de certa forma os diversos modelos de dados apresentados aos usuários.

Com isso, o tradicional modelo cliente-servidor está sendo gradativamente substituído por uma arquitetura multi-camadas com um "cliente magro" (thin client - nome dado ao computador cliente com pouca capacidade de processamento, sendo este processo realizado no computador servidor). Basicamente, temos três camadas na arquitetura *Web* [SAF 2000]:

- uma camada de apresentação, que normalmente é um *browser* que faz a interface com o usuário;
- uma camada intermediária “middleware”, que é um conjunto de objetos ou softwares no servidor;
- uma coleção heterogênea de repositórios de dados.

Esse modelo traz algumas vantagens como requisitos mínimos para o cliente (thin client) e um conteúdo multimídia bastante rico. Aproveitando as vantagens desse modelo *Web*, o PEP baseado na *Web* traz, além dessas, algumas outras vantagens destacadas no QUADRO 1.1. Em contrapartida, traz também novas preocupações bem como ressalta problemas já existentes no PEP convencional (não baseado na *Web*),

sendo muito importante à questão relacionada à segurança e confidencialidade [DET 2001].

<b>O PEP na Web permite:</b>
◆ multimídia (gráficos, imagens, vídeo e som);
◆ flexibilidade para integração com outros sistemas e sistemas legados;
◆ hipertexto: links para guidelines, jornais e fontes de conhecimento médico;
◆ uso de <i>browsers</i> (baixo custo);
◆ mecanismos de visualização e navegação usualmente mais adaptáveis e mais fáceis de usar;
◆ acessibilidade de vários locais, 24h por dia, 7 dias por semana;
◆ menor custo de desenvolvimento, comparado a tecnologias como cliente/servidor ou <i>mainframe</i> ;
◆ suporte à vários sistemas operacionais de diferentes fabricantes;
◆ facilidade de integração de dados dinâmicos com documentos estáticos;
◆ manutenção centralizada, distribuição imediata;
◆ redução do Custo Total da Propriedade

QUADRO 1.1- Vantagens do Prontuário Eletrônico do Paciente baseado na Web.

#### 1.3.3.1 Preparando o Ambiente para o PEP na Internet:

Um sistema de PEP requer uma estrutura onde serão desenvolvidas as atividades, isto é, um ambiente computacional para o uso de tecnologias em um PEP baseado na *Web*. Para tal são necessários [KOH 99]:

1) Uma Equipe de Projeto: no projeto de um sistema baseado na *Web*, a equipe de desenvolvimento deve conter, além dos vários tipos de profissionais normalmente envolvidos, outros com habilidades específicas para a construção de conteúdo e design das páginas *Web*. Webdesigners, webmasters e programadores para a *Web* são exemplos desses tipos de profissionais. Além disso, a equipe deve desenvolver padrões de página que permitam fácil navegação, tenham uma orientação consistente para uma fácil manutenção e um design visual adequado.

2) Identificação das aplicações *Web* mais adequadas para a instituição, e para isto algumas ações devem ser realizadas:

- estabelecer o domínio do projeto;
- identificar as necessidades do usuário;
- priorizar os requisitos do negócio;

- conhecer soluções prontas de empresas da área.

Uma vez que todos esses pontos tenham sido considerados, deve tentar-se visualizar os recursos a serem implementados e desenvolver as regras do desenvolvimento; e, finalmente, definir os hardwares, softwares e serviços a serem desenvolvidos.

3) Determinação dos recursos necessários para o desenvolvimento do projeto, considerando: pessoal, treinamento, hardware e software, upgrades na rede, manutenção, criação de conteúdo, manutenção da informação, monitoramento da segurança e controle de acesso, desenvolvimento das aplicações, esforços para a padronização e upgrades dos softwares.

### 1.3.3.2 Servidores de Aplicativos

Com a *Web*, um novo modelo de utilização de softwares passou a ser possível, no qual o software não é mais um produto e sim, um serviço. Neste modelo, a infraestrutura computacional centralizadora (servidores e softwares) fica remotamente alocada em um ambiente próprio da organização ou empresa servidora de aplicativos, conhecida como *Application Service Provider*, ou simplesmente, ASP<sup>1</sup> (não confundir com ASP<sup>2</sup> de *Active Server Pages*, que é uma tecnologia para desenvolvimento de páginas *Web* dinâmicas da Microsoft). O ASP<sup>1</sup> disponibiliza e gerencia aplicações e serviços computacionais de centros de dados remotos para múltiplos usuários via *Web* ou via rede privada. O software tradicional tem algumas desvantagens, quando comparados ao modelo ASP: o investimento inicial em geral é alto e há uma necessidade contínua para manutenção, atualização e customização. O modelo ASP<sup>1</sup> é uma alternativa viável a clientes que precisam implementar sistemas complexos. Os benefícios de usar um ASP<sup>1</sup> incluem [ADE 2001]:

1) Velocidade para o mercado: o ASP<sup>1</sup> já tem equipamentos, aplicativos e experiência para prover um acesso rápido à comercialização e administração dos serviços.

2) Liberdade operacional: devido ao gerenciamento externo dos aplicativos, o cliente pode focar seus recursos na sua real atividade empresarial.

3) Melhor desempenho: ASP<sup>1</sup>s podem aplicar sua vasta experiência para implementar as melhores técnicas de tecnologia da informação para níveis superiores de disponibilidade, segurança, cópia de arquivos (backups), recuperação em caso de desastre e suporte ao usuário.

4) Flexibilidade financeira: este modelo reduz custos e gastos em hardware, aplicativos e gerenciamento.

5) Redução de riscos: com nenhum capital investido em software, hardware e nem em pessoal, as organizações podem testar uma nova tecnologia com um mínimo impacto em seu ambiente operacional já existente.

6) Acesso de múltiplos usuários sem necessidade de redes privadas.

7) Não há preocupações com atualizações, pois os aplicativos estão centralizados num único servidor.

8) Os dados estão mais seguros, pois contam com uma forte estrutura de hardware e gerenciamento. Considerando-se esses e outros fatores, o uso de Prontuário Eletrônico do Paciente, baseado na *Web*, utilizando o modelo ASP<sup>1</sup>, é uma alternativa para implementar sistemas dessa natureza. Já existem diversas soluções que oferecem esse tipo de serviço, em diversos níveis, ou melhor, desde simples sistemas de PEP para ambulatório até complexos sistemas de informação hospitalar.

### 1.3.3.3 Segurança e Confidencialidade

Talvez os principais obstáculos para o uso de PEP na *Web* sejam as questões relacionadas à segurança e confidencialidade dos dados do paciente. Isso tem gerado grande discussão e, sem dúvida, precisa ser ainda mais discutido, seja para se chegar a conclusões e melhorias no processo de segurança, seja para o esclarecimento do público (usuários, médicos, pacientes, etc.) a respeito da real segurança que os dados armazenados em meio eletrônico oferece e quais são as suas vantagens e desvantagens.

Para evitar acesso por usuários não autorizados, diversos métodos de segurança podem ser implementados. Certificação digital, SSL (*Secure Sockets Layer*) ou tecnologias de firewall (computador ou software que forma uma barreira de proteção, evitando o acesso de usuários não autorizados) garantem ao usuário que o acesso será feito somente por aqueles que possuem o login e a senha corretos.

O SSL permite que a troca de informações entre o cliente “*browser*” e o servidor seja criptografada, garantindo alto nível de segurança. As tendências atuais são de possibilitar as chamadas autenticações fortes, como a impressão digital e a imagem digitalizada da íris que, realmente, identificam o indivíduo de forma unívoca e sem possibilidade de erro.

### 1.3.3.4 Acurácia do diagnóstico

A área de diagnóstico, principalmente aquela que utiliza imagens depende muito da qualidade e precisão dos dados analisados. Há necessidade de proporcionar ao paciente e ao médico exames de alta qualidade e conseqüentemente proporcionar um custo-benefício favoráveis, como são os exames de Ultra-som, Raios-x, Ressonância Magnética, Tomografia Computadorizada, Medicina Nuclear, entre outros. Para um diagnóstico correto é necessário integrar os diversos métodos de diagnóstico por imagem de forma racional. Além disso, se necessita um acesso rápido aos dados clínicos do paciente que, com a análise das imagens, possibilitarão um diagnóstico preciso. O recente advento do padrão DICOM para imagens médicas e o desenvolvimento de programas e computadores específicos para processamento de imagens médicas possibilita a criação de um ambiente computacional juntamente com um conjunto de dados proporcionará uma forma confiável de modo a integrar as imagens provenientes dos diversos equipamentos médicos com os dados clínicos e administrativos. Desta forma o profissional que manipula tipo de dado pode acessar exames realizados na clínica em qualquer equipamento os dados clínicos e exames do paciente. Esta integração possibilita uma maior precisão diagnóstica com melhor utilização dos recursos disponíveis. Ainda, exames anteriores podem ser digitalizados em um digitalizador para fazer parte do conjunto de dados que será avaliado pelo

médico poderem estar disponível na rede. A coluna mestra para o sucesso desse tipo de sistema, é a escolha de um em software que seja desenvolvido preferencialmente em arquitetura aberta. O princípio básico da a sua implantação é que seja de forma gradativa, no qual consiste das seguintes etapas:

- Captura de imagens;
- Transmissão de imagens;
- Processamento e visualização de imagens;
- Impressão de imagens;
- Armazenamento e recuperação de imagens.

Desta forma, se pode simultaneamente analisar imagens de um paciente em uma mesma tela, além de ver a extração dos dados implícitos na imagem, dando mais consistência na identificação entre paciente-imagem. Além de aumentar a acurácia diagnóstica esta integração sistêmica deve permitir ao médico, uma maior dedicação e tempo ao paciente.

#### 1.3.3.5 Vantagens da Interoperabilidade de Sistemas de Informação de Saúde Através das *Web*.

São numerosas as vantagens da troca de informações em relação aquela baseada em papel, as quais vai desde questões ligadas ao melhor acesso, até maior segurança e, principalmente, oferta de novos recursos, os quais não poderiam existir em papel, tais como: apoio à decisão, troca eletrônica dos dados entre instituições, especificamente no caso do PEP. É importante destacar que, ao contrário do que se imagina, o PEP é muito mais seguro e tem maior possibilidade de manter a confidencialidade do que os prontuários em papel, no qual o risco de ocorrer um acesso não autorizado é maior, que num sistema eletrônico. É possível, também, através do PEP, obter-se uma melhoria na qualidade da assistência à saúde do paciente, melhor gerenciamento dos recursos, melhoria de processos administrativos e financeiros e, ainda, a possibilidade para avaliação da qualidade [MCD 90].

## 1.4 PACS - Sistema de Armazenamento e Comunicação de Imagens

Num ambiente hospitalar, as imagens mais comuns em estudo são os Raios-X, compreendendo aproximadamente 70% do total. Atualmente se produzem imagens de radiografia computadorizada (CR), tomografia computadorizada (CT), ressonância magnética (MRI), ultra-sonografia, medicina nuclear (NMI) e angiografia de subextração digital (DAS), entre outras que ocupam 30% do restante.

A grande quantidade de imagens, produzidas para diagnóstico, requer um complicado manejo, quando depende de armazenamento em arquivo. Uma alternativa de manipular imagens digitais, de forma eficiente, é através de dispositivos conectados em rede, que, em conjunto, oferecem uma série de serviços que dão suporte a operacionalidade no ambiente médico. Para obter uma boa aceitação nos resultados médicos, deve considerar-se a facilidade, rapidez, segurança, acesso e a qualidade de apresentação das imagens [MAR 98, WON 97].

O uso de sistemas de comunicação e armazenamento de imagens, também chamados de PACS, são sistemas capazes de armazenar, recuperar e manter informações. Neste contexto, informações podem ser constituídas por textos, imagens, vídeos ou composições destes tipos. Já os dados sob a forma textual são de responsabilidade do HIS (*Hospital Information Systems*) administrar o fluxo de informações referentes ao paciente [PAC 2000].

Devido à portabilidade e adequação a qualquer tipo de imagens, os PACS devem ser criados seguindo alguns critérios desenvolvidos por vários fabricantes, devido a grande quantidade de diferentes sistemas operacionais, ou seja, evoluindo uma arquitetura aberta com diversos equipamentos agrupados em uma topologia distribuída.

Diante destas características e a extensibilidade dos PACS, é necessário encontrar ferramentas de manipulação, processamento e armazenamento da imagem para serem inseridas no contexto de novas tecnologias. É necessário introduzir sistemas que realizem o manejo de imagens, baseado na importância das necessidades de integrar o hospital a este tipo de sistema. Com isso, o ambiente hospitalar precisa estar disposto a explorar alguns pontos, como:

- registros de admissão de um paciente;
- quantidade de pacientes atendidos no serviço de radiologia;
- relação paciente e hospital;
- manipulação de imagens médicas de diferentes áreas do conhecimento;
- existência de outros sistemas de informações existentes no hospital;
- mecanismo de pesquisa e estudo do serviço de radiologia;
- diferenciar áreas do hospital que querem consultar as imagens;
- diferenciar áreas físicas envolvidas;
- retornar problemas de informações perdidas;

- localizar informação quando o paciente sai do hospital;
- proporcionar consultas posteriores da informação;
- intercâmbio de informações com outros hospitais.

Para abranger as necessidades descritas acima se requer um conjunto de dispositivos, cujas responsabilidades são oferecer suporte a todos os elementos operacionais do ambiente que trata com imagens. Esta necessidade cria uma tecnologia de multimídia que tem, como foco, a descoberta em benefício de procedimentos para o diagnóstico. A integração de PACS e HIS visa fornecer de forma transparente para a equipe a médica um conjunto de informações composto por dados e imagens do paciente para a formação do diagnóstico [HUA 97].

### 1.4.1 Componentes

Os sistemas PACS utilizam vários componentes (hardware e software) com funções específicas para poder integrar o ambiente médico com os vários tipos de dados oriundo dos pacientes. Sistemas PACS, como mostra a FIGURA 1-1 são constituídos de vários componentes e serviços compartilhados, oferecendo um ambiente de visualização e armazenamento de imagens. Estes componentes se integram a partir de uma rede de computadores e equipamentos criando um ambiente Cliente/Servidor, a fim de oferecer um serviço onde são conectados todas as modalidades de imagem em radiologia com o objetivo de eliminar o filme.

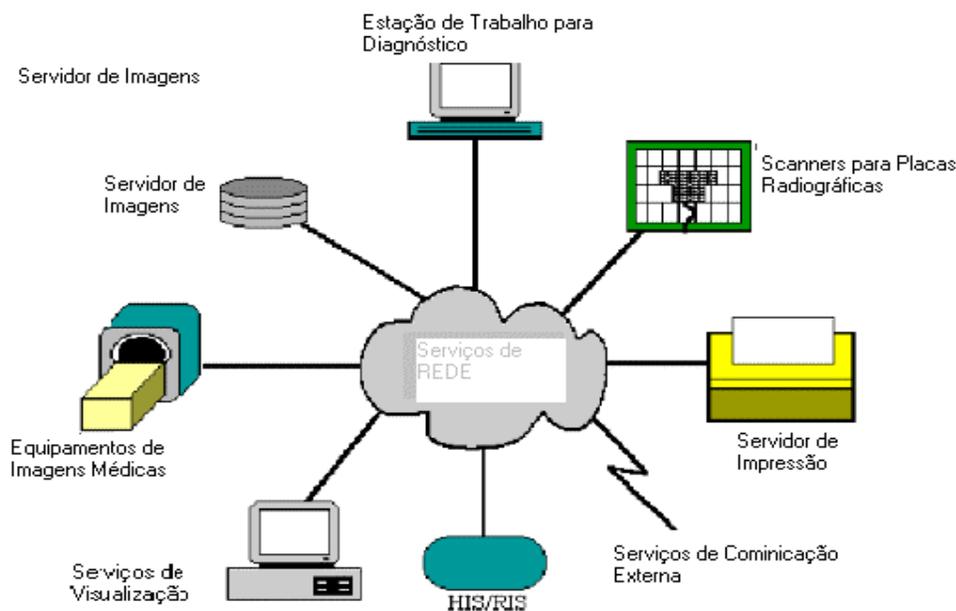


FIGURA 1-1 - Componentes de PACS em ambiente Cooperativo Cliente/Servidor

### 1.4.2 Estrutura

No manejo de informações dentro do ambiente de saúde, por meio de uma rede de computadores, surgiu inicialmente, com o objetivo de desenvolver tecnologias de visualização. Diante disso, o PACS tem como objetivo o armazenamento de imagens, a rapidez de acesso e transmissão das imagens médicas de forma que a sua utilização possa trazer uma qualidade no serviço de saúde com o mínimo de custo.

Nestes últimos anos, a busca em grande escala de bancos de dados de imagens tem atraído a atenção da comunidade de informática. Há algum tempo atrás a capacidade de armazenamento em ambiente de redes eram grandes desafios e deixam a desejar reduzindo a velocidade, dos dispositivos de exibição visuais. A penetração de mercado era muito baixa para a maioria das ferramentas voltadas para o desenvolvimento de banco de dados de imagens.

Recentes aumentos na capacidade de armazenamento, em ambiente de redes, trouxeram um poder de processamento e resolução na exibição e deram um crescimento no desenvolvimento de bancos de dados de imagens, fazendo que o PACS atinja uma evolução no seu conteúdo. Literalmente, centenas de projetos, como a vídeo-endoscopia e vídeo-microscopia, começaram nos últimos anos a ser desenvolvidos [MAR 2000, WON 97].

O PACS é constituído de alguns módulos básicos:

- identificação: identifica o paciente, data do exame e as características da imagem;
- aquisição da imagem: obtenção, seleção e armazenamento da imagem;
- comunicação: comunicação, transmissão e transporte da imagem;
- extração de característica: digitalização da imagem, levando em conta os aspectos visuais;
- resultados gráficos: parâmetros digitalizados e suas características fisiológicas;
- estudo comparativo: comparação simultânea entre vários estudos feitos de imagens para tornar uma análise mais próxima entre resultados anteriores;
- apresentação: mostrar os resultados obtidos para uma posterior apresentação.

### 1.4.3 Crescimento

O expressivo progresso da radiologia nas últimas décadas, associado ao aparecimento e aperfeiçoamento de novas tecnologias, fizeram do diagnóstico por imagem o maior impacto desta abordagem. O principal motivo que desencadeou o desenvolvimento e implantação foi à possibilidade de otimização no processo de disponibilização e consulta de laudos pela rede “ethernet” do hospital [MAR 2000a].

Outro aspecto é o crescimento não ter acompanhado as expectativas dos usuários, devido à falta de investimentos das empresas que controlam este mercado. Diante deste panorama, todos aquelas que de alguma forma utilizam este sistema, sempre recomendam aos interessados, que é possível reduzir custos e melhorar a qualidade dos serviços [BAU 2000].

#### **1.4.4 Tecnologias padronizadas e disponíveis**

Em unidades de radiologia nos hospitais, é muito comum encontrar equipamentos de vários fabricantes, gerando uma grande quantidade de diferentes tipos de imagens. A necessidade de um padrão para o manejo e transmissão de imagens digitais é importante, pois torna possível a utilização de sistemas abertos. Podemos considerar que o uso de um padrão internacional iniciou com a integração de um comitê formado por ACR (*American College of Radiology*), representado pela comunidade de radiologia e a NEMA (*National Electrical Manufacturers Association*), em nome da indústria da área de radiologia de acordo com os procedimentos estabelecidos pela NEMA. Os objetivos iniciais eram trabalhar com diferentes problemas de compatibilidade e integração de interfaces com imagens em ambientes proprietários [ARE 2000].

##### **1.4.4.1 Objetivos**

A padronização tem como princípio:

- Promover a comunicação entre imagens digitais de diferentes fabricantes.
- Oferecer maior flexibilidade dos sistemas de armazenamento e comunicação de imagens
- Facilitar a criação de consulta de sistemas de diagnóstico por diferentes dispositivos de diversos locais remotos.

Os primeiros resultados desta padronização foram, em 1985, ACR-NEMA Versão 1.0, tendo como base idéias obtidas de formatos já existentes. Algumas definições de formatos foram desenvolvidas e adotadas para gravar imagens em fita magnética, criada pela *Associação Americana de Físicos em Medicina* (AAPM). Todavia, como todas as versões, detectaram-se vários erros. Assim em 1988 foi criada a versão 2.0 do ACR-NEMA. Atualmente, a última versão homologada é a 3.0 [HUA 98].

##### **1.4.4.2 Padronização**

A mais importante e necessária padronização é o DICOM, por motivo de união entre a ACR e a NEMA, definindo o formato para a imagem, o hardware requisitado, elementos físicos de conexão e a definição do protocolo de comunicação.

#### **1.4.5 Considerações sobre os PACS**

Um PACS é um sistema de informação com suporte a imagens, permitindo a definição e manipulação uniforme da informação, contida em elementos textuais e

gráficos, na forma de imagens, vídeos e som. Para uma caracterização precisa dos requisitos de sua funcionalidade, deve-se considerar as etapas básicas envolvidas em sua construção, seguindo um conjunto de definições esquemáticas bem como: reconhecimento de padrões, recuperação, estruturação e organização das imagens.

## **1.5 Considerações finais**

Neste capítulo, realizou-se um estudo das possíveis formas de utilização de tecnologias para incluir imagens em prontuário do paciente e dos principais conceitos referentes a sua implementação. Fez-se este estudo com o intuito de apresentar diversas maneiras envolvidas com a sua implementação, bem como necessário o desenvolvimento de um planejamento bem organizado, baseado nos padrões que a comunidade médica utiliza. Também foram abordados aspectos nos quais uma ferramenta de controle de PACS poderia contribuir para melhorar e aumentar a qualidade dos serviços de exames de imagem sem aumentar os custos.

## 2 Tecnologias de Desenvolvimento para a Web

O desenvolvimento de soluções para a *Web* utiliza várias tecnologias que interagem entre si, para fazer a aplicação funcionar. Tais tecnologias envolvem protocolos de rede, server-side applications (aplicações servidoras), bancos de dados e programação de interfaces gráficas para os usuários. CGI, PHP, ASP e Servlets são exemplos de tecnologia para o processamento no servidor (*server-side programming*); CORBA, DCOM e Enterprise Java Beans, por sua vez, são exemplos de tecnologias para Objetos Distribuídos; XML, HTML, SGML, XSL, JAVA, CSS e JavaScript são voltadas para a construção da interface com o usuário via o navegador de páginas para a *Web* (*browser*, tais como Internet Explorer e Netscape). Estas e outras tecnologias serão brevemente descritas neste capítulo, tentando-se, dessa forma, dar uma visão geral do processo de desenvolvimento para a *Web*.

### 2.1 Aplicação de Ambiente Web

Essencialmente, uma aplicação *Web* realiza as seguintes tarefas [FRY 2001]:

- disponibiliza uma interface para a entrada de dados;
- transmite os dados informados pelo usuário para o *Web* server;
- recebe os dados enviados utilizando algum conjunto de "middlewares";
- realiza o processamento no servidor (*Server Side Processing*);
- transmite os resultados de volta ao cliente;
- realiza o processamento no cliente dos dados enviados, mostrando-os ao usuário.

Dessa forma, estas aplicações *Web* são um conjunto de tecnologias chamadas de multicamadas. Primariamente, três camadas se destacam, estando sempre presentes em qualquer aplicação *Web* [SAF 2000]:

- camada de apresentação (interface com o usuário);
- camada middleware (objetos e programas server side);
- camada de banco de dados.

A camada de apresentação utiliza, em geral, um *Web browser* para interpretar as páginas HTML oriundas do servidor. A camada middleware, que pode separar camadas de objetos com finalidades específicas, como objetos que tratam das regras de negócio, é a responsável pelo processamento do sistema, recebendo as solicitações do usuário, interagindo com o banco de dados e remetendo as respostas ao usuário na forma de páginas HTML.

A camada de banco de dados é aquela na qual estão armazenadas as informações do sistema.

Além destas, existe ainda a camada de comunicação, fundamental para que a aplicação funcione. Entretanto, esta camada é transparente para o desenvolvedor, visto que *browser* e o *Web server* (servidor) se encarregam de utilizá-la, sem a necessidade da interferência do desenvolvedor. Na camada de comunicação, os protocolos de rede atuam. O protocolo utilizado na *Web* é o TCP/IP. Na FIGURA 2-1 está ilustrado o funcionamento de uma aplicação *Web*.

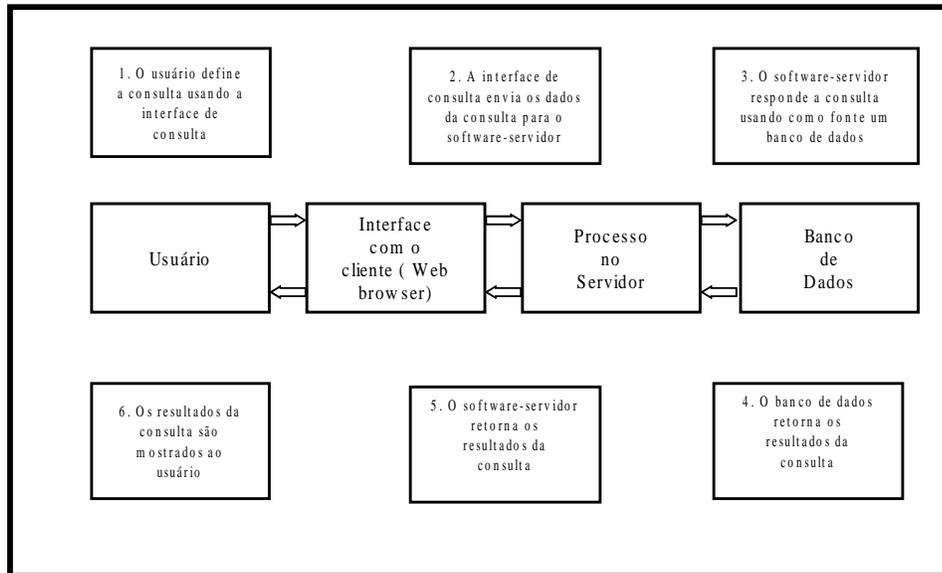


FIGURA 2-1 - Fluxo de uma aplicação na Web.

## 2.2 Tecnologias para a Interface

A aplicação *Web* utiliza-se de uma página em HTML, interpretada pelo *browser*, para interagir com o usuário, formando a Camada de Apresentação. Outras tecnologias podem ser misturadas ao HTML para a construção de uma interface mais poderosa, com um visual mais adequado, além de proporcionar recursos que o HTML isoladamente, não é capaz. A seguir, um breve resumo sobre essas tecnologias, que são responsáveis pela construção da interface com o usuário [FRY 2001]:

### 2.2.1 HTML

O HTML (*HyperText Markup Language*) utiliza os conceitos do HyperTexto e da Hipermídia para apresentar, num mesmo ambiente, dados, imagens e outros tipos de mídia, como vídeos, sons e gráficos. O HTML é um subconjunto do SGML (*Standard Generalized Markup Language*) e utiliza rótulos (tags) que definem a aparência e o formato dos dados, sendo padronizado pelo Object Management Group (OMG). É interpretado por qualquer *browser*, em qualquer plataforma.

### 2.2.2 JAVA

O código Java é compilado e transformado em bytecodes (códigos que são independentes de processador) no qual são interpretados pela JVM (*Máquina Virtual Java*) em tempo de execução. Os programas Java são compilados e executados dentro da JVM podendo ser levadas para a plataforma-alvo, isto significa escrever esta linguagem uma única vez e executá-la em qualquer lugar [MAR 97, CES 96].

### 2.2.3 BROWSERS

São categorias de programas que permitem visualizar um documento criado em um certo padrão, no caso HTML. Atualmente, os *browsers* tem-se tornado complexos, devido à quantidade de padrões existentes (ex. imagens .GIF .JPG, PNG etc). Os *Browsers* são programas que oferecem as ferramentas necessárias para “folhear” as páginas de documentos da *Web*, proporcionando algumas qualidades básicas, como: facilidade de uso, flexibilidade, recursos e outras ferramentas integradas em um único conjunto.

### 2.2.4 CSS

Cascading Style Sheet (CSS) permite que os estilos dos elementos da página (espaçamento, cores, fontes, margens, etc.) sejam especificados separadamente da estrutura do documento, facilitando dessa forma, uma futura modificação no estilo da página.

### 2.2.5 XML

EXtensible Markup Language (XML) é uma linguagem de marcação, tal como o HTML. O XML lida com rótulos (tags) sendo possível definir conjuntos de tags próprios. A definição do padrão de tags possibilita a criação de documentos num formato XML que podem ser facilmente interpretados pelo *Browser*. Diferentemente do HTML, no XML não há tags para a aparência dos dados. O XML é também muito utilizado para padronizar a troca de informações entre sistemas [LIG 99].

XML (Extensible Markup Language) [NAT 2000] é a tecnologia recente, utilizada na *Web*, com tendência a tornar-se padrão com o objetivo de indexar, representar, descrever e divulgar recursos no ambiente WWW. Permite criar tags e elementos, a partir de um subconjunto das linguagens anteriores SGML (Standard Generalized Markup Language) e da HTML (HyperText Markup Language).

Tags XML provêm informação sobre o conteúdo ali contido [CLA 98]. Como vantagens da XML, podem ser citadas as seguintes: baseia-se em padrões internacionais existentes; é totalmente extensível, sem limitações de tags (etiquetas); pode modelar qualquer tipo de dado; suporta controle de validação e editoração; gera automaticamente

ligações e auxílios navegacionais; permite gerar visões configuráveis pelo usuário dinamicamente; apresenta sistema de administração mais simples de *sites Web*; permite fácil redefinição de documentos; apresenta maior velocidade de acesso aos dados.

### **Características da Xml**

XML agrupa os dados entre tags similares aos usados em HTML, com a diferença de que em XML, podem ser criados novos tags, usados para definir os dados nele contidos. Em XML [CLA 98], o termo “element” é usado para descrever os tags (ambos tags de inicialização e finalização). Como, por exemplo, a definição do elemento INFORMAÇÕESGERAIS seria: <INFORMAÇÕESGERAIS> O tópico do documento é XML</INFORMAÇÕESGERAIS>.

Quando servidores *Web* preparam a transmissão de dados com conteúdo XML, tipicamente devem incluir ponteiros para a DTD (Document Type Definition) associada. Clientes *Web* que processam XML devem ser capazes de analisar o contexto de acordo com a DTD e de interpretar a semântica do hipertexto associada a cada um dos diferentes tags do documento [KHA 97]. Embora a utilização de DTD não seja obrigatoriamente utilizada, o XML adapta-se bem ao ambiente da *Web*; ele não precisa conhecer a DTD ou o modelo estrutural que foi usado, para criar um documento XML. Isso não significa que o XML não funcione com DTDs, mas simplesmente que, se um documento não usar uma DTD, ele terá de ser distribuído em sua forma completa [NAT 2000].

A sintaxe adotada neste protótipo foi a de documento bem-formatado, ou seja, permite validar o conjunto de tags para o próprio uso da aplicação. Especifica o conjunto de elementos requeridos e opcionais para os documentos de acordo com o tipo, definindo ainda o nome dos tags e o relacionamento entre elementos em um documento, no qual seguem uma sintaxe bem definida, atendendo às seguintes condições:

- nenhum atributo pode aparecer mais do que uma vez;
- concisão na marcação do XML, sem abreviações;
- devem ser de forma clara, e auto-explicativas.

## **2.3 Tecnologias para o Processamento no Servidor**

Na camada middleware (software intermediário), ocorre realmente o trabalho de programação do aplicativo *Web*, sendo esta camada a responsável por processar a informação enviada pelo cliente (*browser*), processar a regra de negócio (que pode estar em outra camada), interagir com o banco de dados, preparar a resposta (quase sempre na forma de uma página HTML) e enviá-la ao cliente. Os componentes dessa camada estão no *Web Server* e são capazes de utilizar os recursos desses servidores e dos demais recursos conectados para realizar o processamento. É importante perceber que a forma com que todas essas tecnologias trabalham, é similar: recebem uma solicitação do

cliente, processam essa solicitação e respondem na forma de uma página HTML. Existem várias tecnologias para a construção dessa camada. São elas [ZOL 2001]:

### 2.3.1 CGI

Common Gateway Interface (CGI) é um padrão para interfaceamento de aplicações externas com servidores, como um *Web Server*, por exemplo. O CGI é a aplicação mais básica para acessar os recursos do sistema no servidor, e foi também a primeira tecnologia para o desenvolvimento de aplicações *Web*. Pode ser escrito em diversas linguagens, sendo as principais o Perl e o C/C++.

### 2.3.2 ASP

Active Server Pages (ASP) é uma tecnologia da Microsoft, que utiliza os conceitos de SSI e CGI para a construção de conteúdo dinâmico, somente funcionando na *Web Information Server* (IIS), o software servidor *Web* da Microsoft, ou seja, é exclusiva para a plataforma Windows. O código ASP é inserido no HTML e interpretado pelo servidor a cada requisição recebida. O ASP é talvez a mais popular linguagem servidora de script atualmente em uso [PER 2002].

### 2.3.3 PHP

Hypertext Preprocessor (Pré-processador de Hipertexto PHP) é uma linguagem de elaboração de *scripts* embutido, que opera do lado do servidor. Funciona dentro de um documento HTML, com funcionalidades próprias para aplicações de páginas dinâmicas. Segue uma filosofia própria, podendo ser executada por diferentes servidores, principalmente em plataformas consolidadas, como Unix (Solaris, Linux, etc.). PHP utiliza sintaxe baseada em C, Java e Perl. É uma tecnologia não-proprietária [CAS 2001].

### 2.3.4 ISAPI/NSAPI

A tecnologia Information Server Application Programming Interface (ISAPI) é baseada no acesso a Application Programming Interface (API) do *Web Server*, através do qual a aplicação servidora ISAPI ou NSAPI utiliza diretamente a API do *Web Server*, para executar a função desejada. A NSAPI é voltada para o Netscape Server e a ISAPI é a tecnologia para o servidor IIS da Microsoft. Algumas linguagens possibilitam o desenvolvimento de tais aplicativos.

### 2.3.5 Servlets

É um tipo de aplicativo Java que, executado no *Web Server*, permite um funcionamento similar ao CGI. Os Servlets Java são multiplataforma e oferecem bom

desempenho. É um programa do lado do servidor que atende requisições HTTP e retorna resultados como respostas HTTP.

### 2.3.6 JSP

Java Server Pages (JSP) é uma tecnologia baseada em Java que utiliza o mesmo princípio do ASP, com código já mencionado, embutido na página HTML, o qual é interpretado a cada requisição pelo *Web Server*.

### 2.3.7 ColdFusion

Linguagem de script server que também utiliza uma filosofia similar ao ASP e JSP. Possui sintaxe própria e é uma tecnologia proprietária.

### 2.3.8 Banco de Dados

Como banco de dados, as opções são muitas. Existem bancos de dados open-source como o Interbase e *freewares*, como o MySQL, PostgreSQL e comerciais, como Oracle, Informix, DB2 e Sybase, sendo estes mais voltados ao mercado corporativo.

## 2.4 Tecnologia de Protocolo

Com a enorme expansão da *Web* e de seus protocolos, uma mistura entre as tecnologias de objetos distribuídos começa a ocorrer dentro da Rede Mundial de Computadores (WWW). Esta fusão, muitas vezes chamada *Object Web*, possibilita a implementação de arquiteturas que favorecem a interoperação de objetos distribuídos através de protocolos intermediários da *Web*, como o Internet Inter-ORB Protocol (IIOP) [PIS 2001].

Com suas vantagens e limitações destacam-se diferentes níveis de arquiteturas:

HTTP/CGI: é o modelo ainda predominante na *Web* que permite implementar soluções simples de comunicação cliente/servidor;

CORBA: servidores e objetos CORBA possibilitam excelente interoperabilidade entre softwares e seu modelo de distribuição é resultado de um consórcio aberto;

DCOM: representa o cerne da estratégia da Microsoft em estabelecer um modelo de componentes e objetos, totalmente integrados distribuídos na *Web*;

JavaBeans/EJB: comunicação entre servlets e applets representa uma solução multi-plataforma eficiente e de grande abrangência. Atualmente, há

uma grande quantidade de bibliotecas de componentes disponíveis em sua linguagem nativa Java;

Agentes de Software: descrevem a utilização de várias tecnologias para implementar o que há de mais flexível e interoperacional em objetos de sistemas distribuídos.

## **2.5 Tecnologia Utilizada em Sistemas de Imagens Médicas**

A implementação de um sistema distribuído de imagens médicas que satisfaça a contento as características fundamentais descritas por [PIS 2001, PEL 2001], requer a utilização de diferentes tecnologias predominantes na manipulação de imagens médicas e na distribuição de objetos.

### **2.5.1 PACS**

Sistemas de armazenamento e comunicação de imagens – PACS – crescem em popularidade na indústria de tecnologia médica. Por definição, PACS transmite e armazena imagens radiológicas digitais entre equipamentos locais e remotos para interpretação e diagnóstico. PACS, geralmente, são integrados em redes locais (LAN), em sistemas de informações hospitalares (HIS) e sistemas de informações radiológicas (RIS).

Considerando a grande complexidade de desenvolvimento e manutenção, a tendência atual é implementar soluções menos centralizadas, que possibilitem armazenar e processar dados através de uma arquitetura distribuída e que permitam utilizar diferentes equipamentos e ambientes computacionais. Essa arquitetura distribuída não apenas se limita ao armazenamento e transferência de imagens e informação médica de pacientes entre diferentes pontos de uma rede, pois essa funcionalidade também é oferecida pelo PACS tradicional, mas apresenta uma solução baseada em orientação a objetos que deixe de fato implementar interoperabilidade entre aplicações.

### **2.5.2 DICOM**

O padrão DICOM (*Digital Imaging and Communication in Medicine*) é o padrão internacional para a comunicação de imagens biomédicas e de informações relacionadas. O protocolo DICOM é o padrão da indústria para a transferência de imagens nos EUA, Europa e Japão. No Brasil, já podemos notar uma quantidade crescente de equipamentos que suportam DICOM em hospitais e instituições de pesquisa [CLU 2002, HOE 96].

### 2.5.3 HL7

HL7 (*Health Level Seven*) é um padrão sancionado pela American National Standard Institute (ANSI) para comunicação de informação médica, originalmente desenvolvida para troca de dados entre pacientes clínicos. Sua utilização foi estendida para suportar comunicação para a troca de informações e imagens entre sistemas baseados no protocolo DICOM, ou seja, utilizar mensagens HL7, para entregar dados DICOM.

O progresso e popularização da *Web* favoreceram o desenvolvimento de tecnologias para plataformas de computação distribuída. O DICOM e HL7 foram criadas sob várias formas. Um aspecto ainda não completamente resolvido é a conectividade do transporte de informações e arquivamento entre sistemas não-padronizados. Atualmente, a nova forma de utilização desta tecnologia foi completamente solucionada, direcionando-se para a implementação e interoperabilidade entre sistemas médicos.

### 2.5.4 CORBA

As tecnologias ORB oferecem a infra-estrutura de software, necessária para implementar funcionalidade na comunicação de dados e suportar o grande volume das mesmas destes sistemas médicos. Entre as tecnologias ORB destacam-se DCOM da Microsoft Corporation, CORBA da Object Management Group (OMG) e JavaBeans/EJB da Sun Microsystem Corporation.

A arquitetura CORBA é um barramento conceitual de software que permite que aplicações se comuniquem entre si, independentemente do sistema operacional, linguagem de desenvolvimento e fabricante.

O bloco básico de desenvolvimento em CORBA é o objeto CORBA. Um objeto CORBA modela um objeto do mundo real e consiste de dados e métodos que podem ser invocados por este objeto. Uma interface pública para o objeto é definida através de uma linguagem de definição de interface IDL (Interface Definition Language). CORBA IDL é uma linguagem de definição que esconde a implementação subjacente de um objeto. Um cliente do objeto simplesmente utiliza sua interface, para invocar seus métodos – ele não precisa saber sobre sua localização, plataforma ou detalhes de sua implementação. ORB é uma camada intermediária do componente que implemente este barramento de software e age, como se houvesse um intermediário entre o cliente e o objeto.

CORBA foi proposto como padrão pelo OMG, um consórcio com mais de 700 empresas líderes como Microsoft, Nestcape, IBM, Sun, Motorola, Nokia e Boeing. CORBA tem sido utilizado por muitas outras empresas para criar interoperabilidade entre softwares.

### 2.5.5 SGML

O Hypertext Markup Language (HTML) representa um documento *Web*, em ambientes heterogêneos, no qual deu origem ao seu padrão original, o Standard Generalized Markup Language (SGML). SGML é utilizado para criar informação com detalhes em documentos reutilizáveis em muitos domínios técnicos, tais como fichas de pacientes. Com esta tecnologia foi possível então representar o prontuário eletrônico do paciente como um conjunto de documentos baseados em SGML e realizar os padrões de mensagens DICOM e HL7 com sua abordagem de documentos estruturados.

Estudos levam a crer que o sucesso destes padrões: mensagem e informação poderão ser a forma mais própria e original de representar domínios de aplicação e pressionar a indústria, para que esses padrões sejam interoperáveis. A utilização do conceito de meta-registro para pacientes, e conseqüentemente para serviços, é fundamental para a implementação de sistemas distribuídos.

## 2.6 Considerações Finais

Vários autores discutem qual a melhor tecnologia a ser utilizada no desenvolvimento de aplicações *Web*. Alguns defendem o Java (Servlets e JSP) como uma das melhores opções, por ser multi-plataforma e oferecer melhor desempenho. Entretanto, é importante destacar que cada tipo de projeto tem suas necessidades específicas. Em geral, projetos pequenos não exigem soluções complexas e de alto custo; assim, opções mais simples, como ASP ou PHP com banco de dados MySQL, PostgreSQL, por exemplo, podem ser utilizadas [ZOL 2001, CAS 2001]. Por outro lado, para projetos complexos, com grande volume de dados e número de usuários, o uso de tecnologias mais robustas, como Java e Oracle, deve ser considerado.

Para o desenvolvimento de PEPs, a literatura tem mostrado um uso crescente das tecnologias Java, XML, ASP e CORBA [ELB 2000, NAR 2001, GRI 2000], talvez pela complexidade e necessidade de desempenho deste tipo de sistema. Dessa forma, no momento de decidir quais as tecnologias a serem utilizadas no desenvolvimento de um aplicativo *Web*, parâmetros como custo, habilidade da equipe e curva de aprendizado, necessidades de escalabilidade, desempenho, multiplataforma e segurança devem ser considerados, além de adesão a padrões e tecnologias abertas.

### 3 Padrões e Arquiteturas de Metadados

Segundo [KER 97], as pesquisas relacionadas a metadados devem agora ir em direção à integração e interoperação de proposições e padrões existentes. As pesquisas, nesta direção, incluem questões de projeto e desenvolvimento de: modelos de metadados extensíveis que possam ser adaptados a domínios de aplicação específicos; ferramentas permitindo integração e interoperação de conjunto de metadados oriundos de diferentes fontes e representados por padrões diferentes; gerenciadores de metadados extensíveis, oferecendo serviços básicos de suporte a funções essenciais, proporcionando o acesso, transferência, descoberta e análise para o desenvolvimento de aplicações específicas.

Estes direcionamentos requerem um trabalho importante de modelagem de metadados e, mais especificamente, de níveis de modelagem e metamodelos correspondentes. O metamodelo para metadados significa os conceitos usados para definir o modelo de metadados.

Para atender a estes requisitos de integração e interoperação, a literatura de metadados utiliza os termos padrão, modelo e arquitetura. Esse trabalho não leva em consideração uma definição nítida entre os termos padrão e modelo, uma vez que na literatura esses conceitos ainda se confundem. Modelos de metadados são requeridos para descreverem o conteúdo e a semântica dos dados, em qualquer domínio de aplicação. Basicamente, todos os padrões de metadados são modelos que, a partir de um consenso das comunidades de pesquisa passaram a ser utilizados como padrões. Geralmente, esses padrões contêm elementos descritores de dados específicos que são dependentes das finalidades distintas de informações que armazenam. O papel das arquiteturas de metadados é descrever o ambiente de forma a prover a interoperabilidade entre os vários padrões de metadados existentes. Metadados também denominados descritores de dados podem ser definidos em nível de pixel ou a nível semântico.

Neste capítulo, será feito um estudo de padrões de metadados mais relacionados à imagem, que permitem, de alguma forma, descreve-la. Serão apresentadas também as arquiteturas de metadados existentes e uma arquitetura de modelagem.

#### 3.1 Padrões de Metadados

Esta seção apresenta o estudo de dois padrões de metadados, que possibilitam a descrição de imagens, porém não são totalmente adequados à descrição eficiente das mesmas, pois não viabilizam descritores para todas as suas características importantes.

##### 3.1.1 Dublin Core

O Dublin Core Metadata Element Set, ou simplesmente Dublin Core (DC) [BRO97], foi definido inicialmente como um conjunto de treze elementos básicos propostos no primeiro workshop organizado em 1995 pelo OCLC (*Online Computer Library Center*) [OCL 2002, HEE 97] /NCSA (*National Center for Supercomputer*

*Applications*) [WEI95], com o objetivo de desenvolver um registro de metadados para descrever informação eletrônica na rede. Ficou definido que documento DLOs (*Document like objects*) seria o elemento principal, para descrever recursos de informação na rede. Apesar de um documento do ambiente *Web* poder ser composto de textos, imagens, áudio, vídeo ou até outro hiperdocumento, ele deveria ser considerado como um objeto global. O principal objetivo desse padrão é identificar e definir um conjunto de elementos de metadados para determinar os recursos na *Web*.

No workshop de metadados para imagem CNI (*Coalition for Networked Information*)/OCLC [MIL96], acontecido em setembro de 1996, os elementos do conjunto de dados iniciais do DC foram modificados a fim de incluir requisitos que servem para descrever tanto texto como imagem tais como um campo para descrição de conteúdo (DC.description) e um campo para descrição de direitos autorais (DC.rights), ficando o Dublin Core com 15 elementos descritores.

O workshop chamado “W3C Workshop on Distributed Indexing and Searching” [WEI96], que aconteceu em maio de 1996, resultou em uma proposta para embutir metadados em HTML.

O 4º Workshop Dublin Core, que aconteceu em março de 1997, em Canberra [HEE 1997], continuou discutindo questões relacionadas aos elementos e sintaxe do Dublin Core, e formalizou qualificadores adicionais (“The Canberra Qualifiers”): LANGUAGE (especifica a linguagem para o valor do elemento do campo descritor; geralmente o padrão tem sido Inglês), SCHEME (especifica um contexto para a interpretação de um determinado elemento), TYPE (especifica o tipo de um determinado campo e é visto como um nome de sub-elemento).

Exemplo de mapeamento do Dublin Core em HTML:

```
<HEAD>
<TITLE>Metadata: A Imagem ....</TITLE>
<META NAME="DC.title"CONTENT="(TYPE=short) Metadata: A imagem...">
<LINK REL=SCHEMA.dc HREF="http://purl.org/metadata/dublin_core_elements #title">
<META NAME="DC.subject" CONTENT="(TYPE=keyword) metadata, descrição de recurso, ....">
<LINK REL=SCHEMA.dc HREF="http://purl.org/metadata/dublin_core_elements #subject">
  <META NAME="DC.autor"CONTENT="(TYPE=name) Guimaraães, R.R">
  <LINK REL=SCHEMA.dc HREF="http://purl.org/metadata/dublin_core_elements#author">
<META NAME="DC.date"CONTENT="(TYPE=creation)(SCHEME=ISO31 2001-12-29">
  <LINK REL=SCHEMA.dc HREF="http://purl.org/metadata/dublin_core_elements#date">
  <META NAME="DC.autor"CONTENT="(TYPE=email) renato@polairs.ucpel.tche.br">
  <LINK REL=SCHEMA.dc HREF="http://purl.org/metadata/dublin_core_elements#author">
</HEAD>
<BODY>
```

Na TABELA 3.1 a seguir estão apresentados os quinze elementos do Dublin Core [BRO97]:

TABELA 3.1 - Elementos do Padrão Dublin Core.

DC.subject	Sobre o que trata o trabalho; definido com palavra-chaves ou informações tópicas. Assunto: tópico, gênero, ou frases descritivas.
DC.description	Texto descrevendo o conteúdo do recurso.
DC.title	O título que um autor ou editor atribui a um trabalho. Título.
DC.author	A pessoa ou pessoas responsáveis pelo conteúdo do trabalho. Principal criador.
DC.publisher	Um agente ou agentes responsáveis pela disponibilização do trabalho em sua forma atual; normalmente uma editora ou instituição educacional. Agência responsável pela disseminação (por exemplo: instituição hospedeira).
DC.contributors	Outras pessoas responsáveis por aspectos da página.
DC.date	Quando o trabalho se tornou disponível em sua forma descrita. Data da última atualização: YYYYMMDD
DC.type	Tipo de recurso: documento.
DC.format	Formato: texto/html.
DC.identifier	URL ou outros identificadores. Pode incluir ISBN.
DC.source	Fonte original, se originalmente publicado em formato de paper.
DC.language	Código 3-letras para linguagem da página.
DC.relation	(Sob desenvolvimento; pretende ligar por exemplo, gráfico a página principal)
DC.coverage	A localização espacial e/ou características de duração temporais do objeto. (em desenvolvimento)
DC.rights	nota de direitos autorais.

### 3.1.2 SAIF

O padrão SAIF (*Spatial Archive and Interchange Format*) foi desenvolvido como um meio de compartilhar dados espaciais e espaço-temporais, permitindo descrever imagens relacionadas a dados geo-espaciais. SAIF foi aprovado, em 1993, pelo CGSB (Canadian General Standards Board) como um padrão nacional no Canadá. Os principais objetivos do SAIF são [SAI 2001]:

- manipular virtualmente qualquer tipo de dado geográfico;
- endereçar tempo de forma que eventos e relacionamentos temporais possam ser manipulados;
- Direcionar requisitos de gerenciamento de dados (tais como suporte a atualizações, habilidade para integrar dados multimídia, habilidade para interfacear consultas a bancos de dados);
- ser apropriado para realizar operações eficazes em ambientes de rede;
- harmonizar-se com os novos desenvolvimentos de bancos de dados SQL e iniciativas do “Open GIS” (Padrão para Sistemas de Informação Geográficos).

SAIF é capaz de tratar dados geográficos como qualquer outro tipo de dado simples. Introduce um número de construções espaciais e temporais. A informação de metamodelo do SAIF enfatiza a estrutura de dado de objetos, expressa através de um paradigma orientado a objeto.

## 3.2 Arquiteturas de Metadados

Nesta seção serão apresentadas algumas arquiteturas de metadados encontradas na literatura que fazem parte do conjunto das mais importantes e conhecidas. As arquiteturas de metadados descrevem o ambiente de forma a prover a interoperabilidade entre os vários padrões de metadados existentes.

Para existir uma arquitetura de metadados, é necessário um conjunto para agregar suporte à codificação e ao transporte dos dados de forma independente, provendo interoperabilidade do sistema, através de regras comuns que tratem a semântica, sintaxe e estrutura do metadado. Tais regras são:

**Semântica:** Uma arquitetura é capaz de promover interoperabilidade semântica, quando possibilita a compreensão de seus elementos oriundos de padrões distintos. O componente que armazena a semântica de metadados é denominado repositório.

**Estrutural:** Uma arquitetura é capaz de prover interoperabilidade estrutural quando provê representação para modelo de dados distintos (grafos, árvores, registros, pares de nome/valor de um atributo etc).

**Sintaxe:** Uma arquitetura é capaz de facilitar interoperabilidade de sintaxe quando provê um modo padrão para a representação e intercâmbio de metadados (ex: XML)

### 3.2.1 Arquitetura Warwick

O grupo de trabalho Dublin Core criou uma infra-estrutura chamada de Arquitetura Warwick (Warwick Framework) [DEM 96, LAG 97], que tem o objetivo de agregar pacotes de metadados distintos. A arquitetura suporta administração autônoma e acessa pacotes de metadados de padrões diferentes (por exemplo: registro baseado em Dublin Core pode ser um pacote, e SAIF pode ser outro).

Esta arquitetura deve ser modular, para que novos metadados sejam criados, referenciados e associados a outros.

Os componentes básicos da arquitetura são:

- Recipiente: unidade básica onde conjuntos de metadados tipados (os pacotes) são agregados.

- Pacote: conjunto de metadados tipados representando um objeto de primeira classe. O conteúdo do pacote é simplesmente considerado uma seqüência de bits. Eles podem ser divididos em 3 categorias:

Primitivo: formato de metadado primitivo definido separadamente. Podem ser registros Marc, DC e outros;

Indireto: referência a um objeto externo. A marca da referência é um objeto de primeira classe, e poder ter seu próprio metadado e condições para acesso.

Recipiente: coleção de objetos metadados, os quais podem ser pacotes ou outros recipientes.

Abaixo, apresenta-se um exemplo da implementação da Arquitetura Warwick em HTML. Esta implementação propôs as marcas META e LINK do HTML 2.0 com o objetivo de ser usada para embutir metadados e referenciar uma autoridade de esquema (DC, por exemplo), descrevendo um documento:

```
<HEAD>
<TITLE>Metadata Support to Internet Resources Description... </TITLE>
<META NAME="package"    CONTENT="(TYPE=begin) Dublin Core">
<META NAME="DC.title"   CONTENT="(TYPE=short) Metadata Support....">
<LINK REL=SCHEMA.dc HREF="http://purl.org/metadata/dublin_core_elements #title">
<META NAME="DC.autor"   CONTENT="(TYPE=name) Guimaraes, R.R">
<LINK REL=SCHEMA.dc HREF="http://purl.org/metadata/dublin_core_elements#autor">
<META NAME="package"   CONTENT="(TYPE=end) Dublin Core">
</HEAD>
```

### 3.2.2 Arquitetura de Descrição de Conteúdo (MCF)

A MCF (Meta Content Framework), arquitetura de descrição de conteúdo, constitui-se basicamente de uma linguagem de descrição de estrutura baseada em um modelo de dados extensível apresentado pelo Netscape para o consórcio World Wide Web (W3C) [GUH 97].

Por meio da MCF, cada aplicação pode adicionar e usar seus próprios tipos de metadados. O modelo de dados MCF inclui um conjunto de tipos básicos que podem ser estendidos para acomodar novos tipos de dados. MCF determina também um vocabulário básico que inclui um conjunto de termos comuns, usados para descrever o conteúdo de documentos *Web*.

Um banco de dados MCF é um conjunto de DLGs (*Directed Linked Graph*), compreendendo: um conjunto de etiquetas (referidas como propriedades); um conjunto de nós; um conjunto de arcos, onde cada arco é uma tripla, consistindo de dois nós (origem e destino) e uma etiqueta.

Na MCF os nós podem representar coisas como páginas *Web*, imagens, categorias de assuntos, e *sites*, como também objetos do mundo real, tais como pessoas, lugares e

eventos. Etiquetas são as que correspondem a propriedades tais como tamanho ou data da última revisão, para descrever páginas *Web* e relações (tal como hiperlinks). Cada propriedade/etiqueta representa um nó, porém nem todos os nós são propriedades.

A FIGURA 3-1 apresenta um quadro DLG, representando dados MCF. Domain é um tipo de propriedade, e cada ocorrência desse como etiqueta indicada pela seta é uma propriedade. As setas são somente ponteiros de algum tipo, e os objetos são estruturas na memória ou objetos em uma tabela de banco de dados.

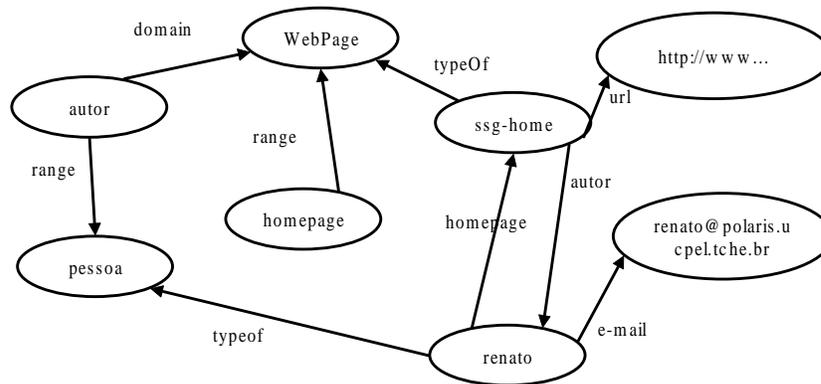


FIGURA 3-1 - Quadro DLG representando dados MCF.

MCF é representada através de uma sintaxe baseada na XML (Extensible Markup Language), a qual conserva as vantagens de extensibilidade da SGML (Standard Generalized Markup Language), considerada como uma linguagem de representação de dados de propósito geral extensível. Os hiperlinks XML são usados para referenciar os blocos MCF, armazenados externamente. Esses são importantes pois favorecem o compartilhamento e reuso de descrições, evitando duplicação.

### 3.2.3 Arquitetura de Descrição de Recurso (RDF)

A RDF (Resource Description Framework), arquitetura de descrição de recurso, especificada pela W3C (World Wide Web Consortium), provê interoperabilidade entre aplicações que trocam informações através da *Web* [LAS 97]. RDF enfatiza facilidades para disponibilizar processamento automático de recursos na *Web*, com o objetivo de definir um mecanismo que descreva recursos de informação sobre qualquer domínio.

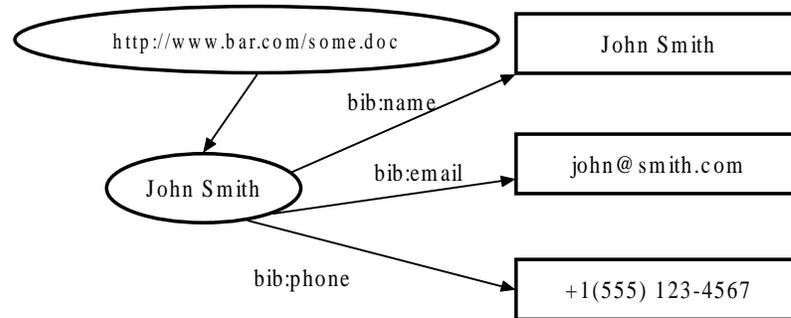
Metadado no contexto da RDF significa “dado descrevendo recursos *Web*”. RDF utiliza XML (Extensible Markup Language) como sintaxe de codificação para os metadados. Os recursos descritos pela RDF podem, geralmente, ser nomeados por um URI (Universal Resource Identifier).

Modelos RDF são grafos direcionados rotulados, onde os nós representam recursos *Web*, e arcos estabelecem propriedades (“Autor”, por exemplo) associadas a um nó. Estas propriedades servem para representar atributos dos recursos e relacionamentos entre eles. Tanto os recursos descritos, quanto os valores que os

descrevem, que são nós. Os arcos, conectando pares de nós, correspondem aos nomes dos tipos de propriedades.

A FIGURA 3-2 mostra exemplos usando RDF [LAS 97], onde os círculos são recursos que estão sendo descritos e os retângulos, os respectivos valores dependendo da propriedade indicada na linha de ligação.

1)



2)

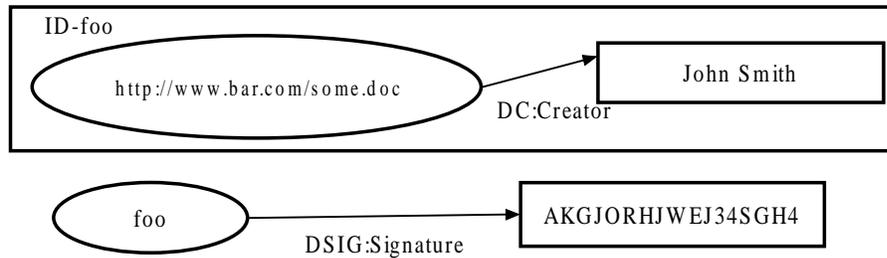


FIGURA 3-2 - Exemplos RDF.

E os exemplos da FIGURA 3-2 podem ser substituídos pela representação de serialização XML:

1)

```
<?namespace href="http://docs.r.us.com/bibliography-info" as="bib"?>
```

```
<?namespace href="http://www.w3.org/shemas/rdf-shemas" as="RDF"?>
```

```
<RDF:serialization>
```

```
<RDF:assertions href="http://www.bar.com/some.dco">
```

```
<bib:author>
```

```

<RDF:resource>
<bib:name>John Smith</bib:name>
<bib:email>john@smith.com</bib:email>
<bib:phone>+1(555) 123-4567</bib:phone>
</RDF:resource>
</bib:author>
</RDF:assertions>
<RDF:serialization>

```

2)

```

<?namespace href="http://purl.org/DublinCore/RDFschema" as="DC"?>
<?namespace href="http://www.w3.org/shemas/rdf-shemas" as="RDF"?>
<?namespace href="http://www.w3.org/shemas/DSig-shemas" as="DSIG"?>
<RDF:serialization>
<RDF:assertions href="http://www.bar.com/some.doc" id=foo">
<DC:Creator>John Smith</ DC:Creator >
</RDF:assertions>
<RDF:assertions href="#foo">
<DSIG:Signature> AKGJORHJWEJ34SGH4</DSIG:Signature>
</RDF:assertions>
</RDF:serialization>

```

### 3.2.4 Arquitetura Segundo Chang e Zhang

Na arquitetura de metadados proposta por [CHA 1997] para tratar metadados em bancos de dados visuais, apresentada na FIGURA 3-3, o metaservidor consiste de quatro componentes: coletor de metadados, meta-banco de dados, construtor de template e módulo de refinamento de metadados. Esta é uma arquitetura diferente das apresentadas anteriormente, pois é operacional, se preocupa apenas em descrever as funcionalidades existentes, de forma a permitir trabalhar com os metadados.

O papel do metaservidor é aceitar consultas de usuário, transformar a informação da consulta para um índice de metadado adequado, distribuir as consultas para os bancos de dados selecionados e retransmitir suas respostas de volta aos usuários.

Imagens no banco de dados visual podem ser categorizadas com base em similaridade com relação a um conjunto de imagens, gerando templates.

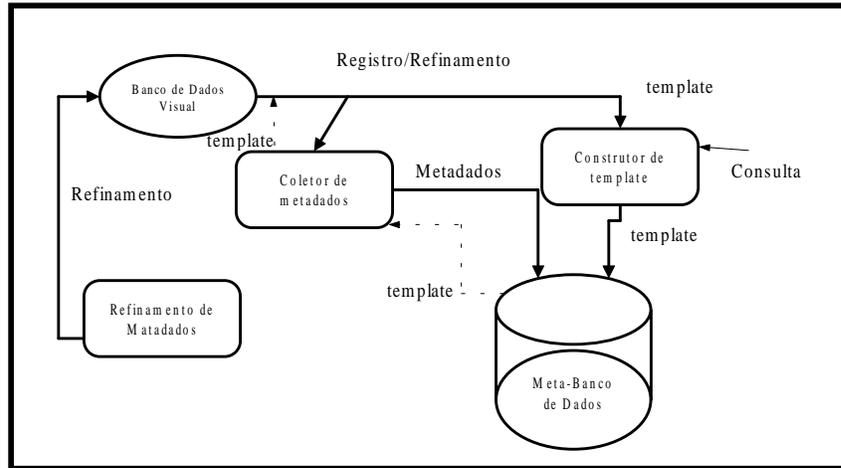


FIGURA 3-3 - Arquitetura de Metadados

O coletor de metadados coleta os metadados do banco de dados de imagem. O banco de dados então computa as similaridades de suas imagens com todos os templates. O coletor de metadados armazena os metadados no meta-banco de dados que registrados no metaservidor, ao serem atualizados permitem que os seus metadados sejam também dinamicamente atualizados. O módulo de refinamento inicia periodicamente as atualizações dos metadados pedindo aos bancos de dados para os resubmeterem ao coletor.

Ao receber uma consulta o metaservidor primeiro extrai várias classes de características desta que, então, é decomposta em subconsultas, cada uma representando uma classe de características a ser recuperada do banco de dados de imagens. Depois o mecanismo de busca do metaservidor casa cada subconsulta com os templates mais relacionados. As similaridades entre a subconsulta e o template são então indexadas com a informação estatística correspondente aos bancos de dados, para derivar a lista dos bancos de dados relacionados à subconsulta. E, finalmente, as listas dos bancos de dados para as subconsultas são agrupadas, com o fim de produzirem o conjunto final de bancos de dados para a dada consulta.

### 3.3 Arquitetura de Modelagem em 4 Camadas Segundo Kerhervé

Nas seções anteriores foram mostradas arquiteturas para tratamento de metadados, tendo cada uma sem enfoque específico para o tratamento da informação. Nesta seção, será apresentada uma arquitetura mais genérica, proposta por [KER 97], apresentando um formalismo que permite uma representação homogênea dos diferentes níveis de modelagem.

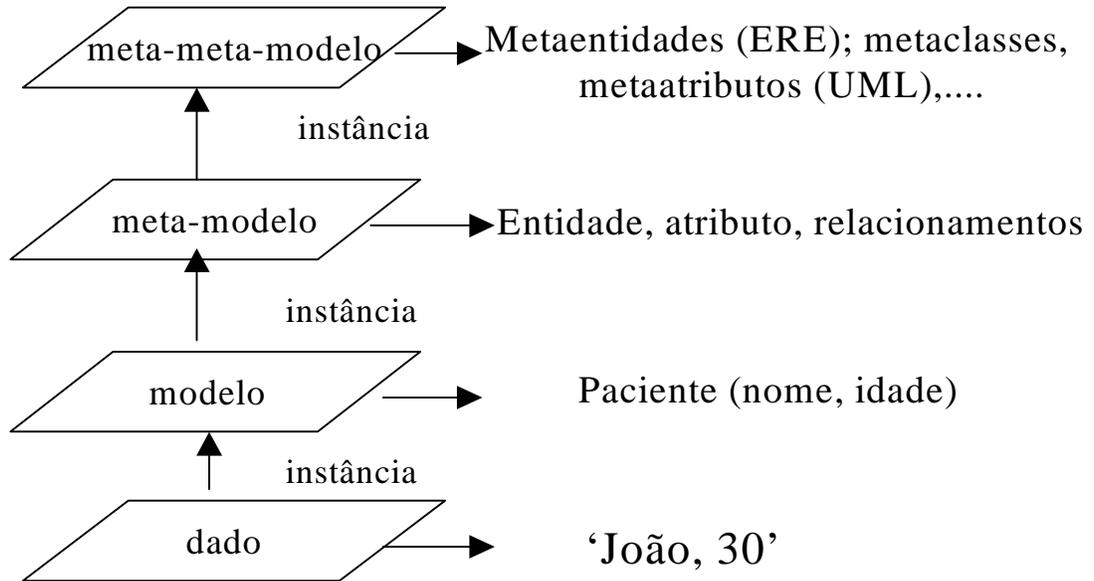


FIGURA 3-4 - Arquitetura de modelagem segundo Kerhervé em 4 camadas

A FIGURA 3-4 apresenta uma arquitetura de modelagem de quatro níveis que permite a representação de metadados a partir da estrutura dos elementos de dados. Os diferentes níveis de modelagem favorecem a incorporação de modelos e metamodelos para metadados.

*Nível de dado* corresponde às instâncias que são manipuladas no sistema. O tipo de metadado que pode ser extraído refere-se aos aspectos físicos e de conteúdo dos dados.

*Nível de descrição do modelo de dados* é o conceito usado para descrever dados e metadados. Nesse nível pode ser encontrado também o modelo para metadados, que é o esquema de metadados ou estrutura dos elementos de metadados.

*Nível de metamodelo* define o formalismo do modelo que é usado no sistema. Descreve os conceitos usados para representar a informação nos níveis mais baixos. Tais metadados são usados para interoperabilidade de ferramentas, métodos ou sistemas.

*Nível de meta-meta modelo* é a raiz do nível de modelagem, permitindo uma representação homogênea dos outros níveis e uma interação mais fácil entre os modelos dos níveis mais baixos.

Nesta arquitetura foram escolhidos grafos conceituais pelo poder de expressão e a possibilidade de manipular tipos e instâncias. Grafo conceitual é um formalismo pelo qual um universo de objetos podem ser modelados por conceitos e relações conceituais. Um conceito representa um objeto de interesse ou conhecimento e uma relação conceitual torna possível associar esses conceitos.

### 3.4 Considerações finais

O ambiente completamente distribuído, provido pela *Web*, requer a interoperabilidade entre padrões e implementações existentes, requisitos importantes para a extensibilidade e adaptação de modelos e metamodelos.

Os padrões e arquiteturas de metadados estudados devem ser considerados quando se pensa na utilização de metadados. Dentre os inúmeros padrões estudados, apenas o DC foi considerado, por dar alguma conotação a dado de imagem. Porém, ele não é totalmente adequado à descrição mais eficiente de imagens, pois não viabiliza descritores para todas as suas características importantes.

Dentro das arquiteturas de metadados estudadas, pode ser visto que todas se preocupam com a extensibilidade de seus modelos e com a possibilidade de utilização em domínios diferentes. A arquitetura Warwick utiliza o conceito de pacote como uma forma de extensibilidade, aceitando, com isso, metadados de padrões variados.

As arquiteturas MCF e a RDF usam uma sintaxe baseada em XML e possibilitam a utilização de metadados em geral, visando a utilização na *Web*, enquanto que a arquitetura, segundo [CHA 97], permite um tratamento especial de metadados para bancos de dados visuais, dando ênfase à parte operacional do tratamento dos metadados.

Os sistemas estudados que utilizam metadados para imagem, não se preocupam em definir todas as características da imagem. Pensando no uso de um banco de dados de imagem na prática, o que realmente interessa, é a possibilidade de efetuar consultas, baseadas em informações referentes ao conteúdo semântico da imagem e não, somente características como cor, textura e formas predominantes. Existem algumas formas, embora limitadas, de utilizar metadados para imagem, porém não foi encontrada uma especificação sistemática e mais completa dos elementos descritores ou de suas classificações.

O objetivo desse trabalho visa à necessidade de se criar um esquema conceitual de metadados para imagem, considerando seu uso para a descrição de metadados extraídos de imagens DICOM.

O próximo capítulo irá apresentar a especificação de uma proposta baseada no esquema conceitual de metadados para imagem, que possibilitará a descrição e recuperação de imagens DICOM, considerando informações técnicas e informações de conteúdo semântico.

## 4 Proposta de um Modelo para Representar Imagens DICOM

A *Web* é uma rede global de computadores que cobre hoje uma boa parte do mundo habitado. Graças ao impressionante progresso tecnológico, representado pelas telecomunicações e pela Informática, tornou-se um fenômeno de massa ímpar na história da ciência, para levar o conhecimento. Um dos fatos de grande importância nos últimos anos e que permitiu a disseminação da informação de forma abrangente foi à criação da *World Wide Web* [SAB 98a]. Porém, devido à grande quantidade de material por ela divulgado, e pela forma de recuperação da informação, seja através de identificadores de recurso, como por exemplo, o URL (Universal Resource Locators) [WEI 95]. Podemos destacar, ainda, outros fatores que dificultam o processo de pesquisa: os métodos utilizados por estas ferramentas para a criação de seus índices; regras rígidas quanto à formulação da consulta; e ainda a falta de padronização no sentido de prover mais recursos que permitam extrair conteúdo das imagens no prontuário do paciente. Desta forma, as principais características serão abordadas neste capítulo, como os documentos podem ser utilizados e como estes são recuperados em um ambiente tão diverso quanto a *Web*.

### 4.1 Modelo para Geração de Interfaces para Imagens

Muitos esquemas de representação com imagens disponibilizam seu conteúdo, através de ligações (links) para diversas páginas, acarretando, com isso, alguns problemas de manutenção. Neste trabalho será utilizado um modelo, mecanismo de grande utilidade para a determinação das formas de manipulação dos dados de uma aplicação, ou seja, nele são incorporadas construções que possibilitam a apreensão do real significado dos dados.

A utilização de um modelo para o controle de interfaces segue uma tendência verificada em outros trabalhos, tais como o Sistema VisualHarness [SHA 97], onde um banco de dados é usado, para descrever e armazenar dados e metadados para a manipulação da estrutura. Outros já enfatizam o uso de um modelo conceitual para especificar a semântica de aplicações da *Web*, fazendo uso de ferramentas específicas, para reduzir o tempo e o esforço de desenvolvimento, e a geração automática de páginas dinâmicas.

Além desta tendência, é necessário justificar o uso de modelos, por ser este um dos princípios de desenvolvimento listados no trabalho para a construção de *sites* com uso de dados intensivo, onde um *site*, caracterizado desta forma, indica que seu objetivo primário é tornar acessível grande quantidade de dados para uma variedade de usuários, o mesmo ocorrendo com um repositório de imagens digitais de pacientes. A seguir alguns destes princípios e a sua aplicação no presente trabalho:

1. todo *site* com uso intensivo de dados deve utilizar um esquema conceitual - como o modelo conceitual provê um melhor entendimento do conteúdo da informação, tal entendimento pode facilitar o processo de geração, uso, manutenção, interoperabilidade, pesquisas e evolução do *site*;

2. deve suportar derivação de dados - que significa o processo de extração dos dados de várias fontes e a produção de dados específicos para as páginas de um *site*.
3. a navegação de dados deve ser suportada – em um nível conceitual, a navegação pode ser modelada por meio de construção de uma metodologia de informação de um tipo determinado;
4. uma composição flexível de páginas deve ser suportada – uma aplicação deve ter várias opções de escolha, permitindo aos usuários o acesso ao conteúdo da aplicação de diferentes modos;
5. estilos de apresentação flexíveis devem ser suportados – um método que permite a definição dos estilos de uma página é considerar a tela como um elemento abstrato, a partir do qual deve possibilitar que outros elementos sejam adicionados;
6. os metadados devem ser gerenciados por um mecanismo – a utilização de uma gerência para o controle dos metadados, facilitando a integração desta meta-informação com o conteúdo do dado;
7. a materialização de páginas e de dados deve ser possível, com o uso de um repositório de dados é possível a geração de páginas instantâneas, sendo que isto é executado no momento da requisição do usuário ou dividido em etapas, onde algumas delas podem ser antecipadas por motivo de performance;
8. paradigmas de pesquisas associativas devem ser suportados, de modo a permitir que o usuário pesquise o *site* sem o conhecimento de sua organização física. Linguagens de pesquisa devem operar em um nível conceitual, através de interfaces gráficas baseadas no esquema conceitual, solicitando destes a definição dos predicados e a definição da navegação;

Destes princípios, os cinco (5) primeiros representam as características metodológicas que devem ser consideradas para um *site*. As regras 6 e 7, dizem respeito à implementação e, a 8 descreve funcionalidades especiais, necessárias a aplicações deste tipo.

## 4.2 Metadados

Metadados são normalmente conhecidos como “informação que descreve informação”. Em um ambiente como o da *Web*, isto pode ser entendido como os dados que podem auxiliar na organização, identificação, segurança, descrição, localização, recuperação, controle de acesso, proteção dos direitos autorais e de propriedade etc., de documentos eletrônicos e não-eletrônicos. O catálogo de uma biblioteca de imagens médicas é um exemplo típico de metadados, pois nele existem várias informações que descrevem os recursos armazenados, auxiliando os usuários na recuperação. Assim, podemos dizer que metadados têm importantes funções e que não estão somente associadas à localização do recurso, podendo abranger outras propriedades tais como: semântica, conteúdo, verificação, controle, gerenciamento do recurso e também a de

padronização de um formato para a utilização entre instituições [MIL 99]. No caso das imagens digitais, sua importância está ligada, principalmente, à descrição das imagens digitalizadas de pacientes e localização dos recursos. Fato importante a ser observado no estudo de metadados é que estes são, em geral, extremamente específicos para a utilização de determinados grupos de trabalho, existindo assim, metadados para descrição de conteúdos, localização de recursos, descrição de dados ambientais etc. Muitos projetos envolvendo metadados têm sido desenvolvidos nos últimos anos [W3C 2000]. A seguir, são apresentados os principais padrões de metadados relacionados a imagens digitais.

### 4.3 Padrão Normalizado

Tradicionalmente, a interoperabilidade de informações heterogêneas e distribuídas de um Sistema de Saúde é obtida através da especificação de estruturas padrões e do mapeamento destas informações para estas estruturas. Esta solução, entretanto, enfrenta dificuldades associadas à complexidade do mapeamento, que está focado nos aspectos estruturais das informações. Essa dependência estrutural também dificulta o desenvolvimento de agentes computacionais para os mais variados propósitos (como agentes de busca), pois estes necessitam de informações que muitas vezes foram mapeadas para estruturas que foram padronizadas para outros objetivos.

A idéia de usar um formato universal para coleta de dados e sua representação não é nova. De fato, há muito que se tenta compatibilizar software entre diferentes computadores, arranjando padrões de troca de dados. No domínio das imagens geradas por equipamentos médicos, é necessário assumir um formato de troca universal para os dados extraídos destas imagens. Num sentido mais amplo, o SGML (Standard Generalized Markup Language) aborda uma tentativa de uniformização.

Por outro lado, o HTML é uma forte ferramenta no domínio da Internet. Aliás, os *browsers* apenas lêem documentos formatados de acordo com esta linguagem. No entanto, o HTML só nos permite representar documentos, não sendo possível, por exemplo, identificar dados de um autor fora do contexto do documento. Do ponto de vista documental, também não se pode identificar isoladamente um autor, porque o HTML não possui esta facilidade de descrever a informação. Daí a importância de criar uma estrutura hierarquizada podendo, por exemplo, identificar o nome do paciente de uma determinada imagem.

Dentro deste contexto, o recurso escolhido para representar os dados textuais das imagens, o foi escolhido o RDF [LAS 97] para descrever dados, juntamente com um vocabulário específico para esta finalidade. A partir deste recurso foi possível descrever uma sentença RDF, que é formada por uma tripla que contém a identificação do recurso (sujeito), a propriedade nomeada (predicado) e um valor (objeto) descrito no item 3.2.3. No contexto das imagens digitais, a estrutura que se está a impor como padrão para a identificação e descrição dos documentos é o Dublin Core, que ainda necessita de alguns qualificadores para melhor definir a função de alguns metadados.

#### 4.4 Dublin Core

Com o objetivo de criar um formato para solucionar o problema da localização de recurso, a Online Computer Library Center (OCLC) (OCLC, 2000) e a National Center for Supercomputing Applications (NCSA, 2000) criaram em 1995 o primeiro Metadata Workshop (DC1, 1995), em 1995 em Dublin, Ohio. Um conjunto de 13 descritores surgiu como resultado deste encontro, formando o padrão de metadados Dublin Core [DUB 2000]. A idéia principal era a de identificar a semântica de um conjunto de descritores que tornasse o processo de localização do recurso na *Web* mais fácil, onde os documentos pudessem ser pensados como objetos, permitindo assim, que autores e provedores de informação pudessem descrever seus recursos por conta própria, sem a necessidade de treinamentos para a criação de registros conforme os padrões pré-estabelecidos. Mais tarde, em janeiro de 1997 (DC3), mais dois (2) elementos voltados para imagens foram adicionados a esse padrão, constituindo o conjunto de descritores apresentados na TABELA 4.1.

Em outubro de 1999, foi realizado o 7º encontro (DC7), com o objetivo de consolidar grupos de trabalho e experiências com interoperabilidade entre sistemas, utilizando metadados heterogêneos. O atual formato é denominado de RFC 2413 [RFC 2413, 1998].

TABELA 4.1 - Descritores do padrão Dublin Core

<b>Escopo</b>	<b>Descritor</b>	<b>Significado</b>
<b>Conteúdo</b>	Assunto e palavras-chave	Tópico relacionado ao conteúdo do recurso. Indicado por palavras-chave e por códigos de esquemas de classificação
	Descrição	Contém uma descrição textual do objeto ou uma referência para essa descrição, tal como um abstract ou resumo
	Língua	Idioma utilizado para expressar o conteúdo intelectual do objeto
	Relação	Indica o relacionamento lógico do objeto com outros recursos
	Fonte	Referência a partir do qual o presente recurso foi derivado
	Título	Nome dado ao recurso ou forma como é conhecido

	Cobertura	Extensão ou escopo do conteúdo intelectual do recurso, tais como: cobertura espacial, período ou jurisdição
Propriedade intelectual	Contribuidor	Entidade responsável por realizar contribuições significativas para o conteúdo intelectual do documento
	Publicador	Agente ou a agência responsável por tornar o objeto disponível, por exemplo, o nome da pessoa, organização ou serviço
	Direitos de	
	Contém ou referencia o direito de propriedade sobre o objeto	
	Criador	Nome da entidade responsável pelo conteúdo intelectual do recurso, por exemplo, autor, organização ou serviço
Instânciação	Data	Data associada a um evento no ciclo de vida do documento, por exemplo, sua criação ou disponibilização. Utiliza-se o formato da ISO 8601, ou seja, ano-mês-dia
	Formato	Formato digital do objeto, utilizado exemplo a partir da lista de tipos MIME, em desenvolvimento
	Identificador	Números ou caracteres usados para identificar o recurso de forma única a exemplo da URL ou ISBN
	Tipo do recurso	Natureza ou gênero do conteúdo do recurso, tal como uma novela, poema ou dicionário

#### 4.4.1 Análise do Padrão

Este padrão tem, como objetivo principal, ser de extrema simplicidade, facilitando a aplicação por parte do usuário na descrição de recursos na *Web*. Quando da aplicação do padrão em imagens digitais, verifica-se que os descritores são muito genéricos, como o descritor tipo de recurso, ou muito restritivos, a exemplo da concatenação de palavras-chave e assunto, por serem itens bastante distintos. Em alguns atributos encontram-se ainda sob definição, dentre eles destaca-se o Tipo do Recurso [DCT 99], possivelmente pelo fato da dificuldade de se criar uma categoria que descreva tipos tão diversos de

recursos. Assim, pôde verificar-se que a aplicação direta do padrão para a descrição de documentos eletrônicos em uma imagem digital é insuficiente, pois neste tipo de aplicação o foco não está na descoberta do recurso, mas sim na sua descrição e, como foi visto anteriormente, a aplicação do padrão para nesse caso, não é suficiente para representar o conteúdo em poucos descritores.

#### 4.4.2 Análise do Formato

Este formato, definido de acordo com as necessidades de representar imagens digitais estáticas neste trabalho, contou com o auxílio do estudo da área de informação que vem sendo utilizada por importantes instituições de saúde. Observou-se ainda que os atributos criados são de extrema utilidade na descrição de imagens digitais. Espera-se, portanto, que este formato seja validado como um padrão para a sua utilização, tendo em vista que não foi imposto por uma instituição, mas sim, fruto das necessidades de representar o conteúdo implícito agregado na imagem.

#### 4.4.3 Formato de Metadado da Ferramenta Proposta

Após o estudo dos padrões de metadados mais utilizados em imagens digitais, verificou-se que para o trabalho proposto, é possível realizar uma composição destes metadados e adicionar outros específicos para as particularidades de descrição de conteúdo de imagens. Para que esta composição seja mais bem compreendida, foram criadas três (3) categorias de metadados necessárias a um estudo completo de documentos eletrônicos neste ambiente, exibidas na FIGURA 4-1.

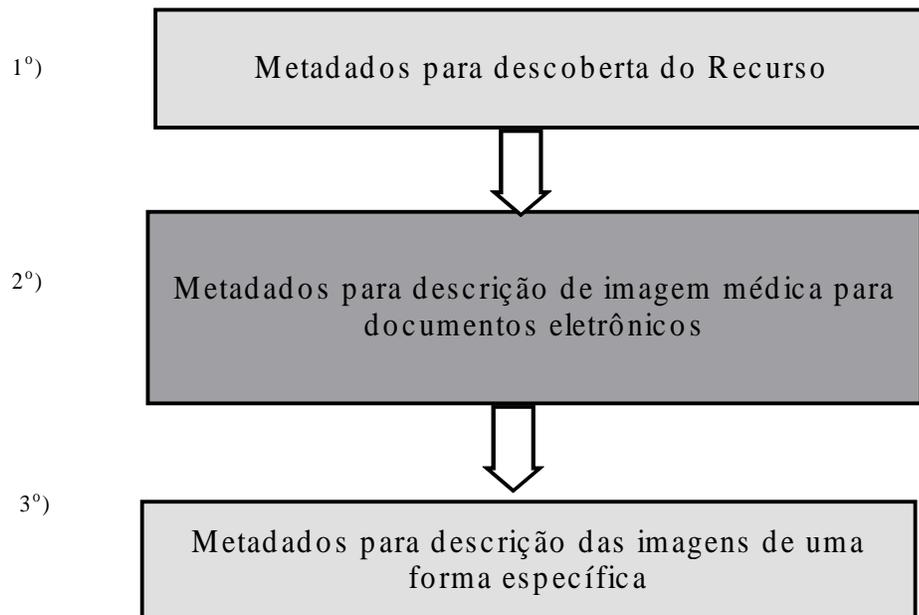


FIGURA 4-1 - Categorias de metadados em imagens digitais em documentos.

Na primeira categoria estão os metadados utilizados para a descoberta do recurso na *Web*; a segunda categoria refere-se ao conjunto de metadados necessários para a

descrição do documento imagem digital ; a terceira categoria refere-se aos metadados que descrevem o conteúdo da imagem, de acordo com as suas especificidades.

Nas duas categorias iniciais é utilizados o padrão Dublin Core, porém, na terceira, o conjunto de descritores fica diretamente associado ao tipo específico de imagem médica mantido pela ferramenta. Assim, como será observado no capítulo 5, o protótipo proposto para o estudo de caso possui objetos específicos para a descrição de uma imagem médica, para a qual é necessário um estudo que verifique o conjunto de descritores específicos que melhor a representam.

Com base nos formatos e padrões estudados, e tendo em vista a aplicação desses descritores em uma representação de imagens médicas digitais, foi necessário considerar um conjunto mais amplo de descritores. Conforme observado anteriormente, o formato DICOM é bastante direcionado para descrição de conteúdo da imagem, enquanto que o Dublin Core se preocupa com a descoberta de recursos. A utilização de uma imagem digital implica não somente na descrição minuciosa de sua semântica, como também das fontes destes documentos, de modo a permitir que uma cadeia de informação possa ser verificada, sendo que esta cadeia é iniciada a partir das referências contidas no documento.

Outra observação importante em relação à utilização de mais de um padrão de metadado, em uma mesma aplicação, é que alguns descritores podem ser comuns a estes padrões, necessitando apenas de um remapeamento [BID 1998] que será descrito no capítulo 2 ou indicação de qual a notação a ser utilizada. Seja por exemplo o descritor Source do padrão Dublin Core, que no formato da imagem DICOM é denominado pela tag VR (Valor de referência) “00080080”.

Para o presente trabalho, foram mantidos os nomes comuns em conformidade com o formato de atributos DICOM, justificando diretamente esta pesquisa para a descrição de imagens médicas. A TABELA 4.2 mostra o mapeamento de atributos DICOM para o formato *Dublin Core*.

TABELA 4.2 - Mapeamento DICOM para Dublin Core

DICOM (VR)	DUBLIN CORE
Localização da imagem (diretório)	Title
0008 0008	Type
0008 0018	Identifier
0008 0020	Date
0008 0080	Source
0008 0090	Creator
0008 1030	Subject
0010 0010	Contributor
Criador (ex: Renato)	Publisher

DICOM (origem)	Format
Fixo	Language
Todos os direitos reservados	Rights

## 4.5 A Recuperação de Informações Usando o Modelo Proposto

### 4.5.1 Modelo

Para tornar estas tecnologias, descritas no capítulo 1 e 2, um padrão que agregue funcionalidades de um sistema distribuído de imagens médicas, é necessário que ele atenda a esses requisitos. Uma forma é definir uma metodologia para que se represente um modelo de informação e uma sintaxe de representação, a fim de definir um mecanismo de transferência entre sistemas. É necessário avaliar alguns aspectos, como:

- O padrão oferece um modelo de dados para o conteúdo da informação?
- O padrão define objetos de interesse em sua área de aplicação? Por exemplo, ele define objetos como imagem, paciente ou relatório?
- O padrão oferece alguma maneira de representar instâncias destes objetos? Que tipo de dados, como números, caracteres ou imagens, são representados diretamente pelo padrão? Como descrever os dados para uma instância específica de uma tomografia computadorizada, por exemplo, incluindo informações do paciente?
- O padrão descreve como a informação deve ser comunicada entre sistemas? Ele define regras de comunicação e como estabelecê-la? Ou o padrão aborda uma técnica em particular?

Através desta metodologia é necessário avaliar a importância de cada padrão e sua contribuição para o desenvolvimento de um sistema de imagem médica que suporte a eficiência, escalabilidade, segurança e produtividade necessárias para a demanda clínica.

### 4.5.2 Resultados

Analisado cada padrão conforme um modelo de informação, representação de dados e mecanismos de transferência. A TABELA 4.3 apresenta um resumo das características de cada padrão. DICOM e HL7 definem um modelo de informação para seus domínios de interesse. CORBA não define um conteúdo particular de informação. Ele oferece a habilidade de representar objetos de informação e um mecanismo para transferir esses objetos. Já, SGML possibilita implementar uma sintaxe de representação que pode ser utilizada por qualquer outro padrão.

TABELA 4.3 - Resumo das Características dos Padrões

	<b>Modelo de Informação</b>	<b>Sintaxe de Representação</b>	<b>Mecanismo de Transferência</b>
<b>DICOM</b>	SIM	SIM	SIM
<b>HL7</b>	SIM	SIM	SIM
<b>CORBA</b>	NÃO	SIM	SIM
<b>SGML</b>	NÃO	SIM	NÃO

Algumas técnicas podem ser utilizadas para implementar interoperabilidade entre padrões. Segundo [BID 1998], apresenta alguns modelos de interoperabilidade:

- **Tunelamento/Encapsulamento:** esta técnica consiste em renomear os atributos uma dada classe (imagem), subordinando-a a uma outra classe que assume sua identidade (nome) e que simula sua interface para os demais componentes do objeto. É considerada como uma técnica de interoperabilidade para tratar um transporte de dados de um padrão para outro como se fossem uma 'carga', e carregá-la, sem interpretação, através de um transporte padrão. Por exemplo, decodificar uma imagem DICOM como um tipo de dado encapsulado através dessa técnica utilizando uma sintaxe específica para transportar essa imagem entre sistemas. A solução 'surge' através das interações entre a sintaxe de transporte da classe sujeita a reflexão a uma outra. A vinculação é definida estaticamente através de um arquivo de configuração, o qual especifica o mapeamento entre classes de um nível para outro.
- **Re-expressão/Mapeamento:** nesta técnica é considerada a interoperabilidade como uma tradução dos elementos de um padrão e convertendo e usando as regras e estilo de um padrão externo. A Re-expressão e representação dos elementos de dados de um padrão em outro, preservando a semântica original. Por exemplo, um objeto de informação DICOM e seus TAGs podem ser expressos como uma entidade SGML, com TAGs do padrão SGML representando as TAGs do padrão DICOM. Não existindo TAGs no padrão SGML para representar algumas TAGs do padrão DICOM, cria-se e preserva-se a semântica DICOM.
- **Harmonização:** esta técnica mostra que um modelo de dados pode ser estendido ou aumentado, usando um outro padrão. Quando os padrões compartilham modelos de dados, a informação pode ser representada e

comunicada sem perda de fidelidade usando ambos os padrões. A nova forma desta abordagem ocorre quando os modelos não compartilham com modelos de dados.

### 4.5.3 Discussão

Analisando os padrões de distribuição de objetos, nota-se que o melhor modelo, relacionando custos e benefícios, para a implementação de um sistema de imagens médicas, considera o uso de DICOM como padrão para a representação e comunicação de imagens médicas. Também é analisado o padrão para visualização na *Web*, de forma que são necessárias a utilização das técnicas de Tunelamento/Encapsulamento e Re-expressão/Mapeamento para traduzir o conteúdo da imagem DICOM para o padrão SGML. Também é usada a técnica de harmonização, para adaptar o modelo de dados DICOM ao padrão próprio desenvolvido na ferramenta. Segundo [SLE 2000] a forma mais livre de representar um objeto imagem está na generalização da estrutura de “tags” SGML.

A necessidade de integrar serviço computacional com imagens apresenta uma tendência conformar-se ao DICOM a uma aplicação dinâmica, de modo que a união destas tecnologias deve estar apta, necessariamente a reconhecer este padrão [PIS 2000]. A arquitetura *Web* com *site* dinâmico representa uma estratégia mais flexível e funcional dentro dos padrões de interoperabilidade disponíveis atualmente. O objetivo é desenvolver e utilizar um sistema integrado para gerenciar imagens médicas que utilize uma metodologia que suporte a produção de programas modulares, reutilização de códigos-fonte, permita útil integração entre diferentes programadores, sistemas operacionais, equipamentos, e melhore a manutenção de longo alcance para o código.

A implementação deste modelo está em sua versão preliminar. Este projeto faz parte das atividades da obtenção de um sistema de apoio à saúde cujos objetivos incluem desenvolver uma vasta funcionalidade para um sistema *Web*, incluindo visualizadores de formatos do padrão de transferência de sintaxe DICOM, repositório de imagem distribuída, e extração de conteúdo de imagem.

## 4.6 Considerações Finais

Neste capítulo verifica-se a necessidade da utilização dos metadados para o tratamento de imagens digitais, porém foi verificado que para a efetivação desta utilização, necessita-se, às vezes, em uma mesma aplicação, de diferentes padrões ou formatos. No caso em estudo, as necessidades são de metadados para: localização do recurso, descrição do conteúdo, semelhança de descritores a partir da transferência de sintaxe em conformidade com o padrão DICOM, podendo ser representada e utilizada na ferramenta proposta. Toda esta representação será posteriormente utilizada em conjunto com a ferramenta sustentada pela tecnologia descrita no capítulo 2.

A visão completa das necessidades é algo muito específico e somente visível após um estudo criterioso das características da imagem e das necessidades de pesquisa e manutenção. Como resultado desta avaliação, foi proposta a utilização de alguns

descritores de DC em conformidade com o padrão DICOM, definindo uma proposta de transferência sem perda de sua característica.

## 5 Protótipo do Sistema Proposto

Para o desenvolvimento deste protótipo foi utilizado o sistema operacional LINUX e, como linguagem de programação, foi escolhida a linguagem PHP (Pré-processador de Hipertexto) [CAS 2001]. Esta linguagem tem como principais características: a habilidade de combinar o código HTML com scripts e rodar estes códigos escritos em linguagens tais como o JAVA, no lado do servidor, possibilitando, além de um menor transporte de códigos, impedindo a visualização desses scripts, ou seja, apenas o código HTML fica visível. O XML também se torna um aliado na criação da página dinâmica, pois ele é responsável pela criação da estrutura dos dados em forma de “tags” publicada em tempo de execução.

Após o estudo da estrutura das imagens DICOM para a implantação de um *site* de visualização na *Web* e também o das tecnologias para a criação desse, foi apresentada a solução de uma ferramenta para a extração de metadados para suporte a profissionais da saúde.

Depois de alcançado o objetivo (solução do problema), a etapa seguinte do trabalho foi colocar em prática a solução proposta, que consistiu na criação do protótipo de uma ferramenta de visualização e troca de informações de imagens médicas através da *Web*. Este protótipo foi desenvolvido nas linguagens Java, PHP, XML.

### 5.1 Ambiente dinâmico

Todas as páginas internas do protótipo possuem um documento que contém o código gerado nas linguagens (PHP, HTML, JAVA, XML) necessário para a sua execução. Assim, se um dia precisar ganhar uma nova função, essa inclusão poderá ser feita apenas como rotina juntamente onde residem estes scripts. Uma vez alterado esses documentos, todas as páginas que o utilizam estarão automática e instantaneamente atualizadas.

Para que o protótipo funcione de maneira sólida, foram utilizadas algumas tecnologias que permitem construir *sites* dinâmicos, ou seja, *sites* onde as páginas são montadas pelo sistema a cada nova visita, de forma que elas possam ser atualizadas automaticamente da escolha do usuário a uma tarefa.

Sem dúvida nenhuma, a utilização de uma filosofia ASP<sup>1</sup>, conforme descrita no item 1.3.3.2 facilitou o desenvolvimento deste protótipo, dando uma maior comunicação e aproveitamento das facilidades desenvolvidas nesta aplicação.

#### 5.1.1 Sistema Operacional

O sistema operacional Linux de fonte aberto foi utilizado por ter a característica. a natureza do fonte aberto, o que significa que a aplicação é de alguma forma exclusiva e construída de forma customizada com vários pedaços de solução colocados em um só lugar. Dessa forma, o protótipo implementado se torna altamente modular dependente apenas do conhecimento da tecnologia de quem o projeta e o implementa. Por ter esta

característica, a escolha do Linux significa buscar uma estratégia centralizada para acompanhar as tecnologias e hardware emergentes. Para tanto, a utilização do Linux, juntamente com o servidor de páginas APACHE, foram de grande importância na integração das linguagens descritas a seguir.

### 5.1.2 Java

O motivo da escolha desta linguagem foi o fato de a mesma linguagem portátil para qualquer plataforma de hardware e poderosa devido a sua estrutura. Um outro motivo foi o de usar o conceito de orientação a objetos, com isso ficaria mais facilitado à utilização de variáveis para manipulação de arquivos das imagens.

Um outro motivo importante, é a sua integração em conjunto com o sistema operacional Linux no qual todo o protótipo foi desenvolvido. Também um fator fundamental na escolha desta linguagem, foi à facilidade de poder utilizar algumas funções para manipular arquivos.

### 5.1.3 PHP

O principal motivo na escolha desta linguagem de script é o fato de poder rodar no lado servidor, facilitando o desenvolvimento da interface para o usuário.

Outro motivo fundamental é o de se criar um *site* multiplataforma, que irá rodar tanto em plataformas Windows como Linux, por exemplo. Sendo assim, o uso desta linguagem de programação script, atingira uma maior amplitude, pois a ferramenta passa a ter a característica de um repositório de dados para diferentes sistemas operacionais. Neste caso, um fator fundamental, é o custo, no do sistema operacional Linux, a sua distribuição é gratuita, fator contrário do que acontece no Windows.

### 5.1.4 XML

A utilização desta linguagem de marcação foi muito importante neste protótipo, porque proporciona uma facilidade de expressar o objeto imagem de forma estruturada, facilitando a interoperacionalidade de sistemas.

Como o XML é um arquivo texto, com uma estrutura que pode ser definida na fase de desenvolvimento, podendo armazenar de forma que os atributos da imagem fiquem armazenados em diretório específico.

Um dos grandes atrativos do XML foi poder separar os dados da interface. Ele exibe os dados no *browser*, simplesmente definindo o conteúdo, valendo-se de delimitadores muito simples para descrição de dados de dados das imagens (metadados). Quanto à apresentação em *browsers*, foi de grande importância, pois a Cascading Style Sheets (CSS) possibilitou assim, o desenvolvimento para de definir a interface do usuário, bastando apenas utilizar diferentes stylesheets (folhas de estilo) sem necessitar a modificação nos scripts em PHP.

A extensibilidade e flexibilidade do XML permitiram que sua descrição se aplique a uma grande variedade de aplicações heterogêneas. Como o XML possui em sua própria estrutura, a descrição dos dados, a descrição e o processamento no *servidor de aplicação*, foi possível ser efetuada a inclusão de instruções dentro de programas Java para ser apresentado no *browser* do usuário.

## 5.2 Características do Protótipo

### 5.2.1 Arquivos Gerados

O objetivo principal da ferramenta é extrair e gerar informações das imagens residentes no *site*, através de um conjunto de formulários. Nestes formulários, existem diversas variáveis que tem a finalidade de armazenar valores (como por exemplo, nome do paciente de um arquivo DICOM) e posteriormente, serão utilizadas, para gerar arquivos do tipo txt, Html, PHP, XML.

#### 5.2.1.1 Arquivos do tipo PHP, HTML, Java e XML

Os arquivos no formato PHP são utilizados no *site* para realizar transações no lado servidor. Entre essas transações, pode citar-se a inclusão de uma imagem no repositório de dados, onde, primeiramente, eles estão inseridos na própria imagem DICOM, partindo do lado cliente e depois são gravados no servidor. Outros exemplos são os arquivos Java compilados, que servem para a extração do conteúdo da imagem e transformados em metadado.

Os arquivos PHP ficam encarregados de receber as informações dos programas Java, no qual transforma em formulários XML e codifica para sua posterior visualização. São realizadas rotinas, a partir de scripts PHP a execução de programas Java para a criação de arquivos auxiliares a serem publicados no lado do cliente. Algumas características da ferramenta em relação à proposta das tecnologias escolhidas e determinantes no seu desenvolvimento foram:

#### PHP:

- toda a montagem das páginas (menu, lista de imagens) foi feita com esta linguagem script;
- programas de configuração dos diretórios específico da ferramenta e armazenamento de variáveis locais do servidor;
- transformação da estrutura de visualização realizada no cliente (*browser*).

#### XML

Um arquivo referente à imagem estudada é gerada em tempo de execução e, logo após, exibida no *browser* do cliente. Esses arquivos são criados e armazenado no

diretório “XML”, sendo, posteriormente, referenciado na execução da escolha da imagem.

## JAVA

Dois diretórios são usados para a execução das classes Java descritas mais a diante. No Item 5.4, são especificados os arquivos que serão gerados por esses programas.

### 5.2.1.2 Arquivos de Configuração da Ferramenta

Toda vez que é feita à execução da ferramenta é lido no arquivo de configuração, onde é determinado o caminho dos diretórios para o armazenamento de dados e execução de programas. A ferramenta lê o arquivo de configuração para que o usuário possa usar de forma transparente para o seu entendimento na questão de localidade dos diretórios

### 5.2.1.3 Arquivos de Imagens

A ferramenta trabalha com um diretório específico para as imagens DICOM. Esse, por sua vez, é determinado pelo arquivo de configuração já descrito anteriormente. Os arquivos das imagens devem ser exclusivamente criados com o padrão DICOM, onde o nome de sua a extensão poderá ser nome qualquer.

Os métodos implementados em Java são todos para consultar às classes do esquema de forma dinâmica. São disponibilizados todas as propriedades de cada classe para que o usuário possa fazer a consulta de acordo com a sua necessidade; a partir daí o método gera dinamicamente a consulta por meio do comando para fazer através da leitura uma verificação do tipo da imagem. O fator determinante é que esta imagem gerada seja originária de um equipamento específico para esta finalidade. Sendo assim, para a ferramenta filtre essa condição é necessário que, no seu “Header”, contenha o sob a forma de texto o prefixo “DICM”. Essa funcionalidade garante que a ferramenta só aceite esse tipo de arquivo, rejeitando qualquer que seja o tipo diferente dessa condição.

## 5.2.2 Diretórios da Ferramenta

O usuário não tem a possibilidade de escolha em qual diretório o repositório ira gravar os dados, a própria ferramenta se encarrega de fazer. Sendo assim, toda vez que a ferramenta é executada existe uma estrutura de variáveis que armazenam os nomes e os caminho para a sua execução FIGURA 5-2. Para tal propósito, foi criada uma estrutura de arquivos para que possam ser copiado as imagens e seus componentes em diretórios específicos para a criação dos metadados extraídos destas mesmas imagens.

Tendo armazenado nas variáveis, os nomes dos diretórios de destino, a ferramenta gera dentro desse, um diretório raiz e mais outros três subdiretórios:

#### 5.2.2.1 Diretório “*raiz*” arquivos php:

- interface do usuário, para proporcionar transações como enviar/receber arquivos;
- exibir informações das imagens;
- configuração de variáveis de caminhos do sistema;
- chamada de classes Java.

#### 5.2.2.2 Diretório “*dicom*” as imagens

Neste diretório se encontram as classes em Java nas quais executam em conjunto a partir dos scripts em PHP.

#### 5.2.2.3 Diretório “*arquivos*” repositório:

Neste diretório são armazenados arquivos resultantes da geração da execução das classes em Java onde se divide em:

- *PNG*: Arquivos de imagens em formato PNG extraído a partir das imagens DICOM através das classes em JAVA.
- *XML*: Arquivos gerados em formato XML a partir das classes JAVA contendo as estruturas para a apresentação no *browser*.

#### 5.2.2.4 Diretório “*gerardf*” arquivos :

Repositório de arquivos no formato XML com a estrutura de acordo com o padrão RDF *Dublin Core*, extraída da imagem DICOM.

### 5.2.3 Estruturação Interna (variáveis)

Para a geração do ambiente a extração de metadados das imagens, é preciso armazenar valores que, futuramente, servirão de forma direta para a criação das páginas. Esses valores são armazenados em variáveis que podem conter o título de uma página, os dados extraídos que deverão ser usados nesta, o diretório e o endereço onde ficará hospedado o dado relativo à imagem.

#### – Variáveis “Externas”

Foi usada esta denominação, para especificar tipos de variáveis, este tipo especificamente trata de variáveis que interagem com o usuário da ferramenta. Ou seja, são todas as variáveis que recebem as configurações do *site*, como citadas anteriormente. Essas são usadas para:

- armazenar as configurações do *site* em um arquivo para o seu uso posteriormente;

- referenciar imagens que foram escolhidas pelo usuário da ferramenta que serão exibidas nas páginas;
- variáveis que armazenam textos escritos na ferramenta, para serem visualizados nas páginas do *site*;
- títulos e cabeçalhos que serão mostrados;
- diretórios do repositório de dados para a hospedagem das imagens.

– Variáveis “Internas”

As variáveis “internas” são aquelas que não possuem a interação direta com o usuário da ferramenta. Ou seja, essas variáveis fazem a comunicação com as variáveis “externas” da ferramenta. Essas variáveis serão utilizadas para:

- armazenar em um arquivo lógico toda a estrutura de uma imagem de modo a demonstrar em uma página específica. Pois primeiramente, é necessário escrever todos os valores de referência da imagem em uma variável do tipo texto. Após, o valor correspondente à variável é escrito dentro de um arquivo de nome “mr10.dc3.xml”, por exemplo;
- variáveis que fazem referência ao nome da imagem que será copiada para um diretório específico do *site*;
- variáveis que têm a finalidade de fazer a cópia da imagem em seus respectivos diretórios equivalentes no *site*.

## 5.3 Estruturação dos módulos

### 5.3.1 Características de Implementação

Este sistema, acessível via *Web*, permite a construção das páginas de forma dinâmica, no momento da sua utilização. As associações armazenadas em diretórios e codificadas com o uso da linguagem de marcação XML (EXtensible Markup Language) páginas criadas, dinamicamente extraídas da imagem DICOM. Um exemplo de código XML é exibido no Exemplo 5-1. Este código tem como objetivo permitir o início da definição da interface do usuário, onde são disponibilizadas as categorias de apresentação inicial que irão compor suas principais telas. A partir da primeira seleção, o conteúdo da imagem é extraído e disponibilizado em “tags” XML. No exemplo, o programa `arquivo.php` é utilizado para extrair e armazenar em diretórios específicos determinados internamente pela ferramenta.

```
<table width="100%">
  <tr>
    <td bgcolor="black">
      <table bgcolor="#ffefdb" width="100%">
        <tr>
```

```

<td bgcolor="#EEDFCC">
  <font face="verdana" size=2>
    <center><b>Lista de Arquivos</b></center>
  </td>
</tr>
<tr>
<td>
  <table>
    <tr>
      <td><font color='darkblue'><b>Nome</b></td>
      <td><font color='darkblue'><b>Tamanho</b></td>
    </tr>
  <?
  $teste="renremrnemrne";
  include('config.php');
  chdir($arqdir);
  $dir = dir(".");
  $dir->rewind();
  while ($file=$dir->read())
  {
    if (is_file($file))
    {
      echo("<tr>");
      echo("<td><a href='download.php?arquivo=$file'>$file</a></td>");
      $stamanho = filesize($file);
      echo("<td>$stamanho</td>");
      echo("<td><form action='extraicab.php' method=get>
        <input type='hidden' name='arquivo' value='$file'>
        <input type='submit' value='Extrair cabeçalho DICOM'>
        </form></td>");
      echo("<td><form action='visualiza.php?arquivo=$file' method=get>
        <input type='hidden' name='arquivo' value='$file'>
        <input type='submit' value='Visualizar imagem'>
        </form></td>");
      echo("<tr>");
    }
  }
  $dir->close();
  chdir($basedir);
?>
</table>
</td>
</tr>
</table>
</td>
</tr>
</table>

```

Exemplo 5-1 Programa em linguagem PHP – arquivos.php

Para uma melhor compreensão do modelo, é possível estruturar o sistema, segundo a representação exibida na FIGURA 5-1, onde observamos os módulos que o compõe:

### **5.3.2 Módulo de Controle da Interface do Usuário**

Neste bloco são providos os recursos para o usuário, através das associações entre as categorias de pesquisa do modelo de interfaces, que objetivam:

- definir uma forma personalizada de trabalho, criando para tal uma interface que contenha as categorias relevantes para suas pesquisas;
- recuperar as interfaces previamente criadas através da identificação do usuário;
- consultar o repositório de imagens por meio de pesquisas diretas.

### **5.3.3 Módulo de Controle dos Metadados**

O nível intermediário é reservado à manipulação do repositório das imagens, onde é representado o modelo de descritores para o controle dos documentos. As seguintes funcionalidades devem ser permitidas para este módulo: geração dinâmica, visualização e funções de gerenciamento de metadados;

### **5.3.4 Módulo de Controle do Repositório de Imagens**

Finalmente, no terceiro nível, encontra-se o sistema gerenciador de arquivos responsável por armazenar as imagens, pois o protótipo tem por objetivo inicial ser um banco de metadados em formato texto e imagens.

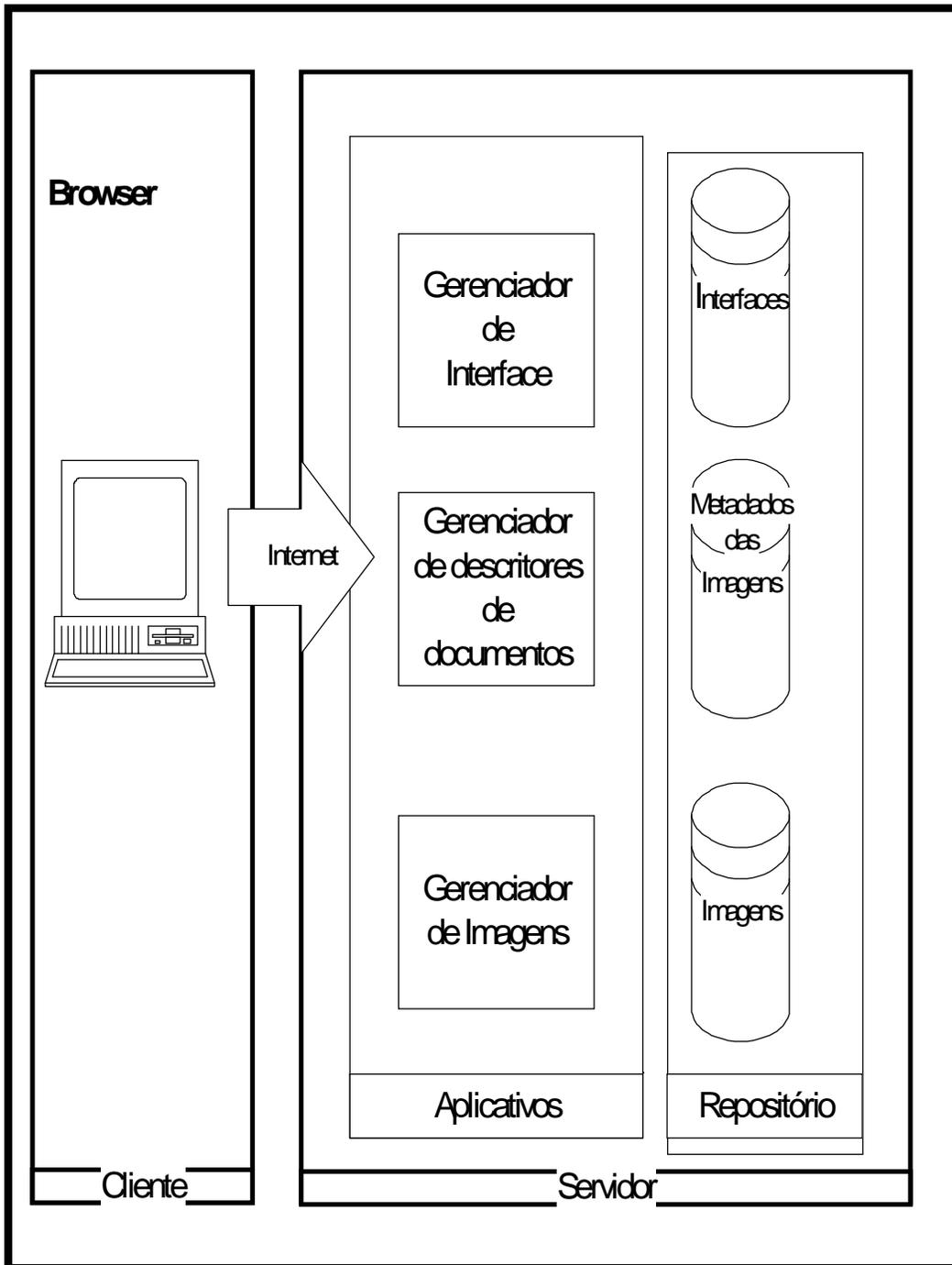


FIGURA 5-1 - Estrutura do sistema

## 5.4 Diagrama do Protótipo

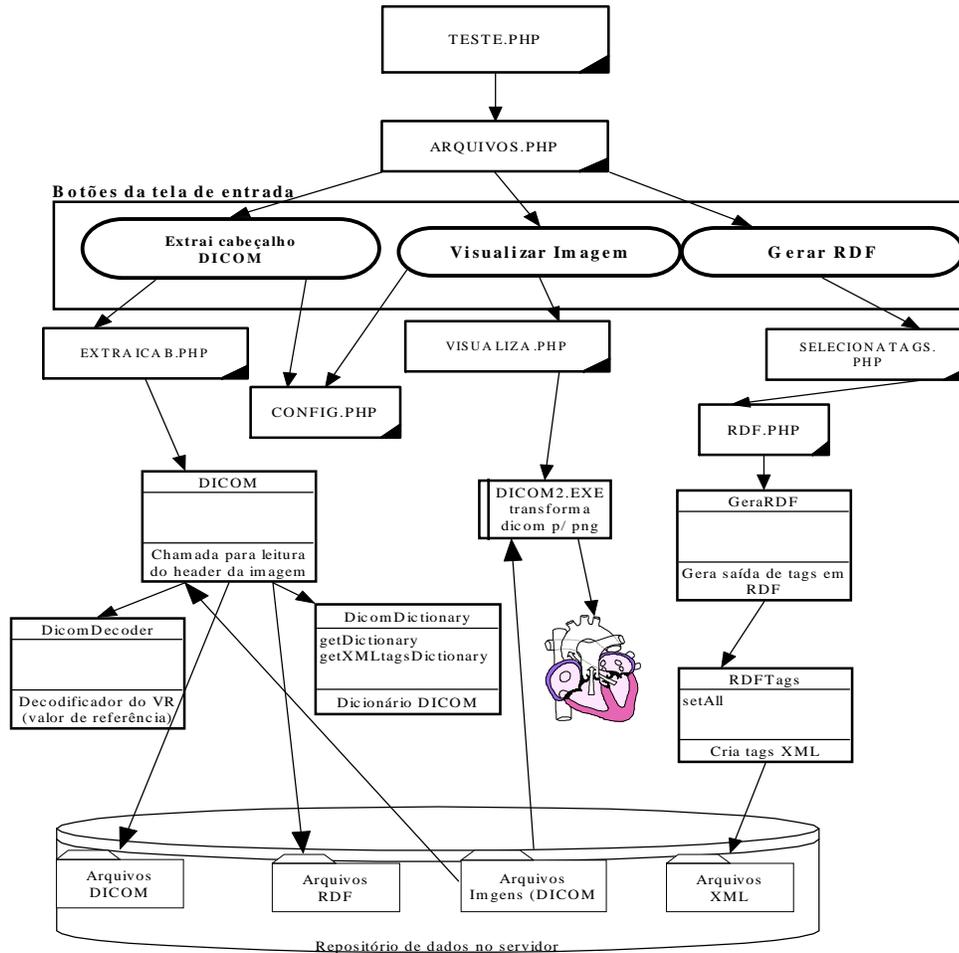


FIGURA 5-2 - Diagrama interno do Protótipo

### 5.4.1 Repositório de dados do Protótipo

#### XML

Os arquivos XML são utilizados com o intuito de fornecer interoperabilidade sintática aos recursos na *Web*. Sendo assim, o repositório possibilita separar o conteúdo do *layout*, proporcionando na aplicação um conjunto extenso de *tags* que são utilizadas para capturar a estrutura do documento em estudo. Com o uso da arquitetura RDF em conjunto com o Esquema RDF (*RDF Schema*) compõem um mecanismo para prover facilidade de descrição de forma genérica. Desta forma, triplas (sujeito, predicado e objeto) denotam relações entre pares de objetos (imagens).

A estrutura do XML FIGURA 5-2 é formada basicamente por três elementos ou arquivos principais que se completam para determinar o armazenamento do conteúdo e sua apresentação:

- Esquema XML;
- CSS (Cascading Style Sheets);
- Conteúdo XML.

A linguagem XML embutida em programas escrito em Java provê o formato dos dados para documentos estruturados, o que irá identificar esse vocabulário será o Esquema XML.

Uma das vantagens da utilização de arquivos XML é o de proporcionar que os dados dos documentos podem ser apresentados de diversas formas. Neste caso, o site da aplicação foi desenvolvido de forma que se o usuário necessitar de mudanças na apresentação dos dados, ele deverá fazer a alteração no código escrito em Java.

A escolha da utilização da linguagem XML foi fundamental para atender duas das três condições para que haja a representação de conhecimento: a interoperabilidade sintática e a interoperabilidade estrutural. A interoperabilidade sintática da linguagem XML foi atingida através do Esquema XML, que provê regras precisas, utilizando uma gramática, que estabelece a sintaxe. A interoperabilidade estrutural da aplicação foi atingida através da linguagem XML que se satisfaz provendo uma representação dos dados e especificando tipos e possíveis valores para cada forma de representação, que também é realizado através do XML . A única restrição quanto à criação dos arquivos, foi à condição que o XML não cria é a interoperabilidade semântica, pois a sua estrutura de dados não dá significado semântico aos arquivos gerados

## **RDF**

Os arquivos RDF (*Resource Description Framework*) são criados através da definição dos metadados (dados que descrevem o conteúdo, estrutura, representação e contexto de algum conjunto de dados extraídos da imagem DICOM). Arquivo RDF é a base para o processamento de metadados, no qual ele providencia a interoperabilidade entre as aplicações, trocando informações compreensíveis das imagens e gravando no repositório as estruturas para serem apresentadas no navegador do usuário.

## **5.5 Definição**

O sistema DICOM Database foi criado como um projeto de Mestrado da UFRGS. Tem, por objetivo principal, prover informações para as pesquisas e estudos na área da Ciência da Computação e, seu público alvo é formado pelos profissionais da saúde. Um outro propósito para o projeto é a criação de um repositório de imagens de acesso a informações locais, onde este que será utilizado para a efetivação da distribuição do banco de imagens entre o repositório e os profissionais. Este conjunto todo é formado de tecnologias e usuários de modo que sua finalidade possui características de um repositório de imagens de pacientes, interagindo com os médicos, caracterizando um serviço de manipulação de informações contidas nas imagens.

### 5.5.1 Avaliação do Desenvolvimento

Nesse protótipo os métodos foram implementados utilizando a forma de captura textual para o armazenamento dos valores das propriedades das classes (DICOM), isto é, são instanciados a partir da extração do seu conteúdo, onde o valor da propriedade é transferido para a ferramenta DICOMDatabase.

Embora não tenha sido implementado nenhum método de captura para outro tipo de imagem, por exemplo: JGP, BMP etc..., esse protótipo poderia ser utilizado juntamente com técnicas das outras abordagens estudadas, para constituir um meio mais rico de recuperação de imagens.

Ficou claro que a arquitetura *Web* é atualmente a principal escolha para a implementação de um sistema de visualização de imagem, acompanhando a tendência mundial de migrar os aplicativos para um modelo Internet, com todos os benefícios que esse tipo de arquitetura oferece como é descrito no capítulo 2.

Os resultados esperados em referencia a padronização de representação de metadados para imagem, se mostrou que alguns não usam nenhum padrão e que há muitos padrões. Apesar disso não ser uma grande surpresa, reforça a necessidade de fortalecer a padronização nos sistemas para este propósito. Por outro lado, o uso do XML na grande maioria dos sistemas já mostra que, apesar de não haver uma forte padronização interna, os sistemas de visualização de imagens na *Web* estão, a cada dia, mais direcionados a interagir com outros sistemas, na chamada interoperabilidade, o qual os força a padronizar as mensagens de troca, utilizando-se padrões aceitos nacional ou internacionalmente. Assim, espera-se que a padronização seja uma consequência da crescente necessidade da troca de dados entre os sistemas.

## 5.6 Problemas Encontrados e Soluções Propostas

O DICOM Database envolveu um estudo sobre algumas técnicas para representação de imagens digitais. Porém, na sua implementação, não foi usado nenhum banco de dados, limitando-se ao uso da ferramenta apenas o repositório em diretórios específicos, criados no ambiente LINUX, com suporte do gerenciador de páginas *Web*, o APACHE.

Outro ponto a ser observado diz respeito ao esquema de mapeamento e atributos da imagem, onde no DICOM Database se optou por um esquema próprio de mapeamento, relativo entre DICOM e metadados DC Dublin Core. Na verdade, tal fato pode vir a criar problemas no futuro, principalmente, quando se deseja compartilhar ou integrar outras imagens ao repositório. Portanto, para o estudo de caso, foram propostos os seguintes itens:

- adoção do esquema de metadados da W3C, por ser esta a mais completa e difundida na área de Ciência da Computação;

- adoção de um servidor de páginas (Apache), para dar o suporte necessário em termos de segurança e facilidade do mecanismo de gerenciamento das páginas geradas e disponibilizadas para clientes *browsers*;
- aplicar o modelo de descritores, para prover maiores possibilidades de representação semântica sobre o conteúdo da imagem, o Re-expressão/Mapeamento, proposto por [BID 1998] DICOM para Dublin Core (DC);
- aplicar ao modelo interfaces, para possibilitar uma qualidade de apresentação para o usuário;
- aplicar ao modelo a possibilidade de extrair conteúdo em tempo real de um objeto incluído no repositório de dados (imagem digital);
- Restringir a funcionalidade da ferramenta a sua manipulação apenas uma imagem de cada vez para, isto é, limitando imagens serializadas.

## **5.7 Funcionalidade e Especificação da Ferramenta e Apoio ao Modelo Proposto - DICOM DATABASE -**

Como resultado do desenvolvimento, foi criado o protótipo DICOM DATABASE a qual é associado à implementação de um sistema de auxílio médico via *Web*. Foram utilizadas técnicas da Engenharia de Software para o desenvolvimento na *Web*.

O Sistema DICOM DATABASE, como pode ser visto neste capítulo, foi desenvolvido, utilizando-se uma metodologia baseada em estudos de abordagens de imagens estáticas mapeadas sob padrões e arquiteturas de metadados resultou em um desenvolvimento detalhado, com a valorização das etapas de requisito, análise e projeto, para o total entendimento daquilo que iria ser implementado, antes de propriamente fazê-lo. Devido à natureza incremental das tecnologias de implementação, foram sendo feitas em partes modulares, devido à complexidade de cada ferramenta utilizada, permitindo assim, um ciclo de implementação mais eficiente.

Dessa forma, implementou-se o protótipo propondo um sistema de visualização e extração de metadados na *Web*, no qual os seguintes conceitos foram implementados:

- acessibilidade pelo médico;
- independência de funcionalidade associada a ferramentas externas;
- integração com outros modelos e padrões existentes;
- utilização totalmente em software livre;
- técnicas de visualização através de ferramentas específicas;

- livre de patentes.

O sistema apresenta os seguintes recursos:

- atualização do *site* através de processo dinâmico;
- um tipo de usuário: profissional de saúde;
- armazenamento de arquivos em formato digital (imagem);
- envio de imagem para um repositório específico;
- resultado através da extração de conteúdo;
- visualização da imagem.

O sistema está disponível para testes na URL: <http://200.132.46.254/site/renato/tmp/teste.php> DICOM DATABASE (home page do sistema na FIGURA 5-3).

## 5.8 Organização da Integração do Protótipo

A fim de exemplificar as funcionalidades do protótipo, serão exibidas, as principais telas do sistema: A tela principal FIGURA 5-3 apresenta as opções de uso da extração de conteúdo da imagem armazenada no protótipo por meio de botões clicáveis. Nesta tela, são disponibilizadas três (3) possibilidades que o protótipo oferece: extração do conteúdo do formato DICOM, visualização da imagem no *browser* e geração de Dublin Core (DC)



FIGURA 5-3 - Tela da aplicação via Browser.

Como já mencionado, as descrições da estrutura interna da ferramenta são utilizadas basicamente na execução de cada botão, conforme podem ser vistas na . A partir desta tela, a ferramenta, um conjunto de botões, referenciando-se a cada imagem existente no diretório específico das respectivas.

O papel da identificação da aplicação e de suas tarefas, bem como das informações agregadas a ela, representadas através de botões, são os meios pelos quais a navegação pela aplicação, é possível. Existem três funcionalidades principais neste contexto, como mostra a FIGURA 5-4:

- visualizar individualmente as propriedades DICOM relativas e decorrentes da execução da leitura do conteúdo da imagem, através do botão *Extrair Cabeçalho DICOM*;
- visualizar a imagem propriamente dita na tela do *browser*, através do botão *Visualizar Imagem*.
- Visualizar as propriedades da aplicação e de cada metadado extraído da imagem, através de inferência sobre esta. Quando da visualização das propriedades da tarefa ou da aplicação, as “tags - RDF” agregadas a elas são trazidas juntas, mantendo-se uma estrutura formada pelos atributos *Dublin Core* através do botão *Gerar RD*.

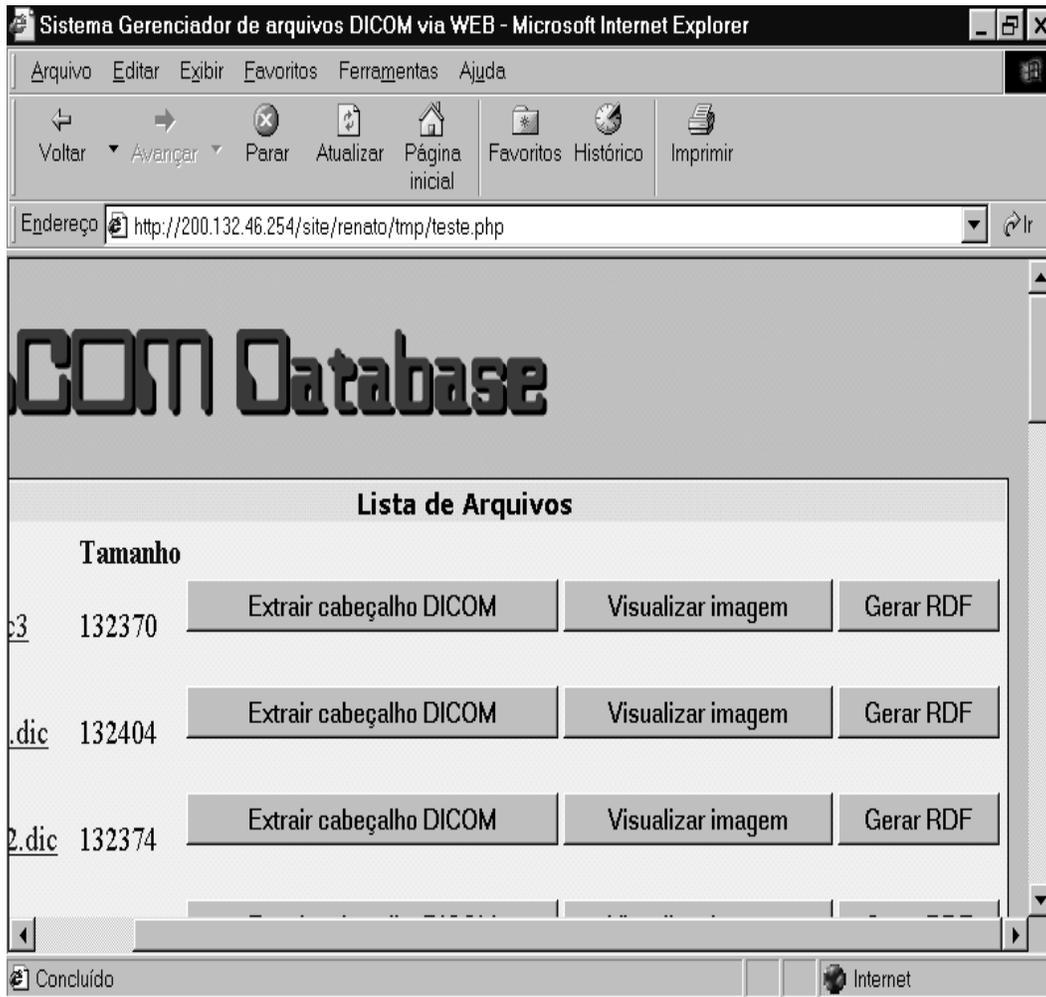


FIGURA 5-4 - Botões de execução

No DICOMDatabase pode ser realizada a operação de enviar a imagem para o servidor, de modo que seja armazenada em diretório específico para esta finalidade. Para operação, a ferramenta permite que o usuário escolha um arquivo de imagem do padrão DICOM e envie de acordo com sua escolha. Para esta finalidade, como mostra a FIGURA 5-5 apresenta a função que enviará para o repositório.

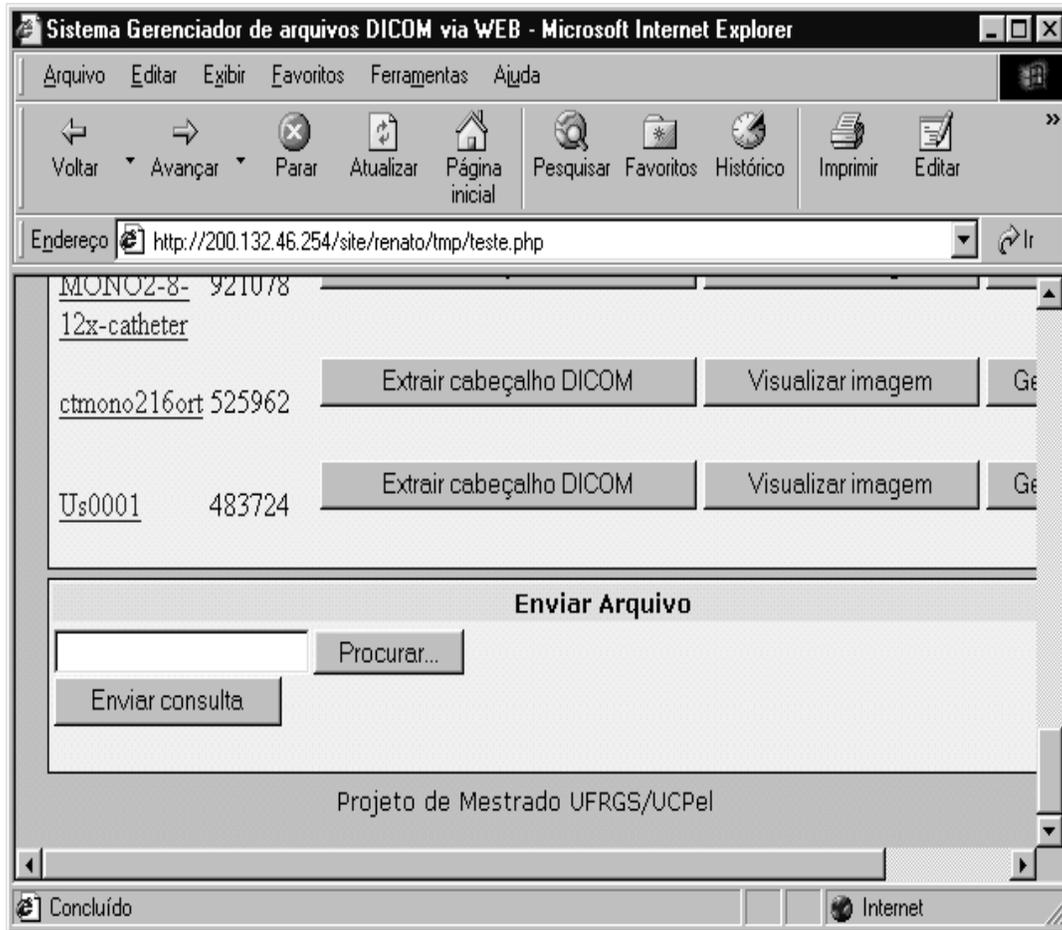


FIGURA 5-5 - Tela de envio das imagens ao repositório

### 5.8.1 Execução da tarefa para Extrair Cabeçalho DICOM

A começar pela criação ou visualização desta tarefa, pode-se gerar um conjunto de propriedades extraídas da imagem DICOM. Nesta situação FIGURA 5-6, o usuário também tem a possibilidade de visualizar de forma organizada a estrutura básica do formato DICOM.



## Execução da tarefa para Visualizar Imagem

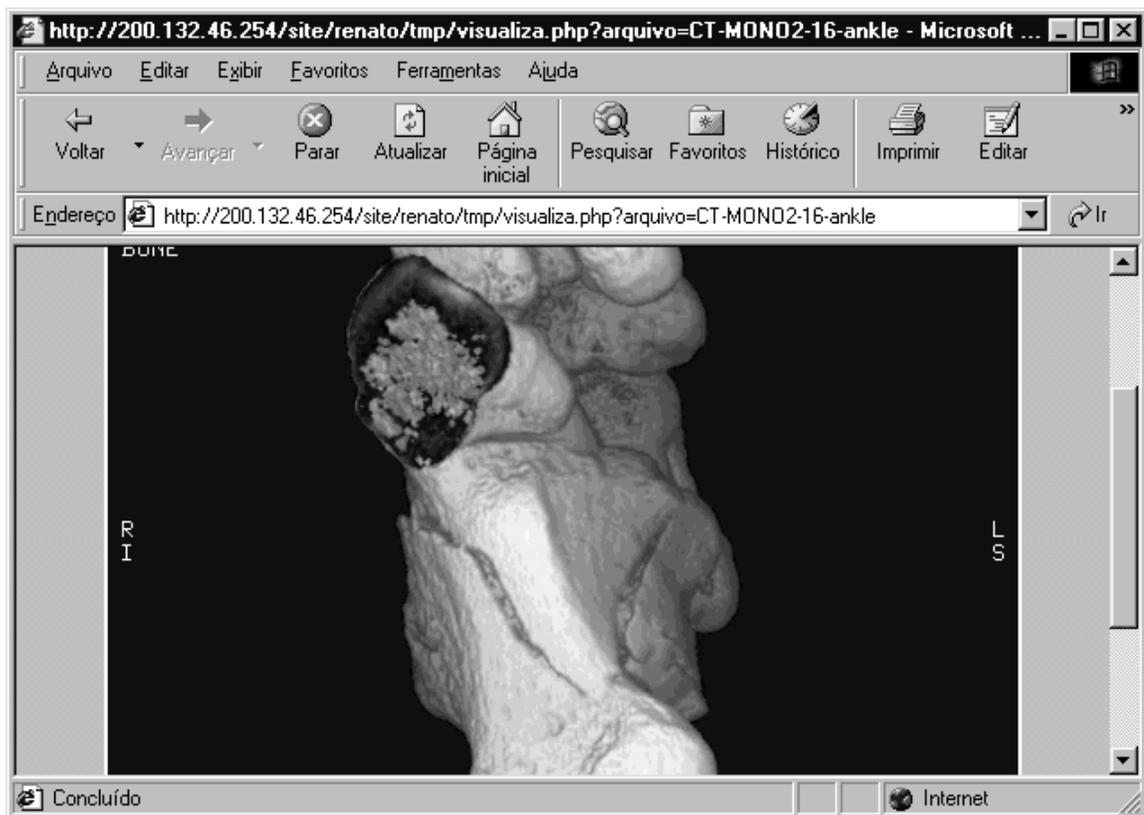


FIGURA 5-7 - Tela de visualização da imagem.

### 5.8.2 Execução da tarefa para Gerar RDF

Motivado pelo grande número de informações relacionadas à aplicação, tarefas e execuções, alguns requisitos foram levantados no **item 4.4.3**. Baseado nisso, optou-se por idealizar uma área de visualização opcional, na qual a referência de “tags” - TABELA 4.2 representam a existência de determinados tipos de informações agregadas à imagem em estudo.

Portanto, a visualização dos metadados de qualquer imagem, dentro do ambiente dar-se-á através da identificação (marcação), ou escolha da propriedade correspondente, como mostra a FIGURA 5-8 os atributos para a escolha do usuário. A partir desta escolha a interação da ferramenta com a imagem armazenada em diretório se dará em conformidade com o escolhido, resultando as estruturas de RDF como mostra a FIGURA 5-9.

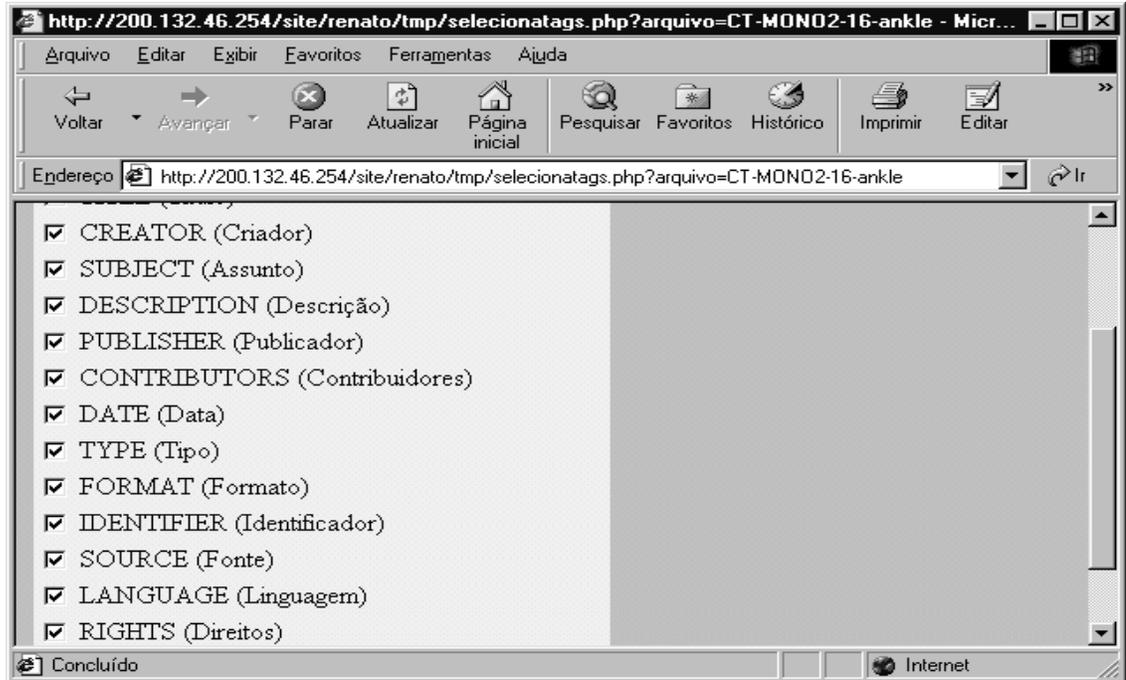


FIGURA 5-8 - Tela de opção de atributos DICOM para RDF –Dublin Core.

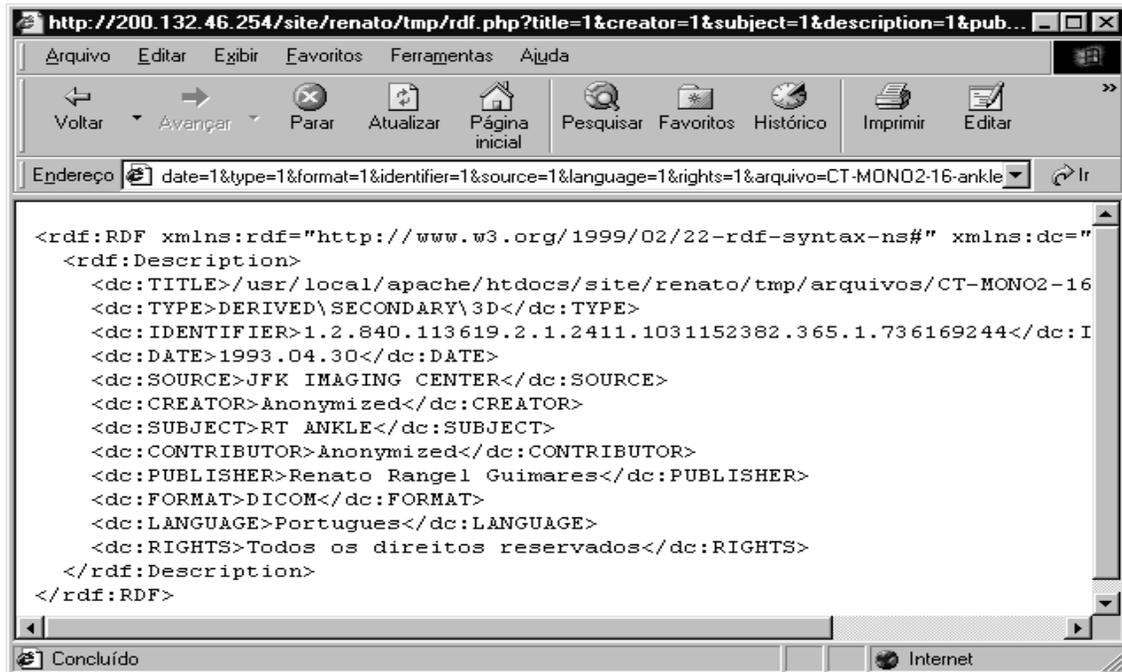


FIGURA 5-9 - Tela de visualização de DICOM para RDF- Dublin Core.

## 5.9 Utilização do Esquema em XML

Em consonância com a relevância da XML, no contexto atual de desenvolvimento de recursos em ambientes distribuídos tal como WWW é importante situar como o presente esquema foi desenvolvido no contexto da XML.

O esquema mostra as vantagens da XML e a possível ligação dessa nova tecnologia com o esquema para documentação e recuperação de imagens, proposto no Capítulo 2.

### Exemplo em XML

O EXEMPLO 5-2 apresenta uma estrutura de metadados extraída da imagem DICOM, na qual é a descrição de uma imagem (Laudo de Ecocardiograma), referente ao do esquema desenvolvido neste trabalho, segundo a sintaxe da XML.

```
<?xml version="1.0" standalone="yes"?>
<?xml-stylesheet type='text/css' href='dicom.css' ?>
<!--Esse é um exemplo de tags XML >
<documentodicom>
  <informacoesgerais>
    <nomesecao>Informacoes Gerais</nomesecao>
    <specificcharacterset>
      <nome>Specific Character Set</nome>
      <valor>ISO_IR 100</valor>
    </specificcharacterset>
    <imagetype>
      <nome>Image Type</nome>
      <valor>ORIGINAL\PRIMARY\OTHER\M\PCA</valor>
    </imagetype>
    <instancecreationdate>
      <nome>Instance Creation Date</nome>
      <valor>19960823</valor>
    </instancecreationdate>
    <instancecreationtime>
      <nome>Instance Creation Time</nome>
      <valor>093811</valor>
    </instancecreationtime>
    <instancecreatoruid>
      <nome>Instance Creator UID</nome>
      <valor>1.3.46.670589.11.0.5</valor>
    </instancecreatoruid>
    <sopclassuid>
      <nome>SOP Class UID</nome>
      <valor>1.2.840.10008.5.1.4.1.1.4 </valor>
    </sopclassuid>
    <sopinstanceuid>
      <nome>SOP Instance UID</nome>
```

```

        <valor>1.3.46.670589.11.0.4.1996082307380009 </valor>
</sopinstanceuid>
<studydate>
    <nome>Study Date</nome>
    <valor>19941122</valor>
</studydate>
<seriesdate>
    <nome>Series Date</nome>
    <valor>19941122</valor>
</seriesdate>
<studytime>
    <nome>Study Time</nome>
    <valor>193438.7200 </valor>
</studytime>
<seriestime>
    <nome>Series Time</nome>
    <valor>202251.2600 </valor>
</seriestime>
<accessionnumber>
    <nome>Accession Number</nome>
    <valor/>
</accessionnumber>
<modality>
    <nome>Modality</nome>
    <valor>MR</valor>
</modality>
<manufacturer>
    <nome>Manufacturer</nome>
    <valor>Philips </valor>
</manufacturer>
<institutionname>
    <nome>Institution Name</nome>
    <valor>NUMATA NS HD HP.</valor>
</institutionname>
<referringphysiciansname>
    <nome>Referring Physician's Name</nome>
    <valor/>
</referringphysiciansname>
<studydescription>
    <nome>Study Description</nome>
    <valor>BRAIN </valor>
</studydescription>
<seriesdescription>
    <nome>Series Description</nome>
    <valor>Angio MIP SAG </valor>
</seriesdescription>
<manufacturersmodelname>
    <nome>Manufacturer's Model Name</nome>
    <valor>T5</valor>
</manufacturersmodelname>

```

```

</informacoesgerais>
<paciente>
  <nomesecao>Paciente</nomesecao>
  <patientsname>
    <nome>Patient's Name</nome>
    <valor>MR/ANGIO/PCA/256</valor>
  </patientsname>
  <patientid>
    <nome>Patient ID</nome>
    <valor>9 </valor>
  </patientid>
  <patientsbirthdate>
    <nome>Patient's Birth Date</nome>
    <valor>19490303</valor>
  </patientsbirthdate>
  <patientssex>
    <nome>Patient's Sex</nome>
    <valor>F </valor>
  </patientssex>
  <patientsweight>
    <nome>Patient's Weight</nome>
    <valor>55</valor>
  </patientsweight>
</paciente>
</documentodicom>

```

#### EXEMPLO 5-2 Exemplo de estrutura de metadados em XML

Vale a pena observar alguns pontos, com relação à sintaxe de definição, lembram muito o documento SGML e HTML. Como vimos, anteriormente, os elementos e atributos estão incluídos, bem como os comentários e outros textos padrões. Um recurso muito forte em comparação com os seus primos de marcação é a capacidade de aninhamento dentro de outros elementos, sem limite de definição de tags.

## 5.10 Considerações Finais

Os documentos gerados no protótipo, são todos em formato texto, decorrente da utilização do XML, que são formados por dados e marcações (markup), que podem ser consideradas como metadados, representados nesta linguagem por tags e armazenados em lugares específicos pela aplicação.

Para que a XML seja utilizada no âmbito do esquema desenvolvido nesse trabalho, faz-se necessário a prévia identificação e definição dos tags a serem criados. O esquema proposto é, portanto, um pré-requisito básico para a utilização dessa linguagem, na descrição de imagens, pois na verdade todas as classes e propriedades, identificadas no esquema, constituem os tags para a criação desse padrão em XML. A descrição desse esquema em XML possibilita, portanto a sua disponibilidade na *Web*, como um padrão para a documentação de imagens. No entanto, para que as imagens

descritas em XML possam ser armazenadas, é necessária a sua veiculação a um repositório das imagens, e ainda, a necessidade de uma aplicação que entenda o suficiente para determinar o que é metadado e a sua respectiva imagem para obtenção das informações ali definidas.

Não conhecemos outro trabalho proposto na literatura que tenha as características do sistema apresentado, e introduza as mesmas funcionalidades, propostas visando facilitar a interoperabilidade de sistemas, e a traço de informações sobre imagens DICOM através da *Web*.

## 6 Conclusões

Este trabalho aborda o problema de representação das imagens DICOM geradas dos exames de pacientes de forma a facilitar o compartilhamento destas informações através da *Web*, e a troca de dados em sistemas de informações na área da saúde. Para tal, foi proposto um modelo com intuito de incorporar aspectos de metadados no projeto, utilizando-se das funcionalidades oferecidas pelo padrão de transferência de imagens DICOM.

Porém, para o uso efetivo do protótipo, faz-se necessária à definição de uma interface que possibilite utilizar a tecnologia disponível. Assim, o estudo sobre os principais projetos deste enfoque revelou que a forma atual de interação entre a tecnologia e a utilização de imagem de paciente, no prontuário, não provêem os requisitos necessários para suportar a flexibilidade exigida. A abordagem descrita neste trabalho representa uma tentativa de resolução deste problema. Neste sentido, os princípios foram aplicados, possibilitando a definição de uma interface personalizada para o usuário.

De modo a ampliar o poder de pesquisa das interfaces, procurou-se estudar os padrões de metadados voltados para imagens digitais. Estes descritores contemplam que podem proporcionar uma nova expressão, utilizando o padrão consagrado para imagens médicas DICOM para um formato Dublin Core, no qual apresenta um total reconhecimento no ambiente *Web*.

### 6.1 Confiabilidade

A implementação de infra-estrutura de segurança é hoje fator imprescindível ao desenvolvimento e implementação de sistemas que permitam criar um registro clínico unificado em formato eletrônico. A abertura das barreiras à informação nas instituições clínicas se pressupõe como requisito fundamental para que exista um trabalho prévio de estruturação interna dos sistemas de informação. Desta forma, é necessário adotar mecanismos seguros de identificação e autenticação dos usuários, impondo o uso de assinatura digital e para garantir a confidencialidade e integridade dos dados. Se estes pontos estiverem assegurados, os sistemas de informação estarão em condições de permitir o acesso da informação remotamente recorrendo de estruturas de confiabilidade (serviços de autenticação), permitindo exclusividade.

Como o mecanismo aqui proposto não está provido de sistemas complexos de segurança, seria interessante agregar à ferramenta uma estrutura de autenticação, permitindo a sua integração entre uma rede pública e hospital, como é comentada no item 1.3.3.3. Desta forma o acesso remoto compartilharia a informação clínica, sobre os sistemas existentes e de forma segura, recorrendo às tecnologias (Certificados Digitais, Smart Cards etc...) para poder acessar Internets/Intranets ou instituições no exterior.

Outro ponto fundamental da aplicação é a forma como a estrutura interna da ferramenta se comporta. Por sua vez, esta estrutura merece uma reformulação no seu conteúdo, pois a sua funcionalidade em relação ao seu resultado foi de maior

importância. Não quer dizer que não foi implementado, mas tecnicamente ela proporciona uma estrutura interna para um estudo futuro mais detalhado desse tema.

## 6.2 Contribuições

São as seguintes as contribuições deste trabalho:

- estudo dos principais sistemas digitais, desenvolvidos no âmbito nacional e internacional;
- desenvolvimento de um modelo, permitindo a criação de interfaces personalizadas para o usuário, que aparentemente é inovador;
- estudo dos padrões e formatos de metadados relacionados com imagens digitais médicas, foram verificados que esses podem dar suporte a funções de descoberta, localização do recurso e descrição de conteúdo;
- desenvolvimento de um modelo de descritores para possibilitar uma gama diversa de pesquisas a um ambiente com um conjunto de informações mais abrangente;
- como consequência do uso do modelo de descritores, foi observada a possibilidade de criação de uma nova categoria de descritores, partindo do Dublin Core denominada de pesquisa adicional, a qual pode ser utilizada para verificar todas as informações complementares a uma imagem. Não temos notícias disto ter sido proposto na literatura antes;
- desenvolvimento de um protótipo contemplando os modelos de interface e metadados criados, voltado para a aplicação de imagens médicas;
- estender o modelo utilizando mais descritores em conformidade com o W3C para descrever conteúdo das imagens geradas do padrão DICOM.

## 6.3 Trabalhos Futuros

O objetivo deste trabalho foi o de verificar as funcionalidades da RDF como suporte à representação de imagens DICOM tornando a sua representação mais Semântica para a *WEB*. Isso foi feito seguindo um caminho por passos. O primeiro deles foi conhecer um pouco mais do significado de cada atributo da imagem, sua arquitetura, objetivos, aplicações etc., descobrindo qual o papel da RDF em todo esse contexto. Depois disso, se fez necessário ter uma noção da sintaxe da RDF. Para tal foi consultado o W3C que tem concentrado esforços em padronizar tudo o que é concernente à *Web*. Através do W3C buscou-se acesso à especificação da linguagem RDF, cujo status é o de recomendação, e também à especificação da RDF Schema, atualmente candidata à recomendação.

Para adequação ao estado-da-arte, este trabalho procurou conhecer um pouco mais sobre o que já foi feito no mundo com respeito à representação do conteúdo das imagens DICOM, particularmente as *tags* contidas na própria imagem.

Na fase de implementação, o trabalho modelou um cenário para uma aplicação de forma que se pudesse entender mais claramente a sintaxe de DICOM e RDF Schema. Para isso foi feito um modelo em forma de remapeamento, prototipando seu código RDF/XML e realizando consultas via aplicação *Web* para o cenário assim estabelecido. O software utilizado, na verdade, ainda esta em fase de construção e testes e, por isso, possui grande limitação.

Apesar da RDF ser uma recomendação do W3C e haver muitos esforços concentrados em sua evolução, há ainda muito por se fazer para que ela esteja consolidada. Soluções práticas incluem:

- derivação de RDF a partir de fontes XML;
- emergência de “motores” e de banco de dados em RDF; e
- interfaces com o usuário (gerais ou específicas) para dados RDF, implementando dessa forma em novos web sites
- representação de conteúdo de imagens médicas. Estas soluções ficam aqui consideradas como sugestões para trabalhos futuros.

As linguagens para dar semântica a *Web*, em geral, ainda são pouco conhecidas da comunidade afim, até porque se tratam de linguagens muito novas, recentemente tornadas como padrão ou ainda em vias disso. Quanto às referências bibliográficas, ainda existem poucos textos em português, a maioria das referências encontrada foi em língua inglesa.

Deve-se considerar ainda que foram tratados aqui somente os níveis de dados e de esquema de arquiteturas de metadados das imagens DICOM. Esta arquitetura prevê ainda um nível lógico em relação com as imagens sendo que a sua representação para RDF Schema envolve mais um esforço para agregar mais semântica, e como consequência a sua visualização. Este esforço irá agregar futuramente mecanismos de busca, promovendo assim a credibilidade da *Web*. Juntamente com a infra-estrutura básica promovida pela HTML e XML e juntamente com uma fonte de riqueza das informações DICOM, a representação da informação na *Web* certamente promoverá novos conceitos e técnicas mais eficientes. Para atingir tais contribuições será necessário desenvolver algumas ações para atingir tal propósito, como:

- verificar e acompanhar o processo da rotina médica ocasionada pela extração de dados do conteúdo das imagens digitais do paciente, em relação à área da Ciência da Informação;
- verificar a prática de utilização da linguagem de marcação XML [LIG 1999] para a troca de informações em ambientes distribuídos, caso seja previsto o armazenamento das imagens em um banco de dados;

- desenvolver de um “Modelo Estendido” para a criação de modelos de descritores para outros sistemas de imagens digitais, que contenham formatos diferentes;
- criar uma estrutura modular para integrar dados cadastrais em conformidade com o paciente;
- proporcionar um serviço integrado em servidores distribuídos para reduzir a carga entre eles, diminuindo o custo de acesso;
- desenvolver um modelo estendido com a integração dos descritores na estrutura relacional, ou seja, criar atributos do modelo para utilização em banco de dados [FRE 96].

O escopo de imagens médicas se estende além de imagens radiológicas. Endoscopia, patologistas, dentistas, e dermatologistas produzem imagens como parte de sua prática. Recentemente, representantes das sociedades profissionais destes grupos têm se encontrado com um Grupo de Trabalho especial do ACR-NEMA para planejar como eles podem aproveitar o trabalho do DICOM. A primeira tarefa empreendida, é a definição do objeto de informação para formação de imagens endoscópicas (incluindo pulmonar, gastrointestinal, geniturinário, e requerimentos de ortopedia).

O Padrão ACR-NEMA DICOM tem sido desenvolvido para atender as necessidades de fabricantes e usuários de equipamentos de imagens (na área médica) na interconexão de dispositivos padronizadas. Suas diversas partes fornecem meios para expansão e atualização, e o projeto do padrão foi direcionado para permitir um desenvolvimento simplificado de todos os tipos de imagens médicas. As futuras adições ao DICOM incluem suporte para criação de arquivos em mídia removível (tal como discos óticos ou fita magnética de alta capacidade) e uma nova estrutura de dados para angiografia e imagens dinâmicas. A versão atual de DICOM tem mostrado um esforço para expandir as comunicações e implementações de diferentes fabricantes usando tanto objetos de informação do DICOM quanto os serviços que eles suportam.

Para resolver a necessidade de criar padrões de troca de mídia o ACR-NEMA reativou a comunidade que fabrica equipamentos, cuja missão é desenvolver mecanismos padronizados e desenvolvendo a parte 10 do DICOM com o propósito de descrever funcionalidades ao armazenamento de mídia e formatos de arquivo. Esta parte é um conjunto de definições de mídia independente. Ela descreve um arquivo genérico e uma estrutura de dados, e em analogia com a camada superior da parte 8, especificando um arquivo básico do DICOM.

Sem dúvida, o DICOM é o maior projeto de padrões de imagens médicas empreendido por indústria e sociedades profissionais. Ele é um padrão complexo por causa do tamanho de seu conteúdo, mas isto é implementável e útil. O padrão oferece um equilíbrio para suportar uma implementação rápida assegurando uma capacidade para desenvolver e responder as necessidades de representar dados textuais em imagens estáticas, ficando um estudo mais aprofundado no futuro para imagens dinâmicas criadas em equipamentos para este propósito.

O estudo focalizado neste trabalho esta totalmente voltado a especificação em tratar imagens estáticas, para que futuramente possa ser inserido em um contexto maior,

para a sua completude. Ainda como sugestão de trabalhos futuros, vale considerar que o estudo merece mais ênfase nos diversos níveis de sua arquitetura, complementando a abordagem aqui tratada.

## Anexo 1 Formatos para Imagens Médicas

### Formatos

Existem vários formatos de arquivo de imagem dentre os quais os mais conhecidos e utilizados são: JPEG (Join Photographic Experts Group), GIF (Graphics Interchange Format), BMP (Microsoft Windows Bitmap) DICOM (Digital Imaging and Communications in Medicine), TIFF (Tagged Image File Format) e PNG (Portable Networks Graphics)[BOU 1997,WIG 2001 , PNG 97].

**JPEG:** padrão desenvolvido pelo Joint Photographic Experts Group para compressão de imagens estáticas, armazena a cor de todos os pixels através de um método de compressão com perda, significando que, ao ser gravada uma imagem nesse formato, sempre se perde alguma informação. O tamanho do arquivo JPEG resultante é muito pequeno, porém os dados que contém são sempre uma aproximação da imagem original. JFIF (JPEG File Interchange Format) é um formato de armazenamento de arquivo para imagens comprimidas com o algoritmo JPEG.

**GIF:** criado pela CompuServe, armazena imagens usando codificação indexada, onde cada pixel é representado por um índice de inteiro dentro de uma paleta de cores. O tamanho da paleta é limitado a 256 cores, fazendo com que o GIF seja mais adequado para imagens que não contenham muitas gradações finas de cores. GIF usa uma técnica de compressão com pouca perda que pode resultar em arquivos muito pequenos.

**BMP:** criado pela Microsoft, suporta até 24 bits para cor e as imagens são mais frequentemente armazenadas sem compressão, resultando em arquivos de tamanho grande.

**TIFF:** originalmente desenvolvido pela Aldus Corporation, suporta até 24 bits para cor e pode ser comprimido com algoritmos como: LZW ou JPEG.

Dois desses formatos mais utilizados evoluíram para novos tipos, apresentando características importantes para esse estudo. Os dois novos formatos são:

**PNG** (Portable Network Graphics), que é sucessor do formato GIF (Graphics Interchange Format);

**SPIFF** (Still Picture Interchange File Format), que é o substituto oficial para o formato de arquivo JFIF (JPEG File Interchange Format) para armazenamento de dados JPEG [SPI 95].

**DICOM:** criado para com a necessidade de agregar tudo em um arquivo (imagem e metadados) de modo a estar de acordo com o sistema PACS desenvolvido e com o protocolo de comunicação de redes. É uma padronização de protocolos de redes, usando convenções tais como o ISO/OSI. Para esse fim, o (NEMA) e o (ACR) cooperaram para criar um formato padrão de dados de imagens, chamado ACR-NEMA que foi o primeiro passo para conseguir a adesão dos fabricantes de equipamentos de imagem digital, facilitando a sua interconexão direta em rede. Atualmente o ACR-NEMA evoluiu para um novo padrão, amplamente adotado, chamado DICOM 3.0 [MAR 98].

### Principais características DICOM:

- intercâmbio de objetos em redes de comunicação em meios de armazenamento através de protocolos de serviços, mantendo independência da rede e do meio físico de armazenamento. Todos os comandos são definidos por uma sintaxe e uma semântica que associa os dados;
- especificação de diferentes níveis de compatibilidade. Explicitamente se define um determinado nível de compatibilidade, para responder somente opções específicas de DICOM;
- informação explícita de objetos através de estruturas de dados. Os objetos são unicamente imagens digitais gráficas;
- identidade dos objetos em forma única, como instâncias com operações permitidas e definidas através de classes;
- flexibilidade em definir novos serviços;
- interoperabilidade entre serviços e aplicações através de uma configuração definida pelo usuário, mantendo uma comunicação eficiente entre serviços;
- representação de aspectos do mundo real, utilizando objetos compostos que descrevem um contexto completo ou seja, objetos normalizados como uma entidade do mundo real;
- seguem as diretrizes do ISO e a estrutura de sua documentação em várias partes, facilitando sua evolução e simplificando a adição de outras.

Os três formatos de arquivo de imagem que serão apresentados a seguir, têm como característica principal de relevância dentro do escopo desse estudo, a possibilidade de embutir metadados no cabeçalho do arquivo.

### **PNG**

PNG (Portable Network Graphics) [ROE 97] é um formato de arquivo de imagem que pretende ser não só o sucessor do formato GIF, como também o mais extensível e livre. Os dois formatos PNG e GIF possuem uma característica importante que é a possibilidade de os dados textuais poderem ser armazenados dentro do arquivo da imagem, o que significa uma forma de armazenar metadados.

Os objetivos para esse formato de imagem são [ROE 97]: portabilidade, não somente através de plataformas e sistemas operacionais, mas também através de softwares; compressão sem perda; compatível com a rede na crescente era da WWW; criação de um novo formato gráfico capaz de substituir o GIF.

PNG, endossado formalmente como padrão pelo W3C (*World Wide Web Consortium*) em 1996, é um formato livre de patentes.

Quando a especificação PNG foi criada, multi-imagens GIFs eram uma curiosidade, e animações estavam completamente excluídas. Foi decidido que PNG seria um formato em imagem simples. Um arquivo PNG pode armazenar texto no cabeçalho do arquivo imagem. Palavras-chave são usadas para indicar o que cada string texto representa.

Um arquivo PNG consiste de uma série de pedaços com as seguintes funcionalidades: assinatura PNG; informações de layout; convenções de nomeação; algoritmo CRC (Cyclic Redundancy Check). Esse arquivo é sempre iniciado por um pedaço IHDR (cabeçalho da imagem) e terminado por um pedaço IEND (fim dos dados PNG).

Especificações dos pedaços:

- Pedaços críticos

IHDR (cabeçalho da imagem): largura, altura, tamanho em bits, tipo de cor, método de compressão, método de filtragem, método de interlace; PLTE (paleta); IDAT (dados da imagem): contém os dados de saída do algoritmo de compressão; IEND: marca o fim dos dados PNG.

- Pedaços auxiliares (opcional)

cHRM (cromaticidade primária e ponto branco); gAMA (imagem gamma); sBIT (bits significantes); bKGD (cor do fundo); hIST (histograma da imagem): frequência usada de cada cor na paleta de cores; tRNS (transparência): valores em cores RGB (red, green, blue); pHYs (dimensões físicas em pixels); tIME (data da última modificação da imagem); tEXt (dados textuais): palavra-chave, separador nulo, texto, título, autor, descrição, direitos autorais, data da criação, software usado para criação da imagem, responsável legal, aviso de natureza de conteúdo, fonte (dispositivo usado para criação da imagem), comentário; zTXt (dados textuais comprimidos): recomendado para blocos de textos grandes contendo: palavra-chave, separador nulo, método de compressão e texto comprimido.

- Pedaço de tipos adicionais

Definido como PNG-EXTENSIONS. Podem ser criados tipos adicionais nesse pedaço, caso os tipos existentes nos pedaços anteriores não sejam satisfatórios à necessidade do usuário.

## DICOM

No DICOM as características do formato estão implícitas no cabeçalho constituído de estruturas de valores de referência. É mostrado no QUADRO 6.1 os atributos contidos no cabeçalho do arquivo DICOM [HOE 1996].

Atributo	Tag	Tipo	Descrição do Atributo
Nome Paciente	(0010,0010)	2	Nome completo
Paciente ID	(0010,0020)	2	Código no sistema
Data nascimento	(0010,0030)	2	Data de nascimento
Sexo	(0010,0040)	2	Sexo M,F,O
Altura Paciente	(0010,1020)	3	Altura em metros
Peso Paciente	(0010,1030)	3	Peso em Kilograma

QUADRO 6.1 - Valores de referência de tags DICOM

### Análise Comparativa dos Formatos

A TABELA 6.1 apresenta, de forma sintetizada, algumas características existentes nos formatos PNG, GIF, DICOM [PNG 97, BOU 97, SPI 95, MAR 98, MUR 96].

No cabeçalho da imagem dos formatos PNG e SPIFF, existem informações técnicas com relação à imagem. No PNG, o pedaço, denominado de cabeçalho da imagem (IHDR), contém somente informações de largura, altura, tamanho em bits, tipo de cor, método de compressão, método de filtragem e método de interlace; outros tipos de informação são colocados em pedaços separados para cada um deles, e a informação textual é armazenada no pedaço padrão texto (tEXt) com palavras-chave adequadas, podendo ser múltiplo.

TABELA 6.1 - Características técnicas dos formatos de arquivo.

Característica	PNG	GIF	DICOM
Imagens coloridas indexadas de até 256 cores	x	x	x
Transparência	x	x	x
Display progressivo	x	x	
Informação auxiliar: comentários textuais e outros dados podem ser armazenados dentro do arquivo da imagem	x	x	
Independência completa de hardware e software	x	x	
Compressão sem perda	x	x (pouca perda)	

Característica	PNG	GIF	DICOM
Imagens múltiplas em um arquivo (para animação)		x	x
Imagens em cores reais, até 48 bits por pixel	x		
Imagens em escala cinza até 16 bits por pixel	x		
Canal alpha completo para transparência geral	x		
Informação gamma da imagem, a qual suporta display automático da imagem com correção de brilho/contraste	x		
Checagem automática de integridade de arquivo e detecção de erros de transmissão	x		
Completamente livre de patentes	x		x
Toolkit implementador padrão disponível	x		x
Benchmarks padrões disponíveis, pontos de referências p/ comparação	x		
Padrão Internet W3C	x		

Entretanto, o formato PNG possui vantagens técnicas a mais do que o DICOM, conforme podem ser observadas na TABELA 6.1. Já existe uma proposta de adição à família PNG, denominado MNG (“Multiple-image Network Graphics”), objetivando armazenar e transmitir animações de múltiplas imagens e composição de quadros (“frames”).

### **O Estudo da Imagem em Diferentes Abordagens**

Imagem digital é uma fonte rica e subjetiva de informação. Pessoas diferentes extraem significados diferentes da mesma imagem. Existe a necessidade de se tratar de forma sistemática essa informação, para que seus significados possam ser mais facilmente armazenados.

A maioria das metodologias usadas para encontrar alguma interpretação de conteúdo na imagem, é ainda limitadas, devido à dificuldade das pesquisas que não utilizam ou não consideram a semântica da informação em imagens. Tipicamente, estas metodologias dependem de IDs de arquivos, palavras-chave, ou textos associados às imagens, e não permitem consultar diretamente baseado nas propriedades visuais de imagens [FLI 95].

Nesta seção pretende-se mostrar as desigualdades de tratamento da imagem dentro de três diferentes abordagens da área de Ciência da Computação, a saber: Recuperação da Informação - Information Retrieval - em Imagem, Banco de Dados de Imagem.

### **Recuperação da Informação (Information Retrieval - IR) de Imagens**

A abordagem de recuperação da informação (IR) consiste na recuperação de todos os documentos cujas propriedades são similares às apresentadas na consulta [GAR 99]. As técnicas para recuperação podem ser classificadas em duas: técnicas de casamento exato e casamento parcial. A técnica de casamento exato apresenta a desvantagem de não conseguir recuperar documentos cuja representação seja parcialmente, por isso essa técnica de casamento parcial é mais poderosa no que diz respeito à eficácia do resultado das consultas.

Em Information Retrieval, a consulta pode ser melhorada, permitindo ao usuário expressar explicitamente um grau de incerteza, e possibilitando a recuperação de imagens, mesmo existindo um casamento parcial entre a consulta e a imagem existente.

#### Características

Nesta abordagem existe a preocupação de possibilitar ao usuário a especificação de consultas imprecisas, para expressar o conhecimento incerto do conteúdo de imagens a serem pesquisadas.

A tecnologia usada para recuperar imagens, baseadas em informação visual, é chamada de Recuperação de Informação Visual (VIR – Visual Information Retrieval). A consulta por conteúdo transforma a consulta do usuário em primitivas da imagem, que correspondem às propriedades do interior da imagem [SHA 97].

Os componentes da consulta devem ter importâncias diferenciadas pelo usuário e a regra associada ao componente da consulta pode ter diferentes preferências. Os valores de preferência e de importância, combinados com o nível de reconhecimento dos objetos correspondentes nas várias interpretações da imagem, podem ser usados para classificar as imagens recuperadas.

A linguagem de consulta que permite a formulação de consultas imprecisas possibilita também limitar o número de imagens recuperadas. Isto é possível pelo fato das imagens recuperadas serem classificadas em ordem decrescente de relevância. A ordem é dada por uma marcação associada a cada imagem. Esta marcação possibilita a medida do grau de casamento entre cada imagem e a consulta. É obtida através da combinação dos valores de preferência e importância associados a cada predicado na consulta especificada.

A linguagem de consulta adotada pela abordagem de recuperação da informação (QL-IR) possui as seguintes características: não é baseada na lógica booleana; permite limitar o número de imagens a serem recuperadas; permite associar valor de importância a cláusulas de objetos diferentes; permite associar valor de preferência a regras posicionais diferentes.

A função de classificação que mede a relevância de cada imagem recuperada é dependente do grau de reconhecimento de cada objeto presente na imagem, e dos

valores de importância e preferência de cada termo da consulta associada a cada regra posicional verificada.

Sistemas de recuperação de informação visual (VIR – Visual Information Retrieval) são classificados, usando os seguintes critérios [CHA 97]: automação, características multimídia, adaptabilidade, abstração, generalidade, categorização e processamento de domínio comprimido.

Uma característica dos sistemas VIR é o fato de a pesquisa ser aproximada, requerendo uma avaliação de similaridade visual. Os itens, retornados no topo da lista dos resultados da consulta têm maior similaridade em relação a sua entrada. Mas os itens retornados raramente têm um casamento exato com os atributos especificados na consulta. Esses sistemas podem também usar entrada humana direta e outros dados suportados para melhor indexar a informação visual.

### **Banco de Dados de Imagens**

1. A origem de bancos de dados de imagem teve início com a comunidade de interpretação de imagem [GRO 92]. No começo dos anos de 1990, as duas comunidades, a de interpretação de imagem e a de banco de dados convergiram para uma concepção comum devido à aceitação de que imagem e texto deveriam ser tratados igualmente. Imagens devem também ser capazes de serem recuperadas por conteúdo e ser componentes integrais de uma linguagem de consulta.
2. Um banco de dados de imagem é um banco de dados que armazena imagens e suporta tipos de dados de imagem com um conjunto de funções associadas.
3. Quando pensamos em banco de dados de imagem, temos que levar em conta que a natureza de imagens como grandes estruturas de arquivo binárias requerem considerações especiais em termos de suas estruturas, organização e capacidades de armazenamento.

#### 6.3.1.1 Características

O tratamento de dados (mídias) do tipo imagem e vídeo, no ambiente de banco de dados, aconteceu primeiramente na arquitetura relacional e depois na arquitetura orientada a objetos [GRO 97].

Após algumas experiências de trabalho com esses tipos de dados na arquitetura relacional, foi descoberta uma limitação: o não casamento/combinção entre a natureza dos dados e a forma que ambos o usuário e o sistema são obrigados a consultar e operar esses dados. Abordagens de indexação padrão não funcionam para consultas baseadas em conteúdo do dado multimídia.

O gerenciamento de dados multimídia, em ambiente de banco de dados, tem levado a idéias conclusivas como: uma consulta multimídia tem que ser vista de maneira diferente da consulta padrão de banco de dados e mais parecida com a consulta utilizada na abordagem de “information retrieval”. Porém, esse conceito ainda não foi totalmente implementado em banco de dados.

As representações lógica e física de objetos imagens são definidas no modelo de dados do domínio. Vários tipos de informação devem ser capturados no modelo de dados imagem: a estrutura detalhada dos vários objetos imagem; propriedades e operações dependentes de estruturas das imagens; relacionamentos entre imagens, e do mundo real; pedaços de imagens com representação de relacionamentos entre eles e os métodos usados para determiná-las; propriedades, relacionamentos e operações de objetos do mundo real.

O processo de recuperação, em banco de dados, de imagem, deveria possuir a característica de consultas imprecisas, até pelo fato de que, no processo de recuperação da imagem descrições semânticas poderão estar associadas a componentes da imagem e múltiplas interpretações poderão ser geradas, dependendo da descrição [GAR 99].

A abordagem de banco de dados adota uma linguagem de consulta booleana, onde o processamento da consulta é baseado no casamento exato entre a consulta e as imagens recuperadas; além disso, não é possível medir a relevância das imagens recuperadas (não é possível a listagem das imagens de acordo com uma ordem de relevância).

## **Anexo 2 Modelagem Conceitual de Metadados para Imagem**

O esquema aqui apresentado vem complementar as abordagens (Recuperação da Informação - Information Retrieval - em Imagem) e os padrões de metadados considerados no capítulo 7. As abordagens utilizam estratégias diferenciadas para conseguirem efetuar consultas às imagens, mas não se preocupam com uma descrição mais completa das características das mesmas. Os padrões considerados no capítulo 7, foi apenas o DC, por dar alguma conotação a dado multimídia. Porém, este padrão está totalmente adequado à descrição mais eficiente de imagem, pois não viabiliza descritores para todas as suas características importantes.

Este esquema viabiliza uma representação sistemática e detalhada dos elementos descritores de imagens estáticas digitais, incluindo tanto elementos para descrever as informações técnicas da imagem, quanto para descrever as informações de conteúdo semântico, podendo com isso ser utilizado em arquiteturas de metadados existentes.

Além dos estudos das diferentes abordagens que tratam a imagem na área de Ciência da Computação, foram também pesquisadas as formas de descrição de imagens na área de Ciência da Informação, já que a catalogação e indexação de imagem também são objeto de estudo desta área.

Na utilização de imagens médicas se exige que sejam descritas as suas informações, principalmente as menos evidentes: o que caracteriza “quem criou”? e “equipamento utilizado para sua criação”? Como descrever e, sobretudo, analisar imagens para fins de catalogação? Como dar conta deste tipo de material? Com exemplos desta ordem, cabe até questionar se é possível “analisar” as informações relacionadas ao conteúdo semântico das imagens [SMI 87]. A partir desses questionamentos, ocorre uma primeira divisão entre as opiniões: considera-se que a análise da imagem não tem nada de específico e que as boas e comprovadas técnicas da análise documentária resolvem perfeitamente a questão; ou parte-se do princípio de que as técnicas de análise de documentos escritas não são apropriadas para criação de atributos para imagens.

Existe um problema entre a descrição e a interpretação, uma vez que a descrição da imagem não traduz o código visual verbal (o que ele diz), criando condições para a sua interpretação. Uma dificuldade é a separação entre a denotação (o que a imagem mostra) e a conotação (o que o leigo e um especialista vêem na imagem).

### **Visão Segundo a Ciência da Informação**

A representação das imagens em Ciência da Informação, é analisada também, estabelecendo níveis para a análise da imagem que ajudam a identificar a questão. Alguns mostram uma visão sob diferentes níveis de detalhes para analisar a imagem [GAR 99]:

Pré-iconográfico: são descritos, genericamente, os objetos e ações representadas pela imagem;

Iconográfico: estabelece o assunto secundário ou convencional ilustrado pela imagem. Trata-se da determinação do significado abstrato ou simbólico da imagem, sintetizando de seus elementos componentes, detectados pela análise pré-iconográfica;

Iconológico: propõe uma interpretação do significado intrínseco do conteúdo da imagem. A análise iconológica constrói-se a partir das anteriores, mas recebe fortes influências do conhecimento do analista sobre o ambiente cultural, artístico e social no qual a imagem foi gerada.

Os metadados podem ser definidos em nível de pixel, nível semântico e nível de domínio (informações específicas de uma determinada aplicação). Os metadados escolhidos para caracterizarem uma dada imagem podem então ser modelados através do Esquema Conceitual de Metadados para Imagem, proposto e descrito ao longo deste capítulo.

### **Características**

O Esquema Conceitual de Metadados para Documentação e Recuperação de Imagens apresentadas, tem como objetivo principal suportar as seguintes características: simplicidade e adaptabilidade a qualquer domínio de aplicação. Este esquema possui classes as quais contêm propriedades referentes às características da imagem.

Nesta proposta, considerando um Objeto Imagem, uma imagem de origem médica (MR, CT, RX, etc...), sabe-se que possuem classes de descritores para o objeto imagem, na qual foram identificadas três classes principais: Objeto Imagem, Conteúdo Semântico e Tipo de Relacionamento (existente entre os componentes da imagem). Cada uma dessas classes associa uma gama de informações com objetivos específicos, a saber:

objeto Imagem: representa em sua maioria um conjunto de propriedades consideradas informações técnicas, relativas ao objeto imagem (imagem digitalizada);

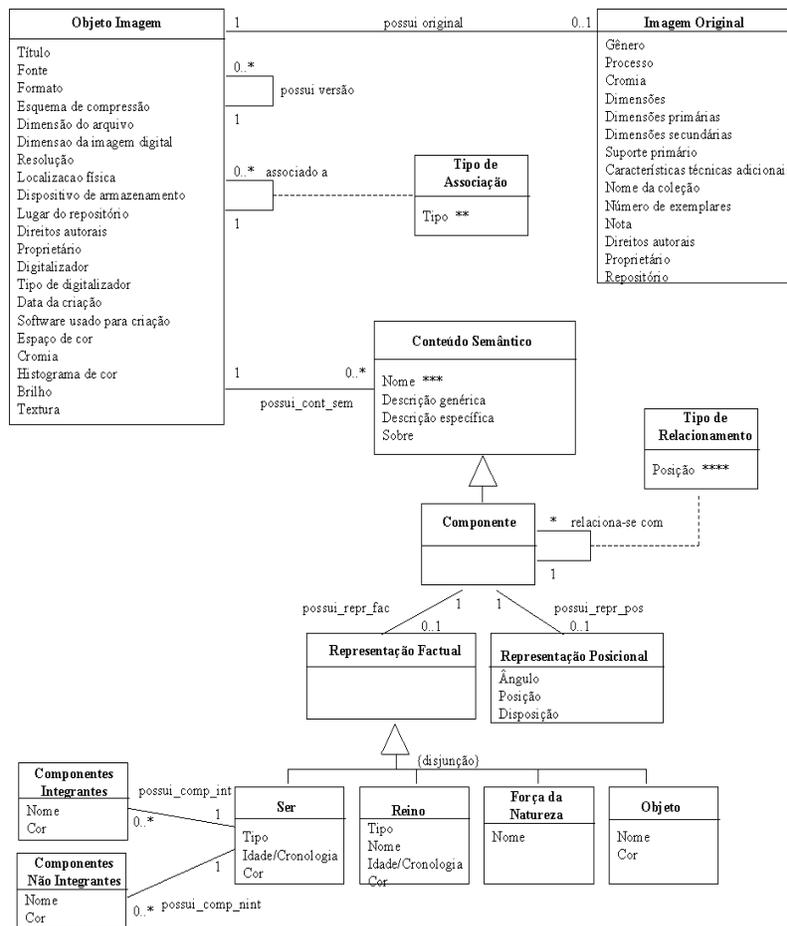
conteúdo Semântico: representa o conjunto de categorias que descrevem o conteúdo semântico do objeto imagem, possibilitando identificar os componentes existentes na imagem;

tipo de Relacionamento: representa o tipo de relacionamento existente entre os componentes da imagem, aspecto importante e necessário para melhor descrever o contexto da informação.

### **Especificação de Classes**

O esquema proposto pode ser visualizado através da FIGURA 6-1, utilizando a linguagem de modelagem de objetos - UML (Unified Modeling Language) [FUR 98]. A descrição das classes e suas propriedades são vistas a seguir:

Este esquema conceitual viabiliza a documentação de imagem estática, descrevendo suas propriedades, possibilitando também descrever os componentes contidos na imagem e os relacionamentos existentes entre eles. O esquema permite também identificar os objetos imagem que são versões de um objeto imagem e os tipos de associações existentes entre os mesmos.



Onde:

\*\**. Tipo em Tipo de Associação* pode assumir os seguintes valores: composto de, parte de, derivado de;

\*\*\**. Nome em Conteúdo Semântico* pode assumir somente os seguintes valores: Ação (O Que), Espaço (Onde), Tempo (Quando), Componente (Quem) e Finalidade (Para Que);

\*\*\*\**. Posição em Tipo de Relacionamento* pode assumir valores como: lado direito, lado esquerdo, frente, atrás, abaixo, sobreposto, diagonal frente direita, diagonal frente esquerda, diagonal atrás direita, diagonal atrás esquerda, centro, etc;

FIGURA 6-1 - Esquema Conceitual de metadados para Imagem

## Objeto Imagem

Um Objeto Imagem representa a imagem digitalizada descrita pelo esquema. As propriedades contidas nessa classe representam um conjunto de características da imagem digital, consideradas informações técnicas nas quais são representadas pelas seguintes propriedades (atributos):

- **Título:** nomeação dada a imagem digital para facilitar a sua identificação;
- **Fonte:** local de origem da imagem digital; informação sobre a fonte individual, agência ou do repositório de onde a imagem foi obtida;
- **Formato:** formato do arquivo da imagem (como por exemplo: JPEG, GIF, PNG, SPIFF, etc);

- Esquema de compressão: tipo de compressão utilizado na criação da imagem digital; como por exemplo: JBIG, LZW (Abraham Lempel, Jakob Ziv and Terry Welch) etc;
- Dimensão do arquivo: tamanho do arquivo em bytes;
- Dimensão da imagem digital: tamanho da imagem em pixels, polegadas, centímetros e/ou milímetros; campo composto contendo o valor e o tipo de unidade utilizada para a descrição da dimensão da imagem digital;
- Resolução: número de unidades do tipo de medida (pixels/bits/etc) utilizadas na descrição da resolução da imagem digital;
- Localização física: local onde a imagem se encontra fisicamente armazenada; pode ser um diretório em uma rede ou uma URL;
- Dispositivo de armazenamento: dispositivo físico onde a imagem está armazenada;
- Lugar do repositório: localização geográfica do repositório;
- Direitos autorais: nome ou identificação do autor ou autores da imagem digital, no caso da imagem não possuir foto original, isto é, se ela tiver sido criada diretamente na forma digital;
- Proprietário: nome ou identificação do proprietário da imagem digital;
- Digitalizador: nome ou identificação da pessoa que digitalizou, no caso de não ter sido criada diretamente na forma digital;
- Tipo de digitalizador: aparelho utilizado para criação da imagem digitalizada;
- Data da criação: data associada à criação ou produção da imagem digital;
- Software usado para criação: software usado na criação da imagem digital;
- Espaço de cor: tipo padrão de cores usadas (por exemplo: RGB (Red, Green, Blue), HSV (Hue, Saturation, Value), CMYK (Cyan, Magenta, Yellow, Black) etc);
- Cromia: cromia da imagem digital, por exemplo: preto e branco, colorida;
- Histograma de cor: distribuição das cores existentes na imagem digital em cada pixel. Suponha que a imagem tenha N cores; a idéia é responder a “quantos pixels a cor X tem associados?”, isto é, contar os pixels referentes a uma dada cor.
- Textura:

- Brilho: proporcional à integral do produto da curva e à função de eficiência de luminosidade, esse cálculo é feito automaticamente por meio de um algoritmo, gerando um vetor.

### **Imagem Original**

A classe Imagem Original contém informações sobre a imagem original (fotografia, pintura, gravura, etc) que passou por processo de digitalização, gerando assim o Objeto Imagem; está sendo representada pelas propriedades (atributos):

Gênero: descrição genérica da imagem original, como por exemplo: fotografia, slide, pintura, gravura etc.

Processo: no caso da fotografia representa a designação específica, isto é, o tipo de processo fotográfico utilizado na produção da foto, com relação aos outros gêneros (pintura, gravura, etc) de imagem original também poderá ser descrito o respectivo processo.

Cromia: cromia da imagem, tal como, por exemplo: preto e branco, colorida.

Dimensões: dimensões da imagem original.

Dimensões primárias: dimensões do suporte primário (papel fotográfico) onde está a imagem.

Dimensões secundárias: dimensões do suporte secundário (moldura) onde está a imagem.

Suporte primário: tipo de suporte primário onde está a foto, os tipos de suporte podem ser: papel, couro, porcelana, tecido, madeira.

Características técnicas adicionais: características técnicas adicionais utilizadas no processo de produção da imagem original, como por exemplo, no caso de fotografia as características podem ser: solarizada, foto montagem etc.

Nome da coleção: no caso da imagem original pertencer a alguma coleção.

Número de exemplares: número de exemplares daquela imagem original existente no acervo.

Nota: indicação da página ou número correspondente para acesso, como por exemplo, no caso de uma foto integrar um álbum.

Direitos autorais: nome ou identificação do autor ou autores da imagem original.

Proprietário: nome ou identificação do proprietário da imagem original.

Repositório: local onde se encontra a imagem original.

### **Tipo de Associação**

O Tipo de Associação representa a associação existente entre os objetos imagem, de acordo com o campo:

Tipo: tipo de associação existente entre objetos imagem. Pode assumir os seguintes valores: composto de, parte de, derivado de. Aqui pode ser representado se um objeto imagem é composto de outros objetos imagem, se ele é parte de um ou se ele foi derivado de outro.

### **Conteúdo Semântico**

A classe Conteúdo Semântico representa informações referentes ao conteúdo semântico da imagem, contendo a descrição desta classe as seguintes propriedades:

- Nome: nome da categoria que estará sendo descrita:
  - Ação: o que os objetos e seres estão fazendo (ações, eventos, emoções).
  - Espaço: onde está a imagem no espaço.
  - Tempo: tempo linear ou cíclico de datas e períodos específicos, tempos recorrentes.
  - Componente: esta categoria é transformada em classe já que possui relacionamentos com outras classes e esses relacionamentos só existem no caso do nome da categoria ser Componente, descrita mais adiante em componente.
  - Finalidade: objetivo de utilização ou armazenamento daquela imagem digital, para que a imagem está sendo armazenada; como por exemplo: “imagem a ser utilizada no cenário de um diagnóstico médico”.
- Descrição genérica: descrição genérica daquela categoria de conteúdo semântico.
- Descrição específica: descrição específica de conteúdo semântico; (De quem, Onde, Quando, Que, para que especificamente se trata a imagem?).
- Sobre: sobre qual assunto a imagem trata de acordo com a categoria (Nome) que está sendo descrita; (representa a manifestação de alguma abstração?).

### **Componente**

A classe Componente identifica os componentes da imagem. Podem ser animados e inanimados, objetos e seres concretos. Componente é uma categoria da classe conteúdo semântico, que foi transformada em uma classe, porque possui relacionamentos com outras, conforme pode ser observado no esquema especificado na TABELA 6.2, mostra exemplos de descrições, utilizando a classe Componente:

TABELA 6.2 - Descrição da Classe

<i>Nome</i> = Componente	<i>Nome</i> = Ecocardiograma
<i>Descrição genérica</i> = Menino	<i>Descrição genérica</i> = Exame
<i>Descrição específica</i> = Renato	<i>Descrição específica</i> = Coração
<i>Sobre</i> = Infância	<i>Sobre</i> = Sistema Vascular

### **Tipo de Relacionamento**

Esta classe representa o tipo de relacionamento existente entre os componentes da imagem. Todos os que foram identificados servem para indicar como um componente está espacialmente localizado em relação a outro ou outros, dentro de uma imagem. Esse tipo de relacionamento é representado por meio do seguinte campo:

**Posição:** em que posição um determinado componente está em relação a outro ou a outros. Pode assumir os seguintes valores: lado direito, lado esquerdo, frente, atrás, abaixo, acima, sobreposto, diagonal frente direita, diagonal frente esquerda, diagonal atrás direita, diagonal atrás esquerda, centro etc.

### **Representação Posicional**

A Representação Posicional identifica em que posição está o componente da imagem que pertence a classe Componente, por meio dos seguintes campos:

- **Ângulo:** em que ângulo está o componente; por exemplo: frente, perfil, costas etc.
- **Posição:** em qual posição está o componente.
- **Disposição:** disposição do componente na imagem, como por exemplo: corpo inteiro, não corpo inteiro, mãos, pernas, cabeça até o busto, cabeça até o joelho.

### **Representação Factual**

A Representação Factual identifica o tipo de classificação do componente da imagem, pertencente a classe Componente, sendo esta classificação dos tipos Ser, Reino, Força da Natureza e Objeto. Esses tipos são subclasses da classe representação factual, descrita como:

#### **Ser**

Ser é uma subclasse da classe Representação Factual, identificando em qual tipo de ser está classificado o componente, de acordo com os seguintes campos:

**Tipo:** o tipo de ser do componente; (homem, mulher, fictício).

**Idade/Cronologia:** qual a idade do componente ou em que faixa cronológica ele está; (bebê, criança, adolescente, adulto, idoso etc).

Cor: cor do componente.

### Reino

Reino é uma subclasse da classe Representação Factual, identificando em qual tipo de reino está classificado o componente, de acordo com os seguintes campos:

Tipo: tipo de reino do componente; (reino animal, reino vegetal, reino mineral).

Nome: nome do reino ao qual pertence; por exemplo: cachorro, árvore.

Idade/Cronologia: qual a idade do componente ou em que faixa cronológica ele está; (bebê, criança, adolescente, adulto, idoso etc).

Cor: cor do componente.

### Força da Natureza

Força da Natureza é uma subclasse da classe Representação Factual, identificando se o componente da imagem é uma manifestação energética e outras vinculadas à natureza, por meio do seguinte campo:

Nome: nome da manifestação da natureza.

### Objeto

Objeto é uma subclasse da classe Representação Factual, identificando se o componente da imagem é um artefato ou manufatura, por meio dos seguintes campos:

Nome: nome do artefato ou manufatura;

Cor: cor do componente.

### **Componentes Integrantes**

Componentes Integrantes identificam os componentes considerados integrantes do ser, isto é, os elementos que entram na composição, fazendo parte da caracterização do Ser. Podem ser peças ou acessórios de indumentária, como por exemplo: óculos, chapéu etc.; esses serão descritos de acordo com os seguintes campos:

Nome: nome do componente integrante do ser;

Cor: cor do componente integrante do ser.

### **Componentes Não Integrantes**

Componentes Não Integrantes identificam os componentes considerados não integrantes do ser, como por exemplo, os componentes que o Ser estiver segurando ou estiverem junto a ele, tal como: homem segurando uma bengala. Esses serão descritos de acordo com os seguintes campos:

Nome: nome do componente não integrante do ser.

Cor: cor do componente não integrante do ser.

Todas as propriedades das classes descritas anteriormente, podem ter as seguintes formas de captura, dependendo do método correspondente, utilizado para a obtenção de seus valores:

- Cabeçalho do Arquivo: a captura do valor da propriedade será feita diretamente por meio dos descritores da imagem embutida no próprio cabeçalho do arquivo digital.

- Textual: a captura do valor da propriedade é feita de forma textual, a partir de interação com o usuário, o valor da propriedade será digitado.

- Captura Automática: a captura do valor da propriedade será feita de forma automática, como por exemplo, através de um algoritmo de busca.

A seguir serão descritos exemplos de imagens feitas a partir do modelo conceitual de metadados proposto nesse capítulo.

### **Exemplo Utilizando o Esquema**

A FIGURA 6-2 apresenta o diagrama de instâncias, sendo descrita através da utilização do esquema conceitual de metadados para imagem especificado anteriormente.



FIGURA 6-2 - Exemplo de Imagem Original

Esta imagem, criada já na forma digital (por isso não possui representação na classe Imagem Original), representa uma imagem médica e contém os componentes: veias, artérias, vasos. Podemos dar o nome de Laudo de Ecocardiograma a este objeto imagem, que será descrito através das classes Objeto Imagem e Conteúdo Semântico. Dentro da classe Objeto Imagem, serão descritas as propriedades formato, direitos autorais, histograma de cor e textura. Na classe Conteúdo Semântico, são identificados: os componentes (Ventrículo, Pericárdio, Aorta) existentes na imagem os tipos de categorias de conteúdo semântico, conforme pode ser observado na FIGURA 6-2

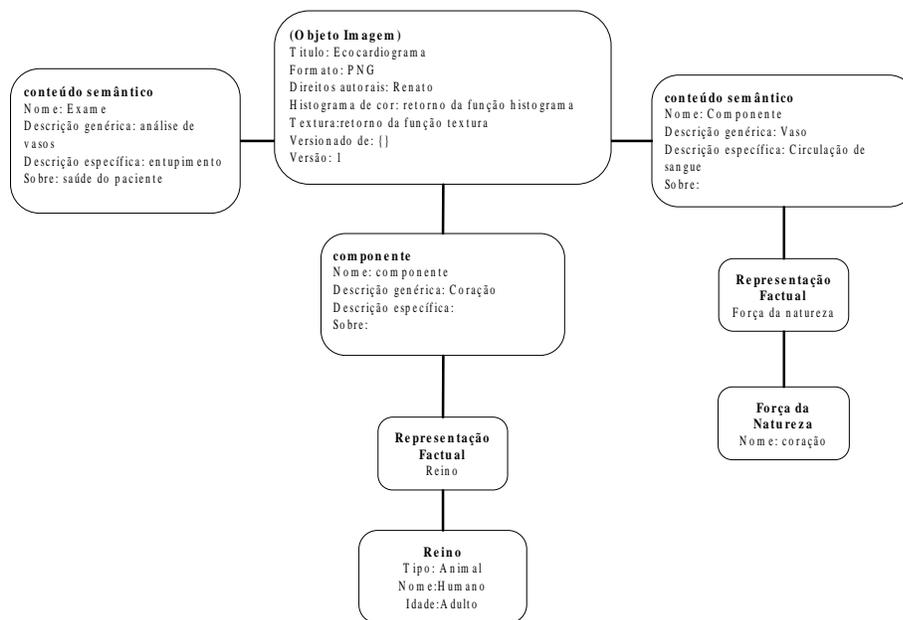


FIGURA 6-3 - Diagrama de Instâncias da figura exemplo FIGURA 6-2

### Considerações Finais

Alguns elementos descritores, especificados nesse esquema podem ser identificados nos padrões de metadados estudados no capítulo 3 e também em sistemas estudados nas abordagens do capítulo 4, por exemplo:

Os elementos DC.title (título), DC.author (autor), entre outros descritores existentes no padrão Dublin Core, são encontrados no esquema aqui apresentado. Com a diferença de que no Dublin Core, existem somente 15 elementos descritores no total, onde somente um deles (DC. description) serve para a descrição de conteúdo semântico, enquanto que, nesse esquema, existe uma forma mais detalhada para o tratamento desse tipo de informação semântica da imagem.

O esquema proposto, nesse capítulo vem viabilizar uma representação sistemática e detalhada dos elementos descritores de imagens estáticas digitais, incluindo tanto elementos para descrever as informações técnicas da imagem quanto para descrever as informações de conteúdo semântico, podendo, com isso, ser utilizado em arquiteturas de metadados existentes.

## Bibliografia

- [ADE 2001] ADEUS ao Software? **Revista Amanhã**, Porto Alegre, jan. 2001. Disponível em: <<http://amanha.terra.com.br/>>. Acesso em: nov. 2001.
- [ARE 2000] ARENSON, Ronald L. et al. Computers imaging and health care: now and in the future. **Journal of Digital Imaging**, Pennsylvania, v.13, n.14, p.145-156, Nov. 2000.
- [BAU 2000] BAUMAN, R.A; GELL, G. The reality of picture archiving and communication systems (PACS): a survey. **Journal of Digital Imaging**, Pennsylvania, v.13, n.4, p. 157-169, Nov. 2000.
- [BID 98] BIDGOOD, W.D. et al. The role of digital imaging and communications in medicine in an evolving healthcare computing environment: the model is the message. **Journal of Digital Imaging**, [S.l.], v.11, n.1, Feb.1998.
- [BOU 97] BOUTELL,T. et al. **PNG (Portable Network Graphics) Specification Version 1.0**. Disponível em: <<ftp://ds.internic.net/rfc/rfc2083.txt>>. Acesso em: mar. 1997.
- [BRU 98] BRUEGEL, R.B. Patient Empowerment - A Trend that Matters. **Journal of American Health Information Management Association**, Sept.1998. Disponível em: <<http://www.ahima.org/journal/features/feature.9809.1.html>>. Acesso em set. 2001.
- [CAI 99] CAIN, M.; MITTMAN, R. The future of the internet in health care: a five-year forecast. **Informatics Review**, [S.l.], June 1999. Written for the California HealthCare Foundation. Disponível em: <<http://www.informatics-review.com/thoughts/future.html>>. Acesso em: jul. 2001.
- [CAS 2001] CASTANETTO, Jesus et al. **Professional PHP Programando**. São Paulo: Makron Books, 2001. 770p.
- [CES 96] CESTA, André Augusto. **A linguagem de programação Java: orientação a objetos**. Campinas: Instituto de Computação, Unicamp, 1996. Disponível em: <<http://www.dcc.unicamp.br/~aacesta>>. Acesso em: fev. 2002.
- [CHA 97] CHANG, Wendy; ZHANG, Aidong. Metadata for Distributed Visual Database Access. In: IEEE METADATA CONFERENCE, 2., 1997, Spring, Maryland. Disponível em: <<http://computer.org/proceedings/meta97/papers/azhang/azhang.html>>. Acesso em: set. 2001.

- [CLA 98] CLARK, Scott. **The WebDeveloper.com Guide to XML**. Disponível em: <[http://www.webdeveloper.com/categories/html/html\\_xml\\_1.html](http://www.webdeveloper.com/categories/html/html_xml_1.html)>. Acesso em: mar. 2001.
- [CLU 2002] CLUNIE, David A. **Medical Image Format FAQ**. START OF PART INDEX. Version: 4.07. Disponível em: <<http://www.dclunie.com/medical-image-faq/html/>>. Acesso em: jan. 2002.
- [DAG 98] DAGHER, A.P.; FITZPATRICK, M.; FLANDERS, A.E. Enhancing Web applications in radiology with Java: estimating MR imaging relaxation times. **Radiographics**, [S.l.], v.18, n.5, p.1287-1293, Sept./Oct. 1998.
- [DCT 99] DCT1, List of Resource Types. Dublin Core Draft Working Group Report, 1999. Disponível em: <<http://purl.org/DC/documents/wd-typelist.htm>>. Acesso em: maio 2001.
- [DEM 96] DEMPSEY, Lorcan; WEIBEL, Stuart L. The Warwick Metadata Workshop: A Framework for the Deployment of Resource Description. Disponível em: <<http://mirrored.ukoln.ac.uk/lis-journals/dlib/dlib/dlib/july96/07weibel.html>>. Acesso em: jun. 2002.
- [DET 2001] DETMER, W. **WWW and the Electronic Medical Record**. Nov. 1994. Medical Informatics Journal Club Section on Medical Informatics Stanford University November 29, 1994. Disponível em: <<http://www.people.virginia.edu/~wmd4n/WWWTalk/WWW-outline.html>>. Acesso em: jul. 2001.
- [DUB 2000] DUBLIN. **Dublin Core Metadata Initiative**. 2000. Disponível em: <<http://purl.org/DC/index.htm>>. Acesso em: jan. 2002.
- [ELB 2000] ELDEIB, Ayman M.; BAHGAT Mohamed M. Standards applied in a Web based Egyptian Asynchronous Telemedicine Project. Disponível em: <[www.cvip.uofl.edu/Papers/ayman1.ps.gz](http://www.cvip.uofl.edu/Papers/ayman1.ps.gz)>. Acesso em jan. 2002.
- [EYS 2001] EYSENBACH, G. What is e-Health? **Journal of Medical Internet Research**, June 2001. Editorial. Disponível em: <<http://www.jmir.org/2001/2/e20/index.htm>>. Acesso em: jul. 2001.
- [FLI 95] FLICKNER, M. et al. Query by Image and Video Content: The QBIC System. **Computer**, Los Alamitos, v.28, n.9, Sept. 1995.
- [FRE 96] FREIRE, Henrique J.P.; PANEPUCCI, H. Banco de dados relacional para tomografia por ressonância magnética nuclear. In: FÓRUM NACIONAL DE CIÊNCIA E TECNOLOGIA EM SAÚDE, 3., 1996, São Paulo. **Anais ...** São Paulo: [s.n.], 1996. p.681-2.

- [FRY 2001] FRYDRYCH, M. **Internet programming**. Disponível em: <<http://www.it.lut.fi/opetus/00-01/010577001/>>. Acesso em: jun. 2001.
- [FUR 98] FURLAN, José Davi. **Modelagem de objetos através da UML: the unified modeling language**. São Paulo: Makron Books, 1998.
- [GAR 99] GARCIA, Simone de Souza. **Metadados para documentação e recuperação de imagens**. 1999. Dissertação (Mestrado em Engenharia) - Instituto Militar de Engenharia, Rio de Janeiro.
- [GRI 2000] GRIMSON, Jane; GRIMSON, William; HASSELBRING, Wilhen. The SI Challenge in Health Care. **Communications of the ACM**, New York, v.43, n. 6, June 2000.
- [GRO 92] GROSKY, W. et al. Research Directions in Image Database Management. In: INTERNATIONAL CONFERENCE ON DATA ENGINEERING, 8., 1992. **Proceedings...** Los Alamitos: IEEE Computer Society, 1992.
- [GRU 99] GRUEN, J. The Physician and the Internet: Observer or Participant? **MD Computing**, [S.l.], v.16, n.6, 1999.
- [GUA 97] GUHA, R.; BRAY, T. **Meta Content Framework Using XML**. 1997. Disponível em: <<http://www.w3.org/TR/NOTE-MCF-XML-970606>>. Acesso em: mar. 2001.
- [HEE 97] HEERY, Rachel et al. Notes from UK participants. In: THE DUBLIN CORE WORKSHOP, 4., 1997, Canberra. **Anais Eletrônicos**. Canberra, 1997. Disponível em: <<http://www.ukoln.ac.uk/metadata/resources/dc4-notes.html>>. Acesso em: mar. 2001.
- [HOE 96] HOEHN, Hervé. **PAPYRUS 3.1 The DICOM compatible image file format**. Working Draft 0.1 - February 1996. Digital Imaging Unit Center of Medical Informatics University Hospital of Geneva. Disponível em: <<http://www.expasy.ch/UIN/html1/projects/papyrus/papyrus.html>>. Acesso em: fev. 2001.
- [HUA 97] HUANG, H. K. Multimedia Application in Health Care. **IEEE Multimedia**, San Francisco, v.4, n.3, p.23, 1997.
- [HUA 98] HUANG, L.J. et al. **A Scalable Architecture for the Storage of Medical Images and Information**. UCLA Dept. Radiological Sciences, 1998. Disponível em: <<http://www.radsci.ucla.edu/>>. Acesso em: mar. 2002.

- [KER 97] KERHERVÉ, Brigitte. Models for Metadata or Metamodels for Data? In: IEEE METADATA CONFERENCE, 2., 1997, Maryland. **Anais eletrônicos...** Maryland: Silver Spring, 1997. Disponível em: <<http://computer.org/proceedings/meta97/papers/bkerherve/bkerherve.html>>. Acesso em: jan. 2002.
- [KHA 97] KHARE, Rohit; RIFKIN, Adam. X Marks the spot: using XML to automate the web. **IEEE Internet Computing**, [S.l.], v.1, n.4, p.78-87, July/Aug. 1997. Disponível em: <<http://www.cs.caltech.edu/~adam/papers/xml/x-marks-the-spot.html>>. Acesso em: jan. 2002.
- [KIL 2000] KILBRIDGE, P.M. E-Healthcare: Urging Providers to Embrace the Web. **MD Computing**, [S.l.], v.17, n.1, 2000.
- [KOH 99] KOHN, D. Preparing the Environment for Internet-Derived Technologies. In: MURPHY, G.F.; HANKEN, M.A.; WATERS, K.A. (Ed.). **Electronic Health Records: Changing the Vision**. Philadelphia: W.B. Saunders Company, 1999.
- [LAG 97] LAGOZE, Carl; DANIEL, Ron. Extending the Warwick Framework. **D-Lib Magazine**, Nov. 1997. Disponível em: <<http://www.dlib.org/dlib/november97/daniel/11daniel.html>>. Acesso em: jan. 2002.
- [LAS 97] LASSILA, Ora; SWICK, Ralph R. Resource Description Framework (RDF) Model and Syntax. Disponível em: <<http://www.w3.org/TR/WD-rdf-syntax-971002/>>. 1997. Acesso em: maio 2002.
- [LIG 99] LIGHT, Richard. **Iniciando em XML**. São Paulo: Makron Books, 1999.
- [MAR 2000 a] MARQUES, Paulo Mazzoncini de Azevedo et al. Implantação de um Sistema de Informação em Radiologia em Hospital Universitário. **Radiol. Bras.**, São Paulo, p. 155-60, 2000
- [MAR 2000] MARTINEZ, Efonso M.; AZIPIROZ, Joaquim Leehan. **Instalación y operación de sistemas PACS (almacenamiento y comunicación de imágenes) en México: características fundamentales**. Disponível em: <<http://itzamna.uam.mx/joaquin/pacs/PACS2.html>>. Acesso em: jan. 2000.
- [MAR 97] MARTINS, Raul H.D. **A variável localização geográfica como elemento integrador de repertórios heterogêneos de dados de saúde pública, no contexto do Rio Grande do Sul**. 1997. Dissertação (Mestrado em Informática) – Instituto Informática, PUCRS, Porto Alegre.
- [MAR 98] MATINEZ M.A et al. **Los sistemas PACS**. Laboratorio de Investigación en Computación y Procesamiento Digital de Señales e Imágenes. Disponível em: <<http://itzamna.uam.mx/alfonso/pacs.html>>. Acesso em: fev. 2000.

- [MAT 2001] MATHIESEN, Finn K. Web technology – the future of teleradiology ? **Computer Methods and Programs in Biomedicine**, Denmark, p. 87-90, 2001.
- [MCD 90] MCDONALD, C.J.; BARNETT, G.O. Medical-Record Systems. In: SHORTLIFFE, E.H., PERREAULT, L.E.(Ed). **Medical Informatics: Computer Applications in Health Care**. New York: Addison-Wesley, 1990. p.181-218.
- [MCM 99] MCMULLIN, S.T; REICHLEY, R.M; WATSON, L.A. et al. Impact of a Web-based clinical information system on cisapride drug interactions and patient safety. **Archives of Internal Medicine**, Chicago, v.159, n.17, p.2077-82, Sept. 1999.
- [MIL 99] MILSTEAD, J.; FELDMAN S. **Metadata: Cataloging by Any Other Name**, 1999. Disponível em:  
<<http://www.onlineinc.com/onlinemag/OL1999/milstead1.html - projects>>;  
<[http://garnet.acns.fsu.edu/~erm3123/iamslic\\_metadata.html](http://garnet.acns.fsu.edu/~erm3123/iamslic_metadata.html)>. Acesso em: jun. 1999.
- [NAR 2001] NARDON, Fabiane B. **Prontuário Eletrônico do Paciente**. Palestra no Instituto do Coração de São Paulo, 2001.
- [NAT 2000] NATANYA Pitts-Moultis; CHERYL, Kirk. **XML: Black Book**. São Paulo: Makron Books, 2000.
- [PAC 2000] Picture Archiving and Communication Systems. **Health Devices**, [S.l.], v.29, n.11, p.385-427, Nov. 2000.
- [PEL 2001] PELIKAN, E; KOTTER, E; TIMMERMANN, U. Networks in the radiology department and the hospital. **European Radiology**, Berlim, v.11, n.2, p. 337-345, May 2001.
- [PER 2002] PEREIRA, Rui Ribeiro. ASP: um Regresso ao Futuro. **ComputerWorld**, São Paulo, Jan. 2002. Disponível em:  
<[http://www.computerworld.iol.pt/cw\\_dictionary\\_02.asp?artigo=5005&palavra=10](http://www.computerworld.iol.pt/cw_dictionary_02.asp?artigo=5005&palavra=10)>. Acesso em: jan 2002.
- [PIS 2000] PISA, Ivan Torres; RUIZ, Evandro E.S. Proposta de Desenvolvimento de um PACS Distribuído Utilizando CORBA para Gerenciamento de uma Base de Imagens Médicas. In: CONGRESSO BRASILEIRO DE ENGENHARIA BIOMÉDICA, 17., 2000, Florianópolis. **Anais...** Disponível em:  
<<http://imagorb.ffclrp.usp.br/old-ivanpisa/download/5261-59635.pdf>>. Acesso em: abr. 2002.

- [PNG 97] PNG.Things that go PNG on the Web. Disponível em:  
<<http://www.cs.man.ac.uk/aig/staff/toby/writing/png.htm>>. Acesso em:  
maio 1997.
- [RFC 2413] RFC 2413, Request for comments 2413, 1998. Disponível em:  
<<http://www.ietf.org/rfc/rfc2413.txt>>. Acesso em: jan. 2002.
- [ROE 97] ROELOFS, Greg. **PNG's Not GIF**. 1997. Disponível em:  
<<http://webreview.com/97/05/09/feature/>>. Acesso em: fev. 2002.
- [SAB 98 a] SABBATINI, R.M.E. Internet em Medicina: os recursos. **Revista  
Informática Médica**, São Paulo v.1, n.1, jan./fev 1998. Disponível em:  
<<http://www.epub.org.br/informaticamedica/n0101/sabbatini.htm>>. Acesso  
em: abr. 2002.
- [SAB 98] SABBATINI, Renato M.E. Informática e a Medicina do Século 21: **Revista  
Informática Médica**, São Paulo, v.1, n.2, mar./abr. 1998. Disponível em:  
<<http://www.epub.org.br/informaticamedica/n0102/hogarth.htm>>. Acesso  
em: abr. 2002.
- [SAB 99] SABBATINI, R.M.E. Os Médicos, os Pacientes e a Internet. **Revista  
Médico Repórter**, [S.l.], set. 1999. Disponível em:  
<<http://www.nib.unicamp.br/papers/reporter-medico-07.htm>>. Acesso em:  
mar. 2000.
- [SAF 2000] SAFRAN , C.; GOLDBERG, H. Electronic patient records and the impact  
of the Internet. **International Journal of Medical Informatics**, [S.l.], v.60,  
n.2, p.77-83, Nov. 2000.
- [SAI 2001] SAIF. Disponível em:  
<[http://epoch.cs.berkeley.edu:8000/sequoia/schema/html/saif/section1.1.htm  
l](http://epoch.cs.berkeley.edu:8000/sequoia/schema/html/saif/section1.1.html)>. Acesso em: nov.2001.
- [SHA 97] SHAH, Kshitij et al. **Black Box Approach to Image Feature  
Manipulation used by Visual Information Retrieval Engines**. Disponível  
em:  
<<http://computer.org/conferen/proceed/meta97/papers/kshah/ksah.html>>.  
Acesso em mar. 1997.
- [SLE 2000] SLEE, V.N.; SLEE, D.A.; SCHMIDT, H.J. **The endangered medical  
record: ensuring its integrity in the age of informatics**. Saint Paul: Tringa  
Press, 2000.
- [SPI 95] SPIFF - "GFF Format Summary: SPIFF". Disponível em:  
<<http://www.ipahome.com/gff/textonly/summary/spiff.htm>>;  
<<http://www.virage.com/market/vir-body.html>>. Acesso em: nov. 2001.

- [TAC 96] TACHINARDI, Umberto. Intranet e sistemas de informações hospitalares. In: FÓRUM NACIONAL DE CIÊNCIA E TECNOLOGIA EM SAÚDE, 3.,1996, São Paulo. **Anais...** São Paulo: FNCTS, 1996. p. 677-678.
- [W3C 2000] W3C, The World Wide Web Consortium. Disponível em: <<http://www.w3.org/>>. Acesso em: jan. 2002.
- [WEI 95] WEIBEL, S. PURLs: Persistent Uniform Resource Locators. **OCLC Newsletter**, 1995, Disponível em: <<http://purl.oclc.org/OCLC/PURL/SUMMARY>>. Acesso em: jan. 2002.
- [WIG 2001] WIGGINS, Richard H. et al. Image file formats: past, present, and future. **RG Radiographics**, Easton, v.21, n.3, p.789-98, May/June 2001.
- [WON 97] **WONG**, Stephen T.C. Networked multimedia for medical imaging. **IEEE Multimedia**, Los Alamitos, v.4, n.2, p.24-35, Apr./June 1997.
- [ZOL 2001] ZOLTÁN, E. **Server-side scripting languages - PHP, Perl, Java servlets - - Which one' s right for you ?** Disponível em: <<http://www-106.ibm.com/developerworks/web/library/wa-sssl.html?dwzone=web>>. Acesso em: jan. 2002.