

JULIANO LUCAS GONÇALVES

**ESTUDO DE GESTÃO DE DADOS NÃO  
CONVENCIONAIS BASEADA EM METADADOS  
PARA O AMBIENTE DE DIPOSITIVOS MÓVEIS**

Dissertação apresentada como requisito parcial à  
obtenção do grau de mestre. Programa de Pós-  
Graduação em Informática, Setor de Ciências  
Exatas, Universidade Federal do Paraná.

Orientadora: Dr<sup>a</sup> Maria Salete M. G. Vaz.

CURITIBA  
2005



## AGRADECIMENTOS

Aos meus pais Olivério e Rita, a minha irmã Lissani e ao meu sobrinho Gabriel por fazerem parte da minha vida e estarem sempre comigo. Mesmo à distância eles estiveram sempre me apoiando para a realização desse mestrado e, sobretudo, para conclusão desse trabalho.

A Rafaela pela companhia e pelo carinho de sempre.

A minha orientadora Professora Dr<sup>a</sup> Maria Salete Marcon Gomes Vaz pelo auxílio, sempre que solicitado, o qual foi imprescindível à realização e conclusão desse trabalho. Muito obrigado pela confiança e incentivo de sempre.

Aos professores membros da banca Dr. João Umberto Furquim de Souza e Dr<sup>a</sup> Laura Garcia Sanches pelas importantes considerações, as quais, contribuíram muito para a entrega da versão final desse trabalho.

Aos professores do Mestrado: Direne, Elias, Silvia e Sunye.

A Jucélia pelo tratamento sempre atencioso.

Aos amigos que fiz no HC: Baggio, Edson, Emerson, Simão.

Aos grandes amigos que fiz durante o mestrado: Cláudio, David, Alexandre, André, Beatriz, Cássio, Diego, Eduardo, João Paulo, José Augusto, Nádia, Paulo, Pedro e Rubens. Desejo a todos vocês muito sucesso.

## SUMÁRIO

<b>LISTA DE FIGURAS.....</b>	<b>VI</b>
<b>LISTA DE TABELAS.....</b>	<b>VII</b>
<b>RESUMO.....</b>	<b>VIII</b>
<b>ABSTRACT .....</b>	<b>IX</b>
<b>1. INTRODUÇÃO .....</b>	<b>1</b>
<b>2. GESTÃO DE METADADOS .....</b>	<b>3</b>
2.1 DEFINIÇÕES DE METADADOS.....	3
2.2 UTILIZAÇÃO DOS METADADOS .....	5
2.3 IMPORTÂNCIA DA CRIAÇÃO DE METADADOS.....	7
2.4 TIPOS DE INFORMAÇÕES CONSIDERADAS METADADOS .....	8
2.5 INDEXAÇÃO, BUSCA E RECUPERAÇÃO DA INFORMAÇÃO.....	11
<b>3. PADRÕES DE METADADOS .....</b>	<b>13</b>
3.1 PADRÃO DUBLIN CORE.....	13
3.2 PADRÃO MARC .....	19
3.3 PADRÃO FGDC/CSDGM .....	21
3.4 PADRÃO TEI.....	27
3.5 PADRÃO GILS .....	30
<b>4. DISPOSITIVOS MÓVEIS .....</b>	<b>36</b>
4.1 INTRODUÇÃO .....	36
4.2 VANTAGENS DOS DISPOSITIVOS MÓVEIS.....	39
4.3 SISTEMAS GERENCIADORES DE BANCO DE DADOS MÓVEIS .....	40
4.3.1 Sybase UltraLite .....	40
4.3.2 Oracle Lite Mobile Server .....	41
4.3.3 DB2 Everyplace.....	42
4.3.4 Microsoft SQL Server CE .....	43

4.4 DESENVOLVIMENTO DE APLICAÇÕES PARA DISPOSITIVOS MÓVEIS.....	44
<b>5. ESTRUTURA DE GESTÃO DE DADOS PARA DISPOSITIVOS MÓVEIS ATRAVÉS DE METADADOS .....</b>	<b>48</b>
5.1 ESTRUTURA DE GESTÃO DE DADOS.....	48
5.2 ESCOLHA DOS METADADOS A SEREM INSERIDOS.....	50
5.3 ESPECIFICAÇÕES DA ESTRUTURA PROPOSTA .....	51
5.3.1 Indexação do Objeto.....	52
5.3.2 Busca e Recuperação do objeto.....	54
5.4. TRABALHOS RELACIONADOS.....	56
5.4.1 Sistema VIMSYS .....	56
5.4.2 Sistema CIARS.....	57
5.4.3 Análise Comparativa .....	57
<b>6. CONCLUSÕES E TRABALHOS FUTUROS.....</b>	<b>62</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>64</b>

## LISTA DE FIGURAS

FIGURA 3.1 - PANORAMA ORGANIZACIONAL DO PADRÃO FGDC/CSDGM.....	23
FIGURA 3.2 - ESTRUTURA GILS.....	31
FIGURA 4.1 - EXEMPLO DE UM PDA APPLE DE 1992.....	36
FIGURA 4.2 - POCKETS PC COM A PLATAFORMA WINDOWS CE .....	37
FIGURA 5.1 - ESTRUTURA DE GESTÃO PROPOSTA.....	49
FIGURA 5.2 - OPERAÇÕES DISPONÍVEIS .....	51
FIGURA 5.3 – DEFINIÇÃO DO OBJETO A SER INSERIDO .....	52
FIGURA 5.4 – ESTRUTURA DE BUSCA E RECUPERAÇÃO. ....	55

## LISTA DE TABELAS

TABELA 2.1 - ATRIBUTOS METADADOS .....	4
TABELA 2.2 - NÍVEL DE GERENCIAMENTO DO CONHECIMENTO.....	5
TABELA 3.1 - ELEMENTOS DUBLIN CORE POR CATEGORIA.....	14
TABELA 3.2 - EXEMPLIFICAÇÃO DOS ELEMENTOS DUBLIN CORE.....	15
TABELA 3.3 - FORMAS DE ARMAZENAMENTO DE METADADOS EM DUBLIN CORE. .....	16
TABELA 3.4 - MARC TAGS.....	19
TABELA 3.5 - SEÇÕES PRINCIPAIS DO PADRÃO TEI.....	27
TABELA 4.1 - DIFERENÇAS ENTRE HTML E XML.....	45
TABELA 5.1 - CONJUNTO DE METADADOS FIXOS UTILIZADOS NO MODELO PROPOSTO.....	50
TABELA 5.2 – METADADOS/VALORES.....	53
TABELA 5.3 - OBJETO .....	53
TABELA 5.4 - ANÁLISE COMPARATIVA SISTEMAS .....	58

## RESUMO

A utilização de dispositivos móveis vem crescendo muito rapidamente ao longo dos anos. Esse crescimento é responsável pela evolução desses dispositivos no que diz respeito, principalmente, à capacidade de armazenamento e processamento, propiciando a manipulação de quantidades e formatos cada vez maiores de dados por seus usuários. Esse aumento traz como consequência à necessidade do usuário poder gerenciar esses dados ditos não convencionais tais como imagens, vídeos e sons de maneira simples e rápida, facilitando sua localização e recuperação. A utilização de metadados surge como uma alternativa para se fazer esse gerenciamento. Metadados são informações geralmente textuais que descrevem outros objetos, tendo como principal objetivo o de possibilitar, posteriormente, a recuperação dos mesmos. Logo o objetivo desse trabalho é realizar um estudo dos aspectos que envolvem a gestão de dados ditos não convencionais (imagens, vídeo, som) em dispositivos móveis, bem como o de apresentar uma estrutura de gestão desses dados através da utilização de metadados. Essa estrutura tem por objetivo permitir ao usuário gerenciar dados não convencionais através de um conjunto de metadados/valores predefinidos, permitindo, também a inserção de novos metadados pelo usuário, sempre que este julgar necessário, facilitando sua recuperação.

Palavras-chave: Metadados, Gestão de dados, Dispositivos móveis.



## **ABSTRACT**

The use of movable devices is growing very quickly along the years. That growth is responsible for the evolution of those devices in what he says respect, mainly, to the storage capacity and processing, propitiating the manipulation of amounts and formats every time larger of data for your users. That increase brings as consequence to the user's need to can administration those said data you don't stipulate such as images, videos and sounds in a simple and fast way, facilitating your location and recovery. The metadata use appears as an alternative to do that administration. Metadata are usually information textual that they describe other objects, tends as objective main the making possible, later, the recovery of the same ones. Therefore the objective of that work is to accomplish a study of the aspects that you involve the administration of data statements you don't stipulate (images, video, sound) in movable devices, as well as the one of presenting a structure of administration of those data through the metadados use. That structure have it goes objective to allow to the user managements don't stipulate through the group of metadata/values, defined before, allowing, also the insert of new metadados goes the user, whenever this to judge necessary, facilitating your recovery.

**Keywords:** Metadata, Administration of data, Movable devices.

## 1. INTRODUÇÃO

Os dispositivos móveis e suas aplicações estão utilizando cada vez mais quantidades crescentes de novos tipos de dados, tais como vídeo, imagem e som. O controle de grandes quantidades de dados introduz um desafio da usabilidade que não pode ser resolvido sem desenvolver novas maneiras de tratamento desses dados. O uso no contexto de metadados é uma maneira de fazer o controle desses dados.

Segundo VAZ [51], o uso de metadados na descrição de dados não convencionais é importante pela dificuldade na pesquisa baseada em conteúdo, isto se dá devido à análise em grande conjunto de dados, pois quando a pesquisa baseada em conteúdo é possível ela fica comprometida por razões de desempenho.

A quantidade de dados que os usuários de telefones móveis podem armazenar está crescendo muito rapidamente. Por muitos anos, os usuários têm armazenado contatos, mensagens do texto, eventos do calendário e notas. Mais recentemente, os telefones móveis tornaram-se capazes de armazenar e até mesmo criar tipos diferentes de dados tais como imagem, vídeo e som.

A finalidade da gerência de dados por meio de metadados é dar suporte a recuperação dos mesmos, organizando-os de modo que procurar manual ou automaticamente se torne possível.

Em dispositivos móveis a utilização dos metadados possibilita ao usuário descrever seus dados por meio dos metadados, o que não é possível atualmente, facilitando a busca dos dados de forma mais rápida e consistente, pois atualmente, imagens, vídeos, sons, são armazenadas em diretórios específicos somente pelo atributo nome, tornando lento o acesso aos dados.

Futuramente, tendo em vista o crescimento acentuado da capacidade de armazenamento desses dispositivos, a tendência é que o volume de dados contido nos mesmo aumente significativamente, sendo, portanto, necessário esse controle.

O objetivo desta dissertação é realizar um estudo dos aspectos que envolvem a gestão de dados ditos não convencionais (imagens, vídeo, som) em dispositivos móveis, bem como apresentar uma estrutura de gestão através da utilização de metadados.

Para tanto esta dissertação está estruturada como segue. No Capítulo 2 são descritos os aspectos inerentes à Gestão de Metadados. No Capítulo 3 são descritos padrões de metadados e alguns exemplos de sua utilização. No Capítulo 4 é apresentada a fundamentação teórica de dispositivos móveis. No Capítulo 5 é apresentado, como contribuição desta dissertação, uma estrutura de gestão de dados não convencionais para o ambiente de dispositivos móveis através do uso de metadados. No Capítulo 6 são feitas as considerações finais e as perspectivas de trabalhos futuros são apresentadas.

## 2. GESTÃO DE METADADOS

Para gerenciar metadados é necessário saber o que eles significam, sua importância e para que servem. Nas subseções seguintes são apresentadas algumas definições para metadados, onde podem ser utilizados e o que envolve a classificação dos mesmos, além de alguns padrões de metadados existentes.

### 2.1 Definições de Metadados

Na literatura, os metadados são definidos de forma clássica como sendo dados sobre dados, ou seja, um conjunto de elementos que possuem uma semântica padronizada, a qual possibilita representar informações eletrônicas e descrever recursos eletrônicos de maneira bibliográfica [21].

Os metadados podem ser entendidos, também, como um meio de se descobrir recursos existentes, bem como eles podem ser obtidos e acessados, evitando a ambigüidade dos dados [21].

Segundo DATE [12], metadados são descrições de outros objetos no sistema, ao invés de simples “dados brutos”, transformando esses dados em conhecimento. Já para BAEZA-YATTES [3] os metadados são atributos de dados ou documentos, os quais contém informações sobre autor e conteúdo, sendo subdivididos em categorias, as quais, são mantidas em um catálogo, algumas vezes registrado de acordo com o *framework*, como dos Padrões de Metadados *Dublin Core* [15] ou *MARC - Machine Readable Card* [31].

Segundo DESIO [14], metadados são dados que descrevem atributos de um recurso. Eles suportam um número de funções: localização, descoberta, documentação, avaliação, seleção, entre outras. Já para WEAVER [54], metadados separam e relatam propriedades, de forma que o usuário possa ser preciso na descrição do conteúdo do documento. Essa precisão é expressa por meio de metadados externos ou metadados baseados em conteúdo [25].

Os metadados externos referem-se a informações associadas à produção e uso do documento. Exemplos como linguagem, data de criação e autor são aplicados ao documento inteiro. Já os metadados baseados em conteúdo são as descrições sobre o que realmente é um documento. Palavras-chave contidas em um artigo enviado para jornais ou revistas, por exemplo, são metadados baseados em conteúdo.

Segundo BALDONADO [4], metadados podem descrever não somente documentos, mas também coleções e bases de dados inteiras, juntamente com os serviços que as mesmas oferecem. O registro de metadados associado a um objeto digital contém informações descritiva e funcional [9]. A Tabela 2.1 apresenta uma lista com exemplos de metadados descrevendo características de serviços gerais.

TABELA 2.1 - ATRIBUTOS METADADOS

<b>Metadados</b>	<b>Valores de metadados</b>
Versão	Versão do objeto metadado
Nome de coleção	Nome da coleção que esta sendo descrita
Nomes Atributos	Atributos que podem ser atribuídos
Linguagens	Origens das linguagens
Data de modificação	Data da ultima modificação do tipo de objeto metadados
Data de expiração	Data em que o tipo de objeto metadados será revisado
Resumo	Resumo da coleção
Restrições Acesso	Restrições para acessar a coleção
Contato	Informação para contato com o administrador da coleção

Segundo MORIARTY [35] estas diferentes interpretações para conceitos de metadados ocorrem devido ao estágio da organização dentro da hierarquia evolucionária de gestão do conhecimento como mostra a Tabela 2.2.

As organizações no nível mais baixo da hierarquia gerenciam dados brutos. Organizações mais avançadas são capazes de gerenciar seus recursos de informação no nível de Informação, Conhecimento ou Sabedoria! No nível de informação, o foco é nos relacionamentos entre todos os componentes do sistema e os papéis individuais que eles assumem no sistema.

Uma organização que está envolvida ao ponto que ela pode mostrar explicitamente as regras de negócio governado, seu comportamento está no nível de conhecimento.

TABELA 2.2 - NÍVEL DE GERENCIAMENTO DO CONHECIMENTO

<b>Estágio</b>	<b>Recurso a ser administrado</b>	<b>Definição de Metadados</b>
Dados	Valores dos dados	Informação necessária para administrar o recurso dos dados
Informação	Valores dos dados e o contexto da informação	Informação necessária para administrar o recurso da informação
Conhecimento	Valores dos dados, contexto da informação e instruções das regras de negócio.	Informação necessária para administrar as regras e políticas de negócio da organização
Sabedoria	Valores de dados, contexto da informação, regras de negócio que podem ser executados, monitoração das regras de negócio e regras e métricas de avaliação.	Informação necessária para administrar o comportamento da organização de acordo com suas regras e políticas de negócio.

Uma organização atinge o nível mais avançado da hierarquia quando ela monitora ativamente seus sistemas para garantir que seu comportamento esteja de conformidade com o planejado. Tal que uma organização pode detectar e diagnosticar qualquer comportamento anormal do sistema.

## 2.2 Utilização dos Metadados

Atualmente, o interesse sobre metadados vem crescendo porque as pessoas precisam encontrar e avaliar informações na Internet e nas intranets, além disso, os sistemas de gerenciamento de conhecimento precisam integrar informações de fontes múltiplas e aplicações precisam oferecer maior facilidade de pesquisa e manutenção [45].

Os metadados podem ser utilizados de duas formas diferentes, podendo estar embutidos nos documentos ou armazenados separadamente. Normalmente, são embutidos em documentos *HTML - Hypertext Markup Language* [26] ou cabeçalhos de arquivos de

imagens. A vantagem de armazenar o documento é a garantia de que os mesmos não serão perdidos, eliminando problemas de perdas de *links* além de garantir que as atualizações dos documentos e dos metadados sempre serão simultâneas. Porém, existem casos em que os metadados não podem ser armazenados juntamente aos documentos, sendo armazenados em banco de dados e sim ligados aos documentos, facilitando o gerenciamento, pesquisa e recuperação dos mesmos.

Na literatura podem ser encontradas umas variedades de taxonomias a respeito de metadados, e após um estudo referente aos mesmos é possível enquadrá-los em duas categorias básicas, metadados técnicos e de negócios.

Os metadados técnicos descrevem os dados necessários pelas várias ferramentas para armazenar, manipular ou movimentar dados. Estas ferramentas incluem banco de dados relacionais, ferramentas de desenvolvimento de aplicações, ferramentas de modelagem, ferramentas de pesquisa em banco de dados, ferramentas OLAP, entre outras.

Segundo SHAEFER [45] para se atingir um nível de integração de metadados e necessário utilizar-se de algumas estratégias como as que seguem.

- Construção de uma ponte proprietária entre vários produtos. A vantagem dessa estratégia é que permite uma alta funcionalidade e compartilhamento de metadados. Porém, como desvantagem, a manutenção consome muitos recursos, especialmente com o crescimento da mesma.
- Fornecer uma solução completa de Tecnologia de Informação. Adquirir ou construir componentes necessários para oferecer um ambiente integrado de uma solução completa. A desvantagem para o usuário é que o obriga a abandonar as melhores abordagens ou se sujeitar a produtos de um único fabricante.
- Criar um padrão internacional para troca de metadados. A utilização de um meio comum para troca de metadados por meio dos padrões internacionais. Existem padrões diferentes de metadados para finalidades distintas de informações.

Os metadados de negócios descrevem os dados necessários pelos usuários de negócio, para entender o contexto do negócio e o significado dos dados.

O estudo dos metadados técnicos é parte importante nesse trabalho, visto que cada dispositivo móvel dispõe de características e particularidades próprias tais como capacidade de armazenamento, transmissão de dados, tempo de funcionamento (duração das baterias), entre outras.

### **2.3 Importância da Criação de Metadados**

Os metadados são importantes na descoberta de recursos, ou seja, permitem a pesquisa de recursos por critérios relevantes, identificação de recursos, agrupamento de recursos similares, diferenciação de recursos não similares e a obtenção de informação de localização [21].

Utilizados na organização eletrônica de recursos, ou seja, devido ao crescimento do número de recursos disponíveis na Internet, torna-se importante o agrupamento de recursos semelhantes facilitando a busca e recuperação dos mesmos pelo usuário.

Outro aspecto importante é a interoperabilidade, ou seja, a descrição de recursos por meio de metadados permitindo que eles sejam compreendidos por programas, facilitando, assim, a interoperabilidade entre os sistemas.

Segundo Gilliland-Swetland [22], os metadados são importantes no arquivamento e preservação de arquivos, ou seja, formatos de arquivo que são padrão, atualmente, podem ser substituídos no futuro e as informações nos metadados podem ajudar a impedir que ocorra a perda das informações contidas em arquivos nesses formatos.

“Muitas organizações não fazem o melhor uso de suas informações porque elas não são bem gerenciadas. Uma das razões porque a informação é tão mal administrada pode ser porque ela não é bem entendida. A informação não obedece às mesmas leis econômicas de outros patrimônios” [35].



Ela possui algumas propriedades únicas que precisam ser entendidas para administrá-la de uma maneira mais efetiva.

A informação é compartilhável infinitamente. Pode-se combinar informações de várias formas para benefício da empresa. Restrição de informação significa perda de oportunidade de negócios, como várias áreas de negócio controlando a mesma informação.

O valor da informação aumenta com o uso. Em muitas organizações encontram-se informações que poderiam ser usadas para adquirir vantagem competitiva. Porém, isso não é feito porque as pessoas não sabem que ela existe ou não podem acessá-la.

O valor da informação diminui com o tempo. Porém, ela varia conforme o tipo da informação. Informação para tomada de decisão tem uma vida útil maior que informações operacionais, por serem usadas mais freqüentemente.

O valor da informação aumenta quando combinada com outra informação. A informação geralmente é mais útil quando ela pode ser comparada e combinada com outra informação.

A informação gera outra informação. O processo de usá-la tende a resultar em mais informação.

“Até hoje a transformação de dados em informação tem sido conduzida pela tecnologia (Dados + Contexto = Informação). Estes avanços tecnológicos estão sendo integrados pela organização. Esta integração ajuda a gerenciar não só a informação, mas também formas em que as pessoas aprenderão e influenciarão a informação e a compartilharão com outras. (Informação + Experiência = Conhecimento)” [35].

## **2.4 Tipos de Informações consideradas Metadados**

Os metadados são utilizados normalmente como um dicionário de informações e sendo assim, devem incluir:

- Origem dos Dados - Todo elemento de dado precisa ter identificado, sua origem ou o processo que o gera. Esta identificação é muito importante no caso de se necessitar saber informações sobre a fonte geradora do dado. Esta informação deve ser única, ou seja, cada dado deve ter uma e somente uma fonte de origem.
- Fluxo dos Dados - Todo elemento de dado precisa ter identificados os fluxos nos quais sofre transformações. É importante saber que dados servem de base para que processos.
- Formato dos Dados - Todo elemento de dados deve ter identificado seu tamanho e tipo.
- Nomes e Alias - Todo elemento de dado deve ser identificado por um nome. Este nome pode ser da Área de Negócios ou um nome técnico. No caso de serem usados aliases para os nomes, pode-se ter os dois. Devem existir padrões para criação de nomes e alias (ex.: convenções para abreviações), evitando assim ambigüidades.
- Definições de Negócio - Estas definições são as informações mais importantes contidas nos metadados. Cada elemento de dado deve conter uma definição do mesmo no contexto da Área de Negócio. O método de manutenção destas informações também deve ser muito consistente, de forma que o usuário possa obter facilmente definições para as informações desejadas. Nestas definições devem ser evitadas referências a outros metadados que necessitem de uma segunda pesquisa para melhor entendimento.
- Regras de Transformação - São consideradas como sendo as Regras de Negócio codificadas. Estas regras são geradas no momento da extração, limpeza e agrupamento dos dados dos Sistemas Operacionais. Cada regra de

transformação codificada deve estar associada a um elemento de Metadado. Se mais de uma aplicação contiver a mesma regra de transformação, deverá ser garantido que estas sejam idênticas.

- Atualização dos Dados - O histórico das atualizações normalmente é mantido pelo próprio banco de dados, mas ter definido um elemento de metadado indicando as datas de atualização dos dados pode facilitar o usuário no momento de verificar a atualidade dos dados e a consistência da dimensão tempo do Data Warehouse.
- Requisitos de Teste - Identifica os critérios de julgamento de cada elemento de dado. Valores possíveis e intervalos de atuação. Deve conter também padrões para procedimentos de teste destes dados.
- Indicadores de Qualidade de Dados - Podem ser criados índices de qualidade, baseados na origem do dado, numero de processamentos feito sobre este dado, valores atômicos, nível de utilização do dado, etc.
- Triggers Automáticos - Podem existir processos automáticos associados aos metadados definidos. Estes processos ou triggers devem estar definidos de forma que possam ser consultados por usuário e desenvolvedores, para que os mesmos não venham a criar situações conflitantes entre as regras definidas nestes processos.
- Responsabilidade sobre Informações - Deve ser identificado o responsável por cada elemento de dados do *Data Warehouse* e também o responsável pela entrada de metadados.
- Acesso e Segurança - Os metadados devem conter informação suficiente para que sejam determinados perfis de acesso aos dados. Deve-se poder identificar que usuários podem ler, atualizar, excluir ou inserir dados na base. Deve

haver também informações sobre quem gerencia estes perfis de acesso e como se fazer contato com o administrador da base de Dados.

Todas essas informações, citadas anteriormente, são aspectos importantes a serem considerados visando um conhecimento amplo dos metadados a serem descritos, com o intuito de garantir sua procedência e, conseqüentemente, a segurança dos mesmos.

## **2.5 Indexação, Busca e Recuperação da Informação.**

A Indexação [46], é a técnica para identificação do conteúdo de um dado para sua posterior recuperação. Tem por objetivo permitir que os objetos possam ser recuperados posteriormente pelos usuários da informação. A indexação permite identificar e representar o conteúdo ou o assunto de um documento de forma a caracterizá-lo. Na indexação não se deve levar em consideração somente o armazenamento, mas principalmente sua futura busca e recuperação.

Na recuperação de informação, os dados não convencionais são representados como uma coleção de aspectos. O usuário irá especificar o que deseja na forma de uma consulta. Existem vários métodos de recuperação de informação textual na literatura, tais como [48]: varredura completa de texto e inversões entre outros.

No método varredura completa de texto, o usuário entra com um padrão a ser pesquisado, e o sistema percorre a base de dados até que um texto igual ao especificado anteriormente seja recuperado. Esse método pode se tornar lento devido à grande quantidade de informações contidas no mesmo.

No método de inversões, cada entrada consiste em um padrão e uma lista de endereços para os objetos onde o padrão ocorre. Esse método torna-se adequado em ambientes estáticos, ou seja, onde ocorram muitos acessos. Contudo, poucas ou nenhuma modificações.

Existem alguns modelos de recuperação de dados não convencionais, dentre os quais podem ser citados: recuperação por meio de um identificador [50] e por sentenças condicionais [50], entre outros.

Na Recuperação por meio de um Identificador esse é dado como índice e o objeto é recuperado. Já a Recuperação por sentenças condicionais os objetos são recuperados somente se eles respondem precisamente às restrições expressas na consulta.

Os objetos não convencionais podem ser descritos por meio de *strings*, ou seja, seqüência de caracteres, as quais necessitam de algoritmos de grande eficiência para manipula-las. Dentre esses algoritmos [33], podemos citar: KMP (*Knuth-Morris-Prat*) e Boyer- Moore, entre outros.

O algoritmo KMP compara um padrão com um texto, caractere a caractere. Porém, quando caracteres de posições correspondentes diferem, ou seja, quando ocorre um *mismatch*, o KMP aproveita a informação em que os caracteres do padrão (ou parte dele) casaram com o texto, evitando assim comparações redundantes. Já o algoritmo Boyer-Moore compara o padrão com o texto, da direita para a esquerda, ou seja, a primeira comparação é feita entre o último caractere do padrão e o caractere do texto correspondente à última posição do padrão. Se estes são iguais, então os caracteres da posição precedente do padrão e do texto são comparados, e assim por diante. Ou seja, a busca por ocorrências do padrão no texto é feita normalmente, da esquerda para a direita, apenas a comparação é feita da direita para a esquerda.

Pelo fato desses dados não convencionais serem representados em sua forma original, acrescidos de informações, as quais, descrevem aspectos inerentes aos mesmos, torna-se importante defini-los utilizando metadados.

### 3. PADRÕES DE METADADOS

“Os padrões de metadados têm como objetivo fornecer as funções e formar uma rede para automatizar registros de propriedades e dados cadastrais de uma forma padronizada e consistente” [10].

O esforço para desenvolver, organizar e padronizar o uso de metadados é desenvolvido por meio de vários programas cooperativos na Internet, orientados pelo *World Wide Web Consortium (W3C)* [53], órgão que regula o desenvolvimento técnico da Internet. Este órgão é responsável pela normalização evolutiva da Linguagem HTML [26] em todas as suas versões e na implementação de linguagens derivadas como o XML (*Extensible Markup Language*) [5]. Dedicados a explorar um padrão documentário para metadados, esforços internacionais conjuntos estão sendo realizados por organizações biblioteconômicas e de normas e padrões técnicos, nacionais e internacionais.

A iniciativa de maior sucesso atualmente originou-se de uma série de seminários, tornando-se conhecida como *Dublin Core Metadata Element Set* [15], Conjunto de Elementos de Metadados Dublin Core [53].

O Seminário sobre o tema realizou-se pela primeira vez em março de 1995. Mas além do Dublin Core existe, ainda, uma variedade de padrões para metadados, criados para diferentes tipos de informações [53].

#### 3.1 Padrão Dublin Core

A Iniciativa de Metadados Dublin Core (*Dublin Core Metadata Initiative - DCMI*) [16], é uma organização que surgiu em Dublin, Ohio (EUA) em 1995 e que está dedicada a promover a adoção de padrões de interoperabilidade em metadados [10]. O Conjunto de Elementos Dublin Core (*Dublin Core Element Set*), [15] é um conjunto de 15 elementos subdivididos em 3 categorias como mostra a Tabela 3.1.

TABELA 3.1 - ELEMENTOS DUBLIN CORE POR CATEGORIA

<b>Conteúdo</b>	<b>Propriedade Intelectual</b>	<b>Manifestações Físicas</b>
<i>Title</i>	<i>Creator</i>	<i>Date</i>
<i>Subject</i>	<i>Source</i>	<i>Type</i>
<i>Description</i>	<i>Contributor</i>	<i>Format</i>
<i>Publisher</i>	<i>Rights</i>	<i>Identifier</i>
<i>Language</i>		
<i>Relation</i>		
<i>Coverage</i>		

1. *Title*: nome pelo qual o recurso é formalmente conhecido.
2. *Creator*: entidade primariamente responsável pelo conteúdo do recurso (pessoa, organização ou serviço).
3. *Subject*: palavras chave, frases chave ou código de classificação preferencialmente selecionado de um vocabulário controlado.
4. *Description*: resumo (*abstract*), sumário ou texto livre sobre o conteúdo.
5. *Publisher*: entidade responsável pela disponibilização do recurso (pessoa, organização ou serviço).
6. *Contributor*: entidades responsáveis por contribuições ao conteúdo do recurso (pessoa, organização ou serviço).
7. *Date*: data da criação ou disponibilização do recurso no formato YYYY-MM-DD.
8. *Type*: natureza ou gênero do conteúdo do recurso.
9. *Format*: tipo de mídia do recurso. Utilizado para identificar o software ou hardware necessário para exibir ou executar o recurso.
10. *Identifier*: identificador único do recurso dentro de um determinado contexto.
11. *Source*: referência a um recurso do qual o presente recurso é derivado
12. *Language*: linguagem do conteúdo intelectual do recurso.
13. *Relation*: referência a um recurso relacionado.
14. *Coverage*: extensão ou escopo do conteúdo do recurso.

15. *Rights*: informações sobre os direitos inerentes ao recurso (propriedades intelectuais, direitos autorais e outros direitos de propriedade).

Para MCCRAY [32], esse conjunto de elementos do *Dublin Core* provê simplicidade, interoperabilidade semântica, útil por meio de domínios potencialmente muito diferentes.

TABELA 3.2 - EXEMPLIFICAÇÃO DOS ELEMENTOS DUBLIN CORE

<b>Tipo de Elemento <i>Dublin Core</i></b>	<b>Aplicação no documento</b>
<i>Title</i>	Modelo de Gerência de dados não convencionais em dispositivos móveis através do uso de metadados
<i>Creator</i>	Juliano Lucas Gonçalves
<i>Subject</i>	Metadados, Padrões de metadados, Dispositivos móveis, Computação Móvel.
<i>Contributor</i>	Maria Salete Marcon Gomes Vaz
<i>Publisher</i>	Universidade Federal do Paraná
<i>Description</i>	Os dispositivos móveis e suas aplicações estão utilizando cada vez mais quantidades crescentes de novos tipos de dados, tais como, vídeos, imagens, e músicas. O controle de quantidades grandes de dados introduz um desafio da usabilidade que não pode ser resolvido sem desenvolver novas maneiras de controlar o tratamento desses dados. O uso do contexto de metadados é uma maneira de fazer o controle desses dados
<i>Date</i>	09/07/2005
<i>Type</i>	Text
<i>Format</i>	Pdf
<i>Identifier</i>	<a href="http://www.inf.ufpr.br/~jlucas/dissertacao.pdf">http://www.inf.ufpr.br/~jlucas/dissertacao.pdf</a>
<i>Language</i>	Português
<i>Rights</i>	Departamento de Informática da Universidade Federal do Paraná.



Na Tabela 3.2 é mostrada uma aplicação dos elementos *Dublin Core* para estruturar as informações do presente trabalho.

Os metadados podem ser armazenados no *Dublin Core* de três formas, como mostra a Tabela 3.3.

TABELA 3.3 - FORMAS DE ARMAZENAMENTO DE METADADOS EM DUBLIN CORE.

<b>Forma de Armazenamento</b>	<b>Características</b>
<i>HTML (Hyper Text Markup Language)</i>	<ul style="list-style-type: none"> <li>- Usa a tag &lt;META&gt; para a definição de metadados</li> <li>- Mantém metadados dentro do documento</li> <li>- É o mais usado atualmente</li> </ul>
<i>XML (eXtended Markup Language)</i>	<ul style="list-style-type: none"> <li>- Permite maior estruturação</li> <li>- É mais trabalhosa: pode requerer programação</li> </ul>
<i>RDF (Resource Description Framework)</i>	<ul style="list-style-type: none"> <li>- Arquitetura de metadados voltada para a descrição de recursos</li> <li>- Poder de expressão mais adequada</li> <li>- Participantes do Dublin Core ajudam a definir</li> </ul>

O HTML é muito utilizado por ser uma representação mais simples e por manter os metadados dentro do documento, o que para dispositivos móveis torna-se mais fácil de gerenciar, pois não há a necessidade da utilização de SGBDs. A seguir é mostrada a definição geral de metadados Dublin Core em html.

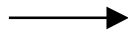
```
<meta name ="PREFIX.ELEMENT_NAME"
content = "ELEMENT_VALUE">
```

onde PREFIX.ELEMENT\_NAME = tipo do elemento Dublin Core a ser inserido (creator, title, subject, etc);  
e ELEMENT\_VALUE = valor do respectivo elemento Dublin Core a ser inserido (Juliano, padrão de metadados, metadados).

Alguns exemplos da aplicação de metadados Dublin Core em Html:

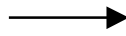
1- <meta name = "DC.Creator"  
content = "Lucas, Juliano">

2- <meta name = "DC.Title"  
content = "The Communist Manifesto">  
<meta name = "DC.Creator"  
content = "Marx, K.">  
<meta name = "DC.Creator"  
content = "Engels, F.">



Repetição de elementos  
Dublin Core (Creator).

3- <meta name = "DC.Creator"  
type = "email"  
content = "jlucas@inf.ufpr.br">  
<meta name = "DC.Creator"  
type = "address"  
content = "Florianópolis, Curitiba">



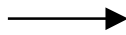
Utilização do elemento  
Type.

4- <meta name = "DC.Title"  
content = "Padrão de metadados para  
Dispositivos Móveis">  
<meta name = "DC.Title"  
content = "Metamida - Um Modelo de  
Metadados na Indexação e Recuperação  
de Objeto Multimídia">



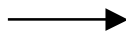
Utilização do elemento  
Title.

5- <meta name = "DC.Subject"  
content = "metadados, dispositivos móveis">



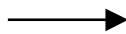
Utilização do elemento  
Subject.

6- <meta name = "DC.Description"  
lang = "pt"  
content = "O autor fala sobre metadados e  
seus padrões. Propõe a criação  
de um padrão de metadados para  
dispositivos móveis, ressaltando  
suas características e limitações">



Utilização do elemento  
Description

7- <meta name = "DC.Publisher"  
content = "Universidade Federal do Paraná">



Utilização do elemento  
Publisher.

- 8- `<meta name = "DC.Contributor"  
content = "Maria Salete M. G. V.">` → Utilização do elemento Contributor.
- 9- `<meta name = "DC.Date"  
content = "2005">  
<meta name = "DC.Date"  
content = "2005-05-14">` → Utilização do elemento Date.
- 10- `<meta name = "DC.Format"  
content = "text/xml">  
<meta name = "DC.Format"  
content = "video/mpeg; 14 minutes">` → Utilização do elemento Format.
- 11- `<meta name = "DC.Identifier"  
content = "http://www/inf.ufpr.br/  
~jlucas/dissertacao.pdf">` → Utilização do elemento Identifier.
- 12- `<meta name = "DC.Language"  
content = "en">` → Utilização do elemento Language.
- 13- `<meta name = "DC.Rights"  
content = "Departamento de  
Informática UFPR">` → Utilização do elemento Rights.

Como mostrado no Exemplo 2, os elementos de metadados do formato Dublin Core permitem que todos os campos sejam repetidos, opcionais e podendo ter associações, quantas conexões forem necessárias. As diferentes comunidades de usuários e as distintas áreas de aplicação requerem diferentes elementos e níveis de complexidade, tornando necessário modelos de metadados complementares que necessitem de uma arquitetura que os acomode. Esta arquitetura deve modular os distintos tipos de metadados.

A falta de extensividade do formato Dublin Core ocasiona problemas no caso de ser necessário acrescentar elementos adicionais à sua estrutura. Neste sentido torna-se preciso conectar a descrição Dublin Core com outros tipos de descritores propostos por outras instituições, preocupadas com a questão da recuperação da Informação eletrônica na Internet. O consenso destes trabalhos deve convergir na necessidade de uma

arquitetura que poderia acomodar a diversidade de modelos e níveis de descrição que caracteriza o mundo heterogêneo das fontes eletrônicas.

De uma forma geral o padrão Dublin Core destaca-se pela sua simplicidade e interoperabilidade, podendo ser utilizado para descrever dados não convencionais em dispositivos móveis, uma vez que o mesmo apresenta um conjunto de elementos capazes de abstrair informações importantes dos objetos tais como: imagens, sons e vídeos. Essas características facilitam o tratamento das informações, facilitando, igualmente, sua busca e recuperação.

### 3.2 Padrão MARC

O padrão *MARC (Machine Readable Card)*, por ser um padrão muito antigo, criado em 1960, vários outros padrões de metadados basearam-se nele. Ele é utilizado para catalogação bibliográfica [31]. Ao contrário do Dublin Core que possui um número pequeno de atributos com o objetivo de ser de fácil utilização, o padrão MARC possui um número bem maior de atributos, pois seu objetivo é ser completo.

Segundo BALDONADO [4] o padrão MARC é mais utilizado em bibliotecas convencionais, enquanto o padrão Dublin Core nas bibliotecas digitais.

O padrão MARC ao contrário do Dublin Core trabalha com registros ao invés de elementos específicos, sendo que o registro é formado pelos campos *tag*, *indicator* e *subfield*.

A Tabela 3.4 contém algumas das *tags* mais utilizadas no padrão MARC, uma vez que são cerca de 856 *tags* [4].

TABELA 3.4 - MARC TAGS

Valor ( <i>Tags</i> )	Representação
010	Marks the Library of Congress Control Number (LCCN)
020	Marks the International Standard Book Number (ISBN)
100	•marks a personal name main entry (author)

245	•marks the title information (which includes the title, other title information, and the statement of responsibility)
250	marks the edition
260	marks the publication information
300	marks the physical description
440	marks the series statement/added entry
520	marks the annotation or summary note
650	marks a topical subject heading
700	marks a personal name added entry (joint author, editor, or illustrator)

A seguir serão mostrados os campos que formam o registro do padrão MARC.

- **MARC indicators:**

- 2 caracteres que seguem cada tag
- Quando uma *tag* não possui indicador, isto é, “undefined”, a posição fica em branco ou com o caracter “#”.
- Cada posição é um número de 0 a 9. Mesmo que dois indicadores juntos pareçam um único número, cada dígito tem seu próprio significado.

Exemplo de Representação de um Registro MARC:

245 14 \$a The emperor's new clothes / \$c  
 adapted from Hans Christian Andersen  
 and illustrated by Janet Stevens.

- O primeiro indicador (valor 1) no campo título (245) significa que deve haver uma entrada separada para o título no catálogo
- O segundo indicador (valor 4) significa que os quatro primeiros caracteres do título podem ser desconsiderados no processo de ordenação (“The ”).

- **MARC Subfields:**

- Códigos: lowercase, precedidos por um delimitador (\$, \_ , @, ...)

- O *tag* 300 (descrição física) apresenta três sub-campos que indicam respectivamente.
  - A extensão (número de páginas);
  - Detalhes físicos (apresenta ilustração);
  - Dimensões (centímetros);

Exemplo 300 ## \$a 675 p : \$b ill ; \$c 24 cm.

Em 1998 o padrão MARC sofreu uma unificação entre os padrões USMARC e CANMARC, utilizados nos Estados Unidos e Canadá, respectivamente. Essa unificação deu origem ao MARC21. Em 1999 a Biblioteca Nacional do Canadá deu início a uma produção francesa do MARC21.

O padrão MARC21 ainda é utilizado atualmente nas bibliotecas de inúmeros países. Porém, vem sofrendo uma queda frente às facilidades disponibilizadas pelo padrão Dublin Core.

O padrão MARC, ao contrario do Dublin Core, possui um grande número de elementos, armazenados em registros. Esse padrão é muito útil quando a quantidade de informações é muito grande, o que não ocorre nos dispositivos móveis, uma vez, que é possível descrever os objetos utilizando-se de um número pequeno de metadados, o que tornaria inviável sua utilização.

### **3.3 Padrão FGDC/CSDGM**

O Padrão *CSDGM – Content Standard for Digital Geospatial Metadata*, do *FGDC – Federal Geographic Data Committee*, começou a ser criado em 1992 e em 1994 foi concluído. Em 1998 teve sua primeira revisão. O mesmo teve como objetivo fornecer um conjunto de terminologias e definições comuns para a documentação de dados espaciais digitais [11].

Esse padrão define:

- Os nomes dos elementos de dados e dos elementos compostos (grupos de elementos compostos);
- As definições desses elementos (de dados e compostos);
- Informações sobre os valores que serão fornecidos para os elementos de dados;
- O grau de obrigatoriedade da informação (obrigatória, obrigatória sobre certas condições e opcional);

A organização do Padrão FGDC/CSDGM é composta de uma hierarquia de elementos de dados e elementos compostos. O padrão tem como ponto inicial à seção zero (metadado). Essa seção possui um conjunto de seções numeradas, onde cada uma representa um elemento composto, cada seção-elemento pode ser definida por outros elementos (compostos ou elementos de dados).

Cada seção é iniciada pelo nome e definição do elemento composto que a define. O nome e a definição são seguidos por regras de produção que definem o elemento composto em termos de elementos de dados ou de outros elementos compostos intermediários. Um panorama geral do padrão [11] pode ser visto na Figura 3.1,

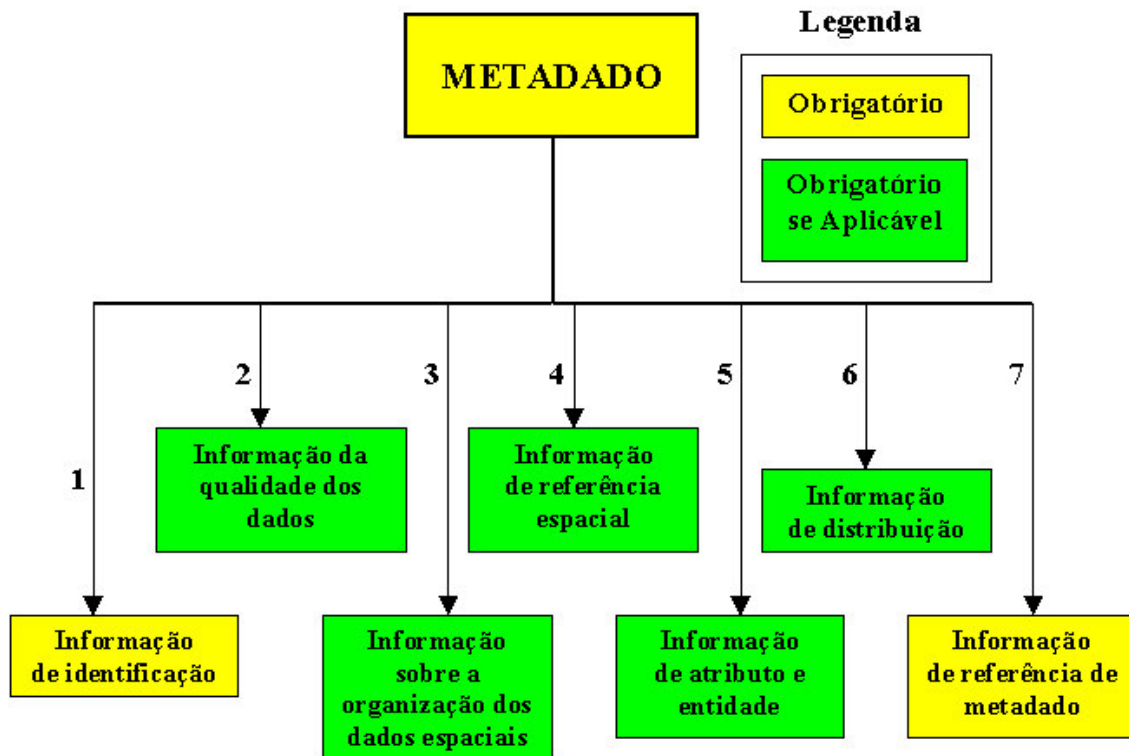


FIGURA 3.1 - PANORAMA ORGANIZACIONAL DO PADRÃO FGDC/CSDGM

- **Metadados**

Os metadados são dados sobre o conteúdo, qualidade, condições e outras características de dados.

Metadata =

Identification\_Information +  
 0{Data\_Quality\_Information}1 +  
 0{Spatial\_Data\_Organization\_Information}1 +  
 0{Spatial\_Reference\_Information}1 +  
 0{Entity\_and\_Attribute\_Information}1 +  
 0{Distribution\_Information}n +  
 Metadata\_Reference\_Information

- **Informação de Identificação**

A informação de identificação corresponde às caracterizações básicas sobre um conjunto de dados, as quais incluem título, informação cronológica, área



geográfica coberta, palavras chave associadas, origem dos dados, regras para aquisição e uso dos dados etc. Na seqüência é descrito o escopo deste tipo de informação.

```
Identification_Information =
    Citation +
    Description +
    Time_Period_of_Content +
    Status +
    Spatial_Domain +
    Keywords +
    Access_Constraints +
    Use_Constraints +
    (Point_of_Contact) +
    (1{Browse_Graphic}n) +
    (Data_Set_Credit) +
    (Security_Information) +
    (Native_Data_Set_Environment) +
    (1{Cross_Reference}n)

Citation =
    Citation_Information (see section 8 for production
rules)

Description =
    Abstract +
    Purpose +
    (Supplemental_Information)
```

Dentre as informações contidas no escopo mostrado anteriormente destacam-se a citação e a descrição que contêm o *abstract* do conjunto de dados a serem descritos.

- **Informação da qualidade dos dados**

Possui diversas informações sobre a qualidade do conjunto de dados, incluindo precisão, fidelidade, critérios de seleção, generalizações, consistência dos dados, definições utilizadas e metadados sobre as fontes de dados.

- **Informação sobre a Organização dos dados espaciais**

A informação sobre a organização dos dados espaciais identifica os mecanismos utilizados para a representação dos dados espaciais, além da sua caracterização como ponto, polígono, entre outros.

Atualmente o padrão possui três mecanismos para a representação da informação espacial. Um genérico para representar os dados em formato *raster* e dois para representar dados em formato vetorial, *STDS - Spacial Data Transfer*, do Departamento de Comércio do Estados Unidos e *VPF - Vector Product Format* do Departamento de Defesa, também dos Estados Unidos.

- **Informação de Referência Espacial**

A informação de referência espacial representa os sistemas de projeção e coordenadas utilizadas, tais como nome da projeção, parâmetros, entre outros.

A informação de referência espacial é uma seção muito importante do ponto de vista espacial, pois permite a identificação do posicionamento dos objetos espaciais em relação a um padrão, impedindo que ocorram distorções.

- **Informação de Atributo e Entidade**

A informação de atributo e entidade representa a descrição do conteúdo das informações do conjunto de dados incluindo tipos de entidades seus atributos e os domínios que seus atributos podem receber.

- **Informação de Distribuição**

A informação de distribuição contém informações sobre o distribuidor do conjunto de dados e sobre como obtê-los. Normalmente essa seção contém, além do distribuidor, os possíveis meios de transferência de objetos (modem, fax, e-mail, etc) e formatos de transferência de objetos (TIFF, ASCII, formato de exportação de um determinado SIG etc);

- **Informação de referência de Metadado**

A informação de referência de metadado contém informações sobre a última atualização do metadado, a pessoa responsável, a última e a próxima revisão, restrições de segurança e acesso etc;

Existem ainda mais três seções que servem como complemento e nunca são utilizadas sozinhas.

- **Informação da citação**

A informação da citação é uma seção especial que especifica a referência recomendada a ser utilizada para o conjunto de dados. É uma seção usada por outras seções do padrão.

- **Informação do período de tempo**

A informação do período de tempo é uma seção especial que representa a data e hora de um determinado evento.

- **Informação do contato**

A informação do contato é uma seção especial que identifica os responsáveis pelo conjunto de dados. (pessoas, organizações, etc);

O padrão FGDC/CSDGM é um padrão mais complexo, possuindo inúmeras seções, cada uma delas apresentando um conjunto de elementos a serem descritos. Isso muitas vezes gera repetição de informação ocupando muita memória, esse motivo torna inviável a utilização desse padrão para dispositivos móveis, visto que, esses dispositivos apresentam limitações de memória. Além disso, ele não especifica a forma na qual a informação está organizada em um computador, nem determina a forma como essa informação é transmitida e apresentada ao usuário, dificultando assim, sua utilização.

### 3.4 Padrão TEI

O Padrão *TEI – Text Encode Initiative* surgiu na Conferência de Poughkeepsie em 1987. Um esforço internacional de pesquisa que teve como objetivo definir um conjunto genérico de normas para representação de material textual em forma eletrônica [40]. Ele possui quatro seções principais [40], como mostra a Tabela 3.5.

TABELA 3.5 - SEÇÕES PRINCIPAIS DO PADRÃO TEI

Seções	Representação
<b>&lt;fileDesc&gt;</b>	Descrição bibliográfica completa, a partir da qual um usuário pode derivar uma citação bibliográfica apropriada, ou um bibliotecário ou arquivista pode usar na criação de uma entrada num catálogo gravando sua presença em uma biblioteca ou arquivo. Inclui informações sobre a fonte que originou o texto
<b>&lt;encodingDesc&gt;</b>	Descrição do relacionamento entre o texto e sua(s) fonte(s). Permite detalhes sobre: se o texto foi normalizado durante a transcrição, como o codificador resolveu ambigüidades na fonte, que níveis de codificação e análise foram aplicados, etc.
<b>&lt;profileDesc&gt;</b>	Informação classificatória e contextual sobre o texto, tais como: assunto, a situação em que foi produzido, os indivíduos descritos por ele ou que participaram em sua produção, etc.
<b>&lt;revisionDesc&gt;</b>	Permite ao codificador prover um histórico de mudanças feitas durante o desenvolvimento do texto eletrônico.

A seguir são mostrados os elementos do padrão TEI:

**<TEIHEADER>** informação descritiva e declarativa caracterizando uma página com título eletrônico a cada texto TEI-compatível.

**<FILEDESC>** descrição bibliográfica completa.

**<TITLESTMT>** agrupa informação sobre o título de um trabalho e aqueles responsáveis pelo seu conteúdo intelectual.

**<TITLE>** contém o título de um trabalho, seja ele artigo, livro, jornal ou série, incluindo quaisquer títulos ou subtítulos alternativos.

**<AUTHOR>** contém o nome do(s) autor(es), pessoa ou corporação; o principal elemento responsável por qualquer item bibliográfico.

**<RESPSTMT>** detalhamento sobre algum responsável pelo conteúdo intelectual, quando os elementos específicos p/ autores, editores, etc. não são suficientes ou não se aplicam.

**<RESP>** frase descrevendo a natureza da responsabilidade intelectual da pessoa.

**<NAME>** contém um nome próprio ou frase nominal.

**<EXTENT>** tamanho aproximado de um texto eletrônico conforme gravado em algum meio de armazenamento, especificado em quaisquer unidades que sejam convenientes.

**<PUBLICATIONSSTMT>** agrupa informação relacionada à publicação ou distribuição de um texto.

**<PUBLISHER>** provê o nome da organização responsável pela publicação ou distribuição de um item bibliográfico.

**<DATE>** data em qualquer formato

**<AVAILABILITY>** fornece informação sobre a disponibilidade de um texto, por exemplo, quaisquer restrições em seu uso ou distribuição, seu status de copyright, etc.

**<ADDRESS>** contém um endereço postal ou outro, por exemplo: de um publicador, uma organização ou um indivíduo.

**<ADDRLINE>** contém uma linha de um endereço.

<SERIESSTMT> agrupa informação sobre a série, se existir, à qual a publicação pertence.

<SOURCEDESC> provê uma descrição bibliográfica da cópia do(s) texto(s) a partir da(s) qual(is) a versão eletrônica foi derivada ou gerada.

<BIBLFULL> contém uma citação bibliográfica altamente estruturada, na qual todos os componentes do arquivo TEI estão presentes.

<PUBPLACE> contém o nome do local onde um item bibliográfico foi publicado.

<ENCONDINGDESC> documenta o relacionamento entre o texto eletrônico e a fonte a partir da qual ele foi derivado.

<EDITORIALDECL> provê detalhes de princípios editoriais e práticas aplicadas durante a codificação do texto.

<TAGSDECL> provê informações detalhadas sobre os tags aplicados ao documento SGML.

<TAGUSAGE> provê informação sobre o uso de um elemento específico dentro de um <text>.

<REVISIONDESC> resume o histórico de revisão do texto.

<CHANGE> resume uma correção ou mudança específica feita a uma versão particular do texto eletrônico que é compartilhado por vários pesquisadores.

<ITEM> contém um item de uma lista.

Segundo BURNARD [6] as principais características do padrão TEI são:

- Formato de metadados SGML;
- É um esquema mais descritivo do que prescritivo;
- Permite que a classificação do conteúdo do documento seja realizada em qualquer grau de granularidade;
- O metadado é fortemente acoplado (fica dentro do documento);
- É extensível podendo se construir *tags*, renomear e excluir elementos existentes, definir estrutura básica, modificar modelos de conteúdo, etc;

É amplamente utilizado em várias partes da comunidade de pesquisa, em particular aquelas engajadas na criação de bibliotecas eletrônicas, em publicações eletrônicas (por exemplo, edições estudantis) e na criação de corpo de linguagens para uso em processamento de linguagem natural.

Nos EUA, projetos líderes de biblioteca eletrônica, como os das universidades de Virginia, Michigan e Indiana utilizam TEI. É muito presente no meio acadêmico. Na Europa o público é ligeiramente diferente, englobando uma série de empresas comerciais de editoração eletrônica [7].

Como objetivos futuros esse padrão ainda pretende criar:

- Extensões para dados geoespaciais, informações históricas sobre arte, descrição de manuscritos;
- Desenvolvimentos em criticismo textual, para incluir tags para bibliografia analítica, codicologia e descrição física de fontes primárias;
- Desenvolvimentos em codificação de dicionários históricos;
- Desenvolvimentos em Sistema de Escrita e problemas com conjuntos de caracteres.

A principal característica do padrão TEI é o forte acoplamento do metadado ao documento. Isso é muito importante quando se trabalha com textos, pois facilita sua atualização. Porém em dispositivos móveis os metadados ficam separados do objeto, isso devido às capacidades de armazenamento e processamento limitadas, ou seja, o objeto só será acessado quando o metadado solicitado for encontrado, com o objetivo de não sobrecarregar o processamento e memória desnecessariamente.

### **3.5 Padrão GILS**

O Padrão GILS - *Government Information Location Service*, Em 1992 por meio de um estudo conduzido pelo *Office of Management and Budget, the National Archives* e pela *Records Administration* foi recomendado o uso do Z39.50 como protocolo apropriado à

recuperação de informação. Em 1994 foi aprovado pelo Departamento de Comércio como padrão federal para processamento de informações. E finalmente em 1995 foi obrigatória (através de uma lei) sua utilização em todos departamentos governamentais dos Estados Unidos da América [19].

O padrão GILS corresponde à (Figura 3.2):

- Serviço de coleta e disseminação de dados;
- Uma arquitetura para projeto físico e implementação
- Um esquema para descrição de metadados

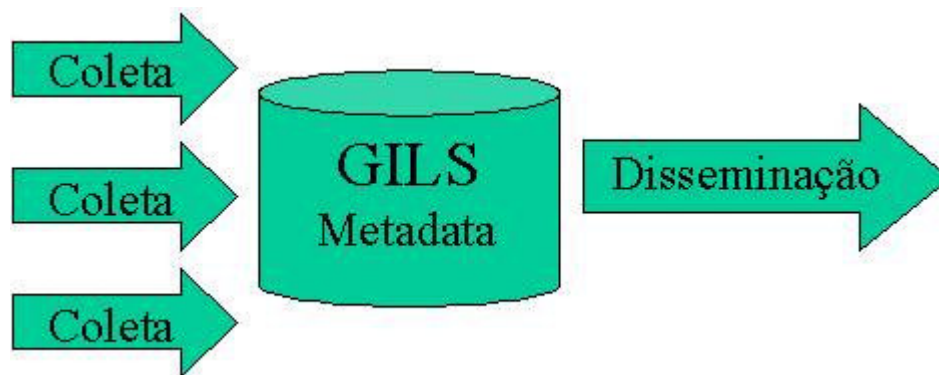


FIGURA 3.2 - ESTRUTURA GILS

Utiliza:

- SGML - Standard Generalized Markup Language é uma linguagem que provê a definição do conteúdo;
- Z39.50 – protocolo de consulta que permite que os dados sejam recuperados de sistemas remotos;
- SQL - *database management Language*, linguagem de consulta da base de dados;
- Gils – Versão 2, especificação de metadados;

O propósito do GILS é ajudar o público a localizar e acessar informações/fontes de informação. Uma fonte de informação é qualquer coisa que provê informação. Isso abrange desde livros até arquivos de som na web e até mesmo os serviços do tipo *heldesk*.



Segundo CHRISTIAN [8] o GILS define cerca de 70 atributos, denominados elementos centrais. Além desses cerca de mais 100 elementos herdados do protocolo Z39.50, sendo que esses padrões podem ser:

- Obrigatório (title, originator, language of record, etc) ou opcional (contributor, abstract, date of publication, etc);
  - Repetível ou não-repetível
  - Controlado ou não controlado
  - Elemento de agrupamento ou não-agrupamento;
- 
- **Obrigatório** - deve existir em todo registro GILS.
  - **Opcional** - só precisa ser incluído num registro, se for adicionar alguma informação útil.
  - **Repetível** - pode ser incluído diversas vezes em um registro, pode ser um elemento ou sub-elemento.
  - **Não repetível** - só pode ser usado uma vez em um registro GILS
  - **Controlado** – o dado está em uma forma específica ou vem de uma lista de valores aceitáveis.
  - **Não controlado** – não está em um formulário específico ou valores restritos a algum domínio.
  - **Elemento de grupamento** – contem sub-elementos; não pode conter outros elementos ou formulário de dados. Não podem ser descritos como controlled ou not controlled já que não contem dados.
  - **Elemento de não agrupamento** - não é um elemento de grupamento; pode conter dados. Se um elemento não for classificado como elemento de grupamento então é um elemento de não grupamento.

A seguir serão mostrados alguns elementos obrigatórios:

- **Title** – Título da fonte de informação.

- **Originator** – nome da agência/departamento que criou o conteúdo da fonte de informação.
- **Language of Resource** – Língua(s) da fonte de informação
- **Record Source** – unidade(s) organizacional que criou ou que modificou por último o registro GILS.
- **Language of Record** – Língua do registro.
- **Date of Last Modification** – Data na qual o registro foi criado ou da última modificação.

A seguir serão mostrados alguns elementos opcionais:

- **Contributor** – qualquer outro autor que não o Originator.
- **Date of Publication** – data em que a fonte foi publicada, enviada ou atualizada dependendo do seu formato.
  - Date of Publication Structured – forma estruturada da data de publicação. Útil em buscas. (sub-elemento de Date of Publication)
  - Date of Publication Textual – Forma textual da data de publicação.
- **Place of Publication** – cidade da publicação, com mais informações se necessário.
- **Abstract** – sumário do conteúdo da fonte de informação.
- **Spatial Domain** – Área geográfica coberta pelo fonte de informação. Agrupa Bounding Coordinates e Place.
  - Bounding Coordinates – definição da área geográfica pelos pontos oeste, leste, norte e sul. Agrupa os elementos West Bounding Coordinate, East Bounding Coordinate, North Bounding Coordinate e South Bounding Coordinate.
  - Place – descrição textual da região coberta. Elemento de Agrupamento para Place Keyword Thesaurus e Place Keyword
- **Availability** – Como conseguir acessar a fonte de informação. Agrupa Medium, Distributor, Resource Description, Order Process, Technical Prerequisites, Available Time Period, e Available Linkage.

- **Medium** – mídia da fonte de informação. Podendo ser : Online, Paper , Braille Large Print, Photographs , Microfiche, Microfilm, Film, Tape, Disk, Optical, Mixed media, Other
- **Distributor** – distribuidor da fonte de informações. Agrupa Groups Distributor Name, Distributor Organization, Distributor Street Address, Distributor City, Distributor State or Province, Distributor ZIP or Postal Code, Distributor Country, Distributor Network Address, Distributor Hours of Service, Distributor Telephone, and Distributor Fax.
- **Resource Type** – Tipo de recurso a ser descrito. Agrupa Resource Type Thesaurus and Resource Type Term.
  - Resource Type Thesaurus – lista controlada ou thesaurus utilizada para determinar o tipo de fonte. Ex: Dublin Core Resource Types

A seguir é mostrado um exemplo de utilização do padrão GILS:

**Title** metamídia – um modelo de metadados na indexação e recuperação de objetos multimídia

**Originator** Maria Salete Marcon Gomes Vaz

**Date of publication** 2000 12 20

**Language Resource** pt

**Abstract:** Um banco de dados convencional é um repositório de dados, que contém valores numéricos, alfanuméricos, booleanos e datas. Um banco de dados multimídia, adicionalmente, contém imagens gráficas, clipes de vídeo, arquivos de som, textos, entre outros.

**Subject term uncontrolled:** metadados, banco de dados, dados multimídia.

**Availability**

Médium: online

**Distributor:**

Name: Maria Salete Marcon Gomes Vaz

Organization: Universidade Federal de Pernambuco

Network Address: [www.ufpe.br](http://www.ufpe.br)

Telephone: 41- 22225588

**Order Process**

Technical Prerequisites: PDF, Microsoft word 6.0

**Record Source:** Universidade Federal de Pernambuco.

**Laguage of Record:** pt.

**Data of last modification:** 2001 03 10

Atualmente no Canadá [23], foi formado um subgrupo - *Eletronic Document Standardas Working Group* com a missão de avaliar o GILS com possível padrão para Tecnologia de Informação do Tesouro Nacional. O GILs foi considerado apropriado e um projeto piloto foi iniciado.

O padrão GILS é utilizado quando a quantidade de informações é muito grande, fato que não ocorre em dispositivos móveis, porque a quantidade de informações é relativamente pequena devido às restrições de memória desses dispositivos. Conseqüentemente, não é necessário um grande número de elementos para descrevê-las, tornando inadequada à utilização desse padrão.

## 4. DISPOSITIVOS MÓVEIS

### 4.1 Introdução

Os dispositivos móveis são aqueles que podem ser operados à distância ou sem fio. Dispositivos que podem ser desde um simples Bip, até os mais modernos Pockets (dispositivos móveis com sistema operacional). Segundo uma projeção divulgada pela INFOEXAME [28] em 2003 no mundo havia cerca de 1 bilhão de dispositivos móveis, sendo que a previsão para 2007 é de que ultrapasse os 2 bilhões. O início da história dos dispositivos móveis pode ter seu marco em janeiro de 1992, quando o CIO da Apple John Sculley falava sobre o novo sonho de consumo da época. O *Free Online Ditionary of Computing* descreve o PDA (Figura 3) como um pequeno computador de mão utilizado para escrever notas, listar apontamentos e auxiliar na organização pessoal. [55].



FIGURA 4.1 - EXEMPLO DE UM PDA APPLE DE 1992

A Microsoft em 1996 foi a 1ª empresa a lançar um sistema operacional para dispositivos móveis, o qual foi chamado de *Windows CE*. Ela criou um consórcio com 40 companhias para fabricarem dispositivos compatíveis com sua nova plataforma. Porém não teve muito sucesso, pois sua interface era muito complicada e consumia muitos recursos dos dispositivos, os quais eram limitados na época.

Em 2000, com o lançamento dos *Pockets PC 2000* (Figura 4), é lançada juntamente a versão 3 do *Windows CE*, a qual chega bem parecida com o ambiente

Windows desktop, agora mais bem elaborada e preparada para trabalhar com dispositivos móveis [55].

Mais uma atualização do sistema operacional garantiu o sucesso do *Windows CE*, nos iPaq Compaq que firmariam a presença da Microsoft no mundo dos dispositivos móveis. Atualmente o Windows CE é o sistema operacional utilizado nos Pockets PC iPaq da Compaq, os modelos Jornada da HP, Axim Pocket da Dell, Pocket Loox da Fujitsu, o lançamento Maestro da Audiovox, vários modelos Toshiba como os E400, modelos Hitashi como o HPW200EC, tendo ainda outros modelos entre outros fabricantes.

Na categoria de celulares, ou *Smartphones* como são chamados, que suportam o Sistema Operacional *Windows CE* há os Modelos Nexio, SPH i500 e i700 e os modelos QCP6035 e 7135 da Samsung, Kyocera [43], Siemens SX56 [47], Motorola MPx [36], Nokia 9500 [37], Sony Ericsson p800 [18], e dependendo da região ainda há outros modelos. No Brasil apenas temos lançado comercialmente pelas operadoras o Gradiente Partner [24] e o Motorola MPX220 [36], mas com preços alcançáveis por poucos.



FIGURA 4.2 - POKETS PC COM A PLATAFORMA WINDOWS CE

Estamos passando por uma transformação tecnológica muito grande nos últimos anos. Paradigmas que antes impediam que uma nova tecnologia chegasse ao mercado estão sendo quebrados, e assim estamos inovando cada dia mais.

No mercado atual, a capacidade de uma organização se adaptar a novas situações mercadológicas utilizando instrumentos tecnológicos e gerenciais para colocar o corpo diretivo em condições de raciocinar, compreender relações entre fatos, descobrir significados e reconhecer a verdade é uma questão de sobrevivência. A tecnologia vem conduzindo esse processo de transformação de dados em informações, e essas em conhecimento. Para tal, será necessário saber pelo menos quais informações estão disponíveis, onde estão localizadas, de que forma são manipuladas e para onde podem ser distribuídas.

A manipulação de grandes volumes de dados, dentro de algumas empresas, tem se tornado uma realidade extremamente comum. Acoplada a ela surgiu uma crescente demanda de uma forma que garantisse o efetivo uso da informação de maneira produtiva e inequívoca.

Existem hoje pessoas que trabalham em ambientes empresariais não tradicionais. Tais ambientes podem ser uma série de locais de trabalho temporários, como no caso dos representantes de vendas ou dos executivos em viagem, ou pode ser que a natureza de uma tarefa exija que o trabalhador esteja sempre se deslocando, como no trabalho em fábricas, entregas de pacotes, serviços de campo ou profissões ligadas à saúde.

Nos últimos cinco anos, vários aparelhos portáteis de informação, de *PC's* manuais a *Pocket PCs*, foram disponibilizados para auxiliar essa força de trabalho móvel. Esses dispositivos não apenas ajudam no gerenciamento de compromissos e contatos como, também, representam uma ferramenta para substituição de processos comerciais feitos em papel por aplicativos baseados em formulários.

Uma maior eficácia e precisão na captura rápida de dados em um dispositivo computacional podem resultar em maior produtividade dos funcionários, maior rapidez na geração de relatórios comerciais para tomadas de decisão, maior satisfação do cliente e custos operacionais reduzidos (por dispensar a entrada de dados por pessoal específico).

Segundo Depiné [13] os dispositivos sem fio oferecem uma conectividade que outros dispositivos não possuem. Em poucos anos o desenvolvimento de aplicações para

esses equipamentos tende a aumentar drasticamente, utilizando-se dos recursos que os mesmos têm a oferecer.

## **4.2 Vantagens dos Dispositivos Móveis**

Dispositivos móveis são completamente diferentes de *desktops*. As pessoas não usam aplicações móveis para acessar a mesma informação ou executar as mesmas funções que elas fariam com um computador fixo. Isso se deve às restrições do hardware do dispositivo móvel (tela menor, entrada de dados limitada, energia limitada, mobilidade, características do ambiente de comunicação sem fio, largura de banda menor, taxa de bits errados maior), e pagamento pelos serviços. Logo, a habilidade, capacidade, e disponibilidade de um usuário acessar e processar informações são menores. Isto significa que um dos aspectos mais importantes no desenvolvimento de aplicações móveis é a usabilidade, ou seja, de nada serve ter softwares e aplicativos extremamente poderosos se não for possível utilizar todas as suas funções.

Muitos trabalhos hoje em dia exigem do profissional elasticidade e versatilidade. Com o objetivo de auxiliar estes profissionais com elementos da área de TI, eles precisam dispor dos mais modernos aparelhos portáteis existentes no mercado.

Segundo SHAEFER [45], do ponto de vista empresarial, os dispositivos móveis são ótimos geradores de informação, podendo ser utilizados na automatização do processo até nas coletas de informações estratégicas, pois com suas reduzidas dimensões podem estar sendo transportados e estar presentes em todas as situações em que um profissional dessa área pode atuar.

Para estes dispositivos, cada vez mais surgem novos aplicativos exclusivos para o ambiente. Junto a isto cada vez mais os usuários estão usufruindo as facilidades de um mundo interligado por redes sem fio onde o usuário pode ter a qualquer hora e em qualquer lugar, a informação que desejar bastando estar apenas conectado a estas redes.

Um problema em se tratando de dispositivos móveis é a falta de uma plataforma padrão para desenvolvimento de softwares e aplicativos, o que acabaria assim como o excesso de linguagens de programação existentes atualmente.



Outro aspecto importante dos dispositivos móveis diz respeito aos SGBDs móveis, ou sejam, sistemas gerenciadores de base de dados para esses dispositivos, pois é preciso ter acesso aos dados, bem como poder gerenciá-los mais facilmente.

### **4.3 Sistemas Gerenciadores de Banco de Dados Móveis**

Os maiores fabricantes de sistemas gerenciadores de banco de dados atuam no mercado da computação móvel, oferecendo um conjunto de produtos que incluem software para o desenvolvimento de aplicações e sistemas de banco de dados adaptados para este ambiente, os quais buscam uma integração dos dados destas aplicações com os servidores de bancos de dados localizados na rede fixa. Uma característica comum desses produtos está no fato de serem desenvolvidos, principalmente, para atuarem na desconexão dos clientes, como bancos de dados locais, sempre necessitando de uma carga antecipada dos dados (*hoarding*) para dar suporte às aplicações. Quando conectados à rede fixa, executam a sincronização com seus servidores, evitando trabalharem conectados como se estivessem em um ambiente de banco de dados distribuído.

Segundo uma pesquisa do *Gartner Dataquest* de 2002, publicada na *Pc Magazine* [30], a Sybase com o seu produto *Sybase SQL Anywhere* dominava, aproximadamente, 65% do mercado de banco de dados móveis. Outros produtos relevantes no mercado são o Oracle Lite da Oracle, o DB2 Everyplace da IBM, o SQL Server 2000 CE da Microsoft e o SQLBase da Gupta Technologies. Todos esses produtos são considerados como *Small Footprint DBMS*, ou seja, SGBDs projetados para minimizar a utilização da memória dos hosts móveis. Alguns desses produtos são mostrados na Seção 4.3.1.

#### **4.3.1 Sybase UltraLite**

O produto *SQL Anywhere Studio* da Sybase é um pacote de software que prove a gerência de dados e possui uma solução de banco de dados embutida chamada de *Sybase UltraLite*. Além do seu sincronismo com os SGBDs da Sybase, pode se comunicar com produtos de outros fabricantes [49]. Suas principais características são:

- Pode ser utilizado nas plataformas Windows 95/98/Me, NT, 200, XP, Windows CE (Pocket PC/Handheld PC), Palm Computing Platform, Wind River VxWorks;
- Acesso aos dados através de JDBC, SQL embutida e API baseada em C++;
- As funcionalidades da SQL incluem processamento de transações, integridade referencial e operações de junções de várias tabelas;
- Suporta *Binary Large Object* (BLOB);
- Prove serviço de encriptografia para segurança dos dados;
- O tamanho máximo do banco de dados é de 2Gb;
- O tamanho máximo de cada linha é de 4k;
- O número de linhas de cada tabela depende do tamanho do banco de dados;
- Pode ter até 1.000 tabelas por banco de dados;
- Pode ter até 65.535 linhas por tabela;
- Possui suporte para sincronização em redes sem fio;
- O espaço em disco ocupado pelo no cliente e cerca de 50 kb.

O SBGD da Sybase se destaca pela interoperabilidade com inúmeros fabricantes, o que o torna muito flexível. Outro fator importante é que o cliente ocupa o menor espaço em disco dentre os SBGS estudados, cerca de 50 kb, fator importante, quando se encontram limitações na capacidade de armazenamento. Suporta o campo BLOB (*Binary Large Object*) que permite o armazenamento completo do objeto (imagem, vídeo ou som) na base de dados, tornando mais rápida sua localização e recuperação.

#### **4.3.2 Oracle Lite Mobile Server**

O Produto *Oracle Lite Mobile Server* está construído sob o Sistema Oracle9i Application Server. Possui uma estrutura XML e suporta o desenvolvimento de aplicações que podem utilizar voz, troca de mensagens e o acesso sem fio. As aplicações móveis podem acessar seus dados localmente. A versão *Oracle 9i Lite* possui dois componentes principais [38]:

1. *Oracle 9i Lite Mobile Server*, responsável pela gerência dos bancos de dados móveis e pelo sincronismo das aplicações móveis com um amplo número de equipamentos móveis; e
2. *Oracle 9i Lite Mobile Development Kit*, produto para desenvolvimento de aplicações para várias plataformas. Suas principais características são:
  - Pode ser utilizado nas plataformas Palm, CE, Symbian EPOC e Windows 95/98/NT/2000;
  - Possui suporte para sincronização em redes sem fio;
  - Suporta as funcionalidades da SQL padrão;
  - Possui serviços de mensagens em aparelhos de telefone, pagers e computadores portáteis;
  - Possuem serviços de mensagens *Push-based* e *Pull-based* para *broadcast*;
  - Possui serviços de voz;
  - Prove serviço de encriptografia para segurança dos dados;
  - Não é divulgado o espaço físico ocupado pelo cliente.

Esse SGBD tem como vantagens à utilização em várias plataformas. Porém o espaço físico ocupado pelo cliente não é divulgado e também não dá suporte ao campo *BLOB* (*Binary Large Object*), o que facilitaria muito o armazenamento de dados não convencionais em dispositivos móveis.

#### **4.3.3 DB2 Everyplace**

O *DB2 Everyplace* é um banco de dados relacional com aproximadamente 150K. Pode ser utilizado como um banco de dados local quando seu *host* está desconectado ou como um cliente acessando o servidor durante a conexão com a rede fixa. Possui sincronismo bi-direcional com os SGBDs corporativos [27]. Suas principais características são:

- Pode ser utilizado nas plataformas PalmOS, Microsoft Windows CE/Pocket PC, Symbian EPOC, embutido no Linux, QNX Neutrino e Microsoft Win32;

- Possui suporte para sincronização em redes sem fio;
- Pode ser sincronizado como cliente com o IBM DB2 Universal Server nas plataformas Unix, Windows, OS/2, OS/390 e AS/400;
- Suporta as funcionalidades da SQL padrão;
- Possui interface *QBE - Query-By-Example* como interface de consulta;
- Não suporta sub-consultas, visões, *triggers*, *stored procedure*, tipo de dados *BLOBs* funções definidas pelo usuário;
- O espaço em disco ocupado pelo no cliente e cerca de 150 kb.

O DB2 é muito semelhante ao *Oracle Lite Mobile Server*, pois também pode ser utilizado em diversas plataformas. Contudo, seu cliente ocupa um espaço físico de 150 kb, espaço muito superior ao ocupado pelo SGBD da *Sybase*. Além disso, esse SGBD também não suporta o campo *BLOB (Binary Large Object)*, não permitindo assim o armazenamento completo do objeto na base de dados.

#### 4.3.4 Microsoft SQL Server CE

O *Microsoft SQL Server CE* é um banco de dados relacional para o desenvolvimento de aplicações para os equipamentos móveis. Pode ser utilizado como um banco de dados local quando seu *host* está desconectado ou como um cliente acessando o servidor durante a conexão com a rede fixa [34]. Suas principais características são:

- Suporta a sintaxe de consulta da Linguagem SQL;
- Fornece uma *API* para acesso ao *SGBD Microsoft SQL Server*;
- Possui otimização de consultas e processamento de transações;
- Projetado para ser integrado com a Plataforma *Microsoft .NET*;
- Possui seu tamanho em torno de 1 Mb;
- Prove serviços de encriptografia para segurança dos dados;
- Mantêm interoperabilidade com os produtos da *IBM, Oracle e Sybase*;

- Dá suporte às plataformas Win32 e Windows CE, Linux, Palm OS, QNX Neutrino e Symbian EPOC;
- O espaço em disco ocupado pelo no cliente varia entre 800 Kb-3 Mb.

O Microsoft SQL CE também suporta diversas plataformas, contudo, tem desvantagens em relação ao SGBD da Sybase, pois ocupa de 800 kb a 3Mb de espaço em disco contra 50 kb da Sybase. Além disso, também não suporta *BLOB (Binary Large Object)*, o que dificultaria o armazenamento em dispositivos móveis.

Existem ainda diversos trabalhos sendo propostos na área de banco de dados móveis:

VLACH em [52] apresenta o projeto MDBAS que é um protótipo de um sistema de gerência de múltiplos bancos de dados baseado em agentes móveis, integrando um conjunto de bancos de dados autônomos distribuídos sob a rede, possibilitando a criação de esquemas globais de bancos de dados, gerenciando a transparência de execução para os clientes móveis.

PREGUICA em [42] apresenta um modelo de transação para o projeto MobiSnap, o qual é um projeto que tem por objetivo dar suporte ao desenvolvimento de aplicações baseadas em um banco de dados relacional para o ambiente móvel.

PENG end CHEN em [39] propõem um algoritmo de *data mining* para descoberta de padrões na movimentação dos clientes. O interessante desta pesquisa é que combinam as técnicas de *data mining* de reconhecimento de padrões com o objetivo de melhorar o desempenho total de um sistema móvel.

Esses demais SGBDs estão em estudo ou teste, sendo que existem poucas informações técnicas a respeito dos mesmos, dificultando, assim, maiores comparações.

A diversidade de SGBDs é grande, cada qual com seus pontos fortes e fracos, ficando a cargo do usuário a escolha do mais adequado às suas necessidades.

#### **4.4 Desenvolvimento de Aplicações para Dispositivos Móveis**

As aplicações e serviços para dispositivos móveis são normalmente desenvolvidos baseados em arquiteturas específicas, usando componentes não-reutilizáveis, ou seja uma

determinada aplicação desenvolvida para um dispositivo em uma plataforma utilizando uma dada ferramenta [29].

A portabilidade dessa aplicação para outros ambientes é feita através de um processo que, em muitas situações, é necessário refazer quase que totalmente a aplicação gerando com isso altos custos tornando-se por muitas vezes inviável.

Segundo FRIDAY [20], ferramentas capazes de facilitar o desenvolvimento de aplicações para os diversos dispositivos móveis e que possam atender às várias formas de aplicações serão extremamente úteis para integradores e desenvolvedores de novos serviços para ambientes móveis. A variedade de dispositivos móveis traz consigo uma grande variedade de plataformas e, conseqüentemente, diferentes ferramentas de desenvolvimento.

O desenvolvimento de aplicações para celulares e PDAs tem características diferentes. No caso de celulares têm-se utilizado linguagens de marcação, como WML, HDML, CHTML, e VoiceXML [17]. Em geral, PDAs têm mais recursos computacionais quando comparados aos celulares, o que faz com que suas aplicações e serviços tenham mais funcionalidades. Existem linguagens de marcação disponíveis que têm sido muito utilizadas sendo normalmente subconjuntos de HTML. Seu maior concorrente é a linguagem XML (*eXtensible Markup Language*). As diferenças entre as duas podem ser vista na Tabela 4.1.

TABELA 4.1 - DIFERENÇAS ENTRE HTML E XML

<b>Características</b>	<b>HTML</b>	<b>XML</b>
Extensibilidade	Conjunto fixo de etiquetas	Conjunto de etiquetas extensível
Propósito de Etiquetas	Etiquetas descrevem apresentação	Etiquetas descrevem conteúdo
Visão	Apresentação única	Múltiplas visões do mesmo documento (por XLS)
Orientação	Documentos	Documentos + dados semi-

		estruturados
Pesquisa	Somente por palavra chave	Por palavras chaves e também consultas por campos sensíveis a utilização de letras maiúsculas e minúsculas nas expressões

Quanto à característica “extensibilidade”, a linguagem HTML [26] oferece um conjunto fixo de etiquetas tais como "<B>", "<I>", "<P>". Ao contrário, um conjunto de etiquetas extensível, refere-se ao fato de o desenvolvedor poder criar suas próprias marcações. No caso da marcação de um livro, pode-se gerar etiquetas padronizadas, como por exemplo "<TÍTULO>", "<AUTOR>" e "<SEÇÃO>", entre outras.

O padrão HTML [26] descreve características de exibição de um determinado documento, tais como tamanho da margem e fonte, ao passo que a marcação XML descreve o conteúdo dos documentos, ou seja, identifica o conteúdo de forma semântica.

Pelo fato de a linguagem HTML possuir marcação de exibição misturada com conteúdo, o mesmo documento pode ter somente uma forma de apresentação. Caso seja necessário exibir este documento de outra maneira, ele terá que ser etiquetado novamente [26]. No caso do padrão XML [5], apenas um documento é marcado, e as folhas de estilo XSL possibilitam múltiplas visões deste documento. Esta característica ainda implica a facilidade de manutenção.

Segundo PITTS-MOULTIS [40], XML é orientado a documentos mais dados semi-estruturados, pois em um processo de consulta os dados estão identificados semanticamente. Além disso, com o padrão XML, é possível fazer distinção entre caracteres maiúsculos e minúsculos nos termos digitados para consulta.

Porém, antes de escolher uma ferramenta para desenvolver aplicativos para dispositivos móveis é necessário ver se ela atende a alguns requisitos importantes [40].  
como:

- Capacidade de descrever aplicações independentes do dispositivo móvel que será utilizado. Isto permite que a mesma descrição possa ser utilizada em qualquer dispositivo.
- Inserção de código. A descrição deve permitir a inclusão de código externo para podendo ser inserido pelo projetista ou usuário da ferramenta. Seria uma forma de atualização ou um plugin para realizar alguma tarefa extra.
- Sincronização de aplicações. A ferramenta deve prover facilidades para a criação de mecanismos de sincronização,
- Importação de dados. Deve ser possível adquirir dados e conteúdo de fontes diversas como XML, bancos de dados e protocolos de aplicação da Internet.
- Extensibilidade. Deve ser extensível, possibilitando que novas interfaces, novas aplicações e novos dispositivos sejam incorporados facilmente.
- Curva de aprendizado aceitável. Deve ser de fácil utilização. Este é um requisito ligado a usabilidade. A interface da ferramenta deve ser intuitiva o suficiente para que o projetista possa usa-la facilmente.

Esses requisitos apresentados visam aumentar a adaptabilidade da aplicação independentemente do dispositivo utilizado.

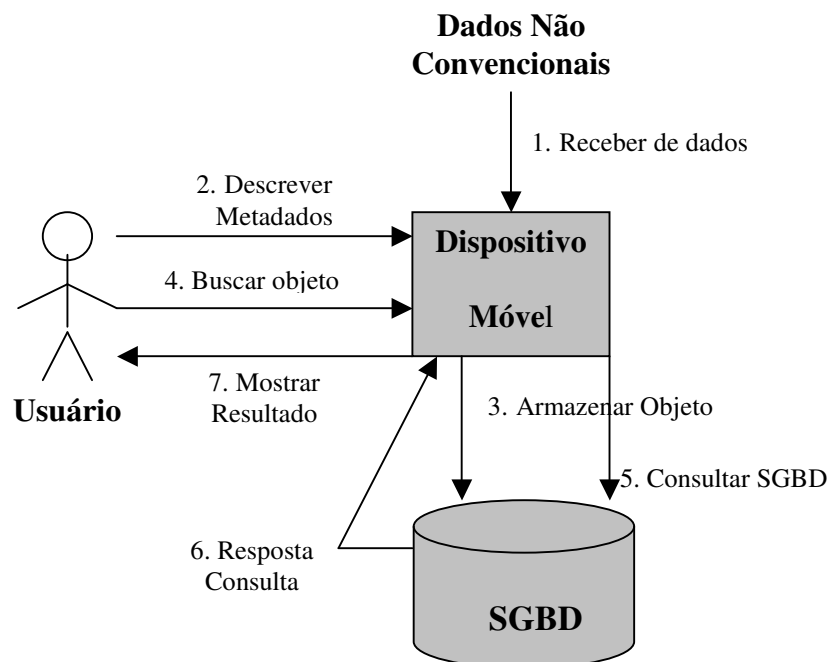


## 5. ESTRUTURA DE GESTÃO DE DADOS PARA DISPOSITIVOS MÓVEIS ATRAVÉS DE METADADOS

Os dispositivos móveis, com o passar dos anos, vêm dando suporte a formatos e tamanhos de dados cada vez maiores, tais como, imagens, sons e vídeos. Contudo, a gestão desses dados não acompanha a mesma evolução, e com a disponibilização de novas tecnologias é necessário que seus usuários precisem que esses dados sejam gerenciados de forma rápida e precisa. Uma alternativa de gestão desses dados em dispositivos móveis, resultado desta dissertação, corresponde a uma forma de gerencia-los utilizando metadados, com o objetivo de facilitar sua indexação, busca e recuperação. Para apresentar este modelo de gestão, este capítulo está estruturado como segue. Na Seção 5.1 é apresentada a estrutura de gestão proposta. Na Seção 5.2 mostra a escolha de um padrão de metadados a serem inseridos e a Seção 5.3 mostra uma especificação de gestão.

### 5.1 Estrutura de Gestão de Dados

Essa forma de gestão proposta tem por objetivo permitir a descrição dos metadados pelo usuário, facilitando com isso a busca e recuperação do mesmo como mostra a Figura 5.1.



### FIGURA 5.1 - ESTRUTURA DE GESTÃO PROPOSTA

A estrutura de gestão proposta para dispositivos móveis é composta de sete etapas, as quais são descritas a seguir, pois, elas serão aprofundadas mais detalhadamente nas Seções 5.2 e 5.3.

#### 1. Receber de dados

Esta fase consiste no recebimento dos dados pelo dispositivo, ou seja, maneira pela qual os dados serão incluídos no dispositivo móvel. Esses dados podem ser enviados por *e-mail* ou adquiridos através de um *download*, ou podem ser gerados pelo próprio usuário (tirar uma foto, gravar um vídeo, gravar um som);

#### 2. Descrever metadados: consiste no preenchimento dos valores dos metadados pelo usuário;

3. Armazenar objeto: após os valores dos metadados serem atribuídos pelo usuário, serão indexados na base de dados. Essa indexação dos objetos é feita com o objetivo de recuperá-los posteriormente.

4. Buscar objeto: essa etapa consiste na busca de determinado (s) objeto (s) pelo usuário. Essa busca é feita por meio dos valores dos metadados correspondentes ao objeto armazenado na base de dados;

5. Consultar SGBD: através da escolha do (s) metadado (s) especificado (s) pelo usuário na etapa anterior é feita uma busca na base de dados. Essa busca procura o valor, ou valores existentes do metadado solicitado pelo usuário, na base de dados. Após a identificação do valor e/ou valores do metadado, são encontrados os objetos representados pelos mesmos.

6. Resposta consulta: após realizar essa busca na base de dados terá como retorno o resultado da mesma, podendo encontrar um ou mais objetos que atendam a característica do metadado especificado, ou mesmo, não encontrar objeto algum;

7. **Mostrar Resultados:** Consiste em mostrar ao usuário o objeto recuperado por meio dos resultados obtidos pela busca;

Essa estrutura de gestão tem por objetivo priorizar o ganho em termos de tempo, sendo que não será necessário ao usuário acessar inúmeros objetos até encontrar o objeto desejado, pois somente os objetos que atenderem o valor do metadado fornecido pelo usuário serão recuperados.

## 5.2 Escolha dos Metadados a serem inseridos

Por meio do estudo dos padrões de metadados foi possível estudar algumas características de descrição e armazenamento utilizadas pelos mesmos a fim de fornecer as características (metadados) que serão utilizados no modelo. O padrão sugerido é o Dublin Core pela simplicidade, por ser um consenso internacional, visto que a Internet é um recurso global. Inicialmente é apresentado um conjunto previamente definido de metadados disponíveis ao usuário, os quais são um sub-conjunto do padrão Dublin Core. Esse conjunto de metadados visa atender as características básicas dos objetos armazenados pelos usuários de dispositivos móveis. Inicialmente a estrutura possui cinco metadados (características) e pode ser observada na Tabela 5.1.

TABELA 5.1 - CONJUNTO DE METADADOS FIXOS UTILIZADOS NO MODELO PROPOSTO

<b>Metadados</b>	<b>Significado</b>
Título	Nome para o objeto
Autor	Dono do objeto
Assunto	Descreve uma ocasião especial, a qual o objeto possa representar para o usuário;
Data	Data na qual foram inseridos os metadados
Formato	Imagem: bmp; Som: mp3; Vídeo: avi
Tipo	Se o objeto e uma imagem, som ou vídeo

**Título:** nome dado à imagem, vídeo ou áudio recebido;

**Autor:** dono ou responsável pelo objeto;

**Assunto:** descreve uma ocasião, a qual o objeto esteja ligado;

**Data:** data na qual foi recebido o objeto, no formato (DD-MM-AAA);

**Formato:** formato do tipo de objeto, como bmp para imagem, avi para vídeo e mp3 para áudio;

**Tipo:** tipo do objeto, ou seja, se ele é uma imagem, vídeo ou áudio;

Poderiam ter sido utilizados outros tantos metadados, porém como esses dispositivos tem capacidade de armazenamento limitada e com o intuito de não sobrecarregar o processamento julgou-se mais adequado um conjunto predefinido de metadados/valores e permitir ao usuário inserir novos metadados, caso, ache necessário para posteriormente recupera-los. Essa flexibilidade permite ao usuário criar metadados (informações) considerados importantes por ele, respeitando assim, a individualidade de cada um.

### 5.3 Especificações da estrutura proposta

Inicialmente o usuário vai poder optar pela escolha entre três operações: a indexação, exclusão e busca/recuperação de objetos através dos metadados descritos por ele. A Figura 5.2 mostra as operações que estarão disponíveis ao usuário.



FIGURA 5.2 - OPERAÇÕES DISPONÍVEIS

O usuário optará por uma das operações mostradas anteriormente. A operação de busca e responsável também pela recuperação do objeto. Esse gerenciamento se faz necessário à medida

que ocorre o crescimento da quantidade de dados que podem ser manipulados por esses dispositivos propiciando ao usuário encontrar o objeto que procura de maneira automática tornando-se mais eficiente e rápida do que uma busca manual, visto que não haverá necessidade de ficar procurando em todos as pastas e diretórios existentes.

### 5.3.1 Indexação do Objeto

Após o dispositivo móvel receber o objeto, o usuário vai visualizá-lo, no caso de imagem ou vídeo, e ouvi-lo no caso de áudio nos aplicativos padrões de cada dispositivo tais como visualizador de imagem e vídeo e reproduzidor de sons. Feito isso ele irá escolher o tipo de objeto que ele deseja inserir na base de dados, como mostra a Figura 5.3.

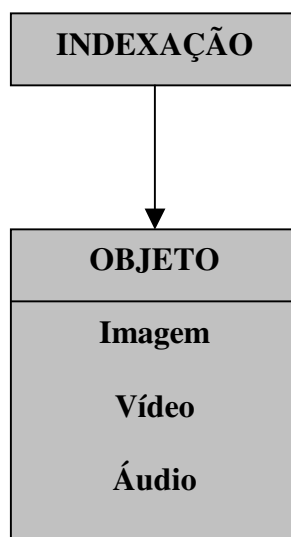


FIGURA 5.3 – DEFINIÇÃO DO OBJETO A SER INSERIDO

Após definir o objeto a ser inserido, o usuário irá atribuir valores ao conjunto de metadados fixos disponíveis no modelo, e se julgar necessário, poderá acrescentar novos metadados/valores para representar o objeto. Após o usuário determinar os valores dos metadados o objeto e seus metadados serão armazenados na base de dados.

Para armazenar esses objetos pode ser utilizado o SGBD *Sybase UltraLite*, pois como já mencionado anteriormente os dispositivos móveis apresentam limitações no que diz respeito principalmente à memória, sendo que atualmente essa capacidade varia entre 60 mb até 220 mb.

Além disso, esse SGBD permite a comunicação com produtos de outros fabricantes, o que o torna o mais versátil dentre os SGBDs estudados. Outro fator importante é o fato de suportar *Binary Large Object (BLOB)*, campo utilizado para armazenar objetos do tipo imagens, sons e vídeos;

O armazenamento consiste inicialmente em utilizar duas tabelas, uma contendo os metadados/valores e outra os objetos propriamente ditos. Essa estrutura pode ser vista nas Tabelas 5.2 e 5.3 respectivamente.

TABELA 5.2 – METADADOS/VALORES

<b>TABELA METADADOS</b>							
Cód	Metadados						Cód Objeto
	Título	Autor	Assunto	Data	Formato	Tipo	
1	Colação	Juliano	Formatura	02-12-2002	bmp	imagem	1
2	Vamos Fugir	Skank	Skank ao vivo	27-11-2000	mp3	som	2
3	Parainfo	Cavalin	Formatura	04-12-2002	avi	vídeo	3

TABELA 5.3 - OBJETO

<b>TABELA OBJETO</b>	
Cód objeto	Objeto (Tipo blob)
1	000000000000111111010000100
2	01010000000011101000010010011
3	0011000000001111110100001111110

A Tabela Metadados contém um valor (cod) para gerenciamento interno e posteriormente um conjunto de metadados com seus respectivos valores. É importante ressaltar que esse conjunto de metadados pode ser expandido pelo próprio usuário sempre que julgar necessário. Existe também um cód objeto, o qual, contém o mesmo valor do cód objeto da Tabela Objeto.

A tabela objeto possui um cód objeto que é enviado para a Tabela Metadados, ou seja, é responsável pela ligação do objeto com seus respectivos metadados/valores. O outro campo existente na Tabela Objeto é o campo Objeto, campo esse do tipo *Binary Large Object (BLOB)*, o qual é utilizado para armazenar objetos do tipo imagens, sons e vídeos na forma binária.

O armazenamento do objeto em uma tabela separada de seus metadados se justifica principalmente em termos de desempenho, visto que quando for feita a busca por um determinado metadado não vai precisar carregar a todo o momento o objeto, só irá buscar o objeto que atender a solicitação do metadado/valor realizado pelo usuário. Em se tratando de dispositivos móveis a questão desempenho de processamento é muito importante visto que esses dispositivos possuem desempenho limitado quando comparados à micro computadores bem como *notebooks*.

### **5.3.2 Busca e Recuperação do objeto**

Essa etapa consiste em disponibilizar ao usuário a busca de determinado objeto contido na base de dados. A busca poder ser feita por meio de qualquer um dos metadados pré-definidos ou ainda por algum metadado inserido pelo próprio usuário, sendo que a busca por título retornará um único valor (objeto), pois não pode ser duplicado. As demais opções (data, autor, assunto, formato e tipo) podem trazer tantos quantos resultados forem encontrados na base de dados durante a consulta. A estrutura de busca proposta para dispositivos móveis é mostrada na Figura 5.4.

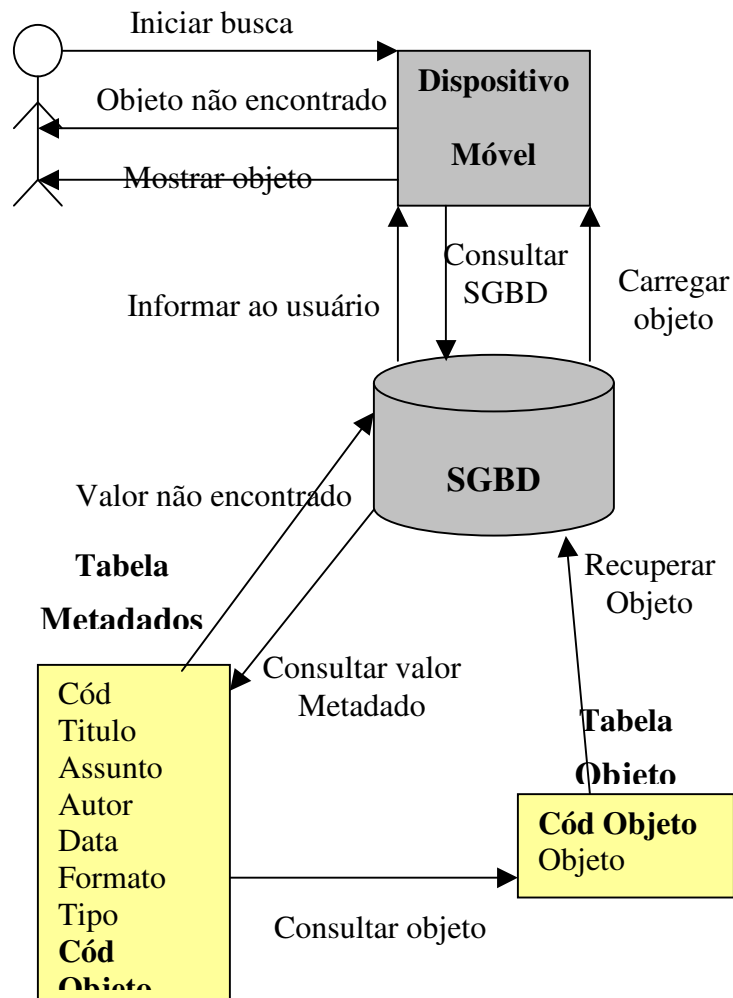


FIGURA 5.4 – ESTRUTURA DE BUSCA E RECUPERAÇÃO.

Antes de iniciar a busca é preciso que o usuário determine o valor do metadado que será responsável pela busca e recuperação do objeto.

Após ser determinado um valor para o metadado, será feita uma consulta ao SGBD. Primeiramente é percorrida a tabela metadado com o objetivo de encontrar ocorrências do padrão (valor do metadado) informado pelo usuário. Caso não seja encontrado um valor igual ao determinado inicialmente, o usuário receberá uma mensagem de objeto não encontrado. Se o valor determinado inicialmente é encontrado



na tabela metadado, então será realizada uma consulta na tabela objeto para recuperação do mesmo e conseqüentemente mostrado para o usuário.

O algoritmo de pesquisa sugerido é o KMP porque irá casar caractere a caractere do valor do metadado informado pelo usuário com os valores disponíveis na tabela metadado, garantido que somente aqueles em que ocorrerem o casamento exato entre os valores serão recuperados. Esse processo minimiza o desempenho do processador bem como a memória utilizada pelo dispositivo, uma vez que a busca somente se dará na base de metadados, ou seja, não terá que ficar carregando o objeto na memória. Isso somente acontecerá quando o valor do metadado determinado pelo usuário for igual ao valor do metadado disponível na tabela metadado, ai então a tabela objeto será acessada para recuperar e mostrar o objeto para o usuário.

#### **5.4. Trabalhos Relacionados**

O armazenamento de dados não convencionais (imagens, sons e vídeos) passou a ser uma necessidade dos usuários. Esse armazenamento tem por principal objetivo propiciar ao usuário a recuperação desses dados de forma fácil e rápida. A eficiência dessa recuperação pode depender dos tipos de dados que são tratados, dos métodos de recuperação utilizados e dos algoritmos de pesquisa implementados.

Nesta seção são apresentados dois trabalhos relevantes, dando ênfase aos tipos de objetos armazenados, utilização dos metadados na indexação, pesquisa e recuperação dos mesmos.

##### **5.4.1 Sistema VIMSYS**

O Sistema VIMSYS (*Visual Information Management System*) [2], foi projetado com o objetivo de gerenciar informações visuais para reconhecimento de faces humanas. Esse armazenamento se dá por meio de um conjunto de atributos básicos, onde a partir desses o usuário interage com o sistema, introduzindo novos atributos. Suporta três tipos principais de recuperação sendo elas:

1. Recuperação através de um identificador: utiliza-se consultas a qualquer um dos atributos das imagens na base de dados;
2. Recuperação por sentenças condicionais: permite que a pesquisa seja realizada em estágios, ou seja, permitindo que o usuário ajuste os critérios das consultas a cada estágio e progressivamente consiga os resultados desejados;
3. Recuperação por similaridade: é feita através da descrição de uma imagem. O sistema recupera as imagens que possuíam maior grau de similaridade à descrição feita pelo usuário.

O sistema VIMSYS utiliza um algoritmo de segmentação para localizar cada objeto em uma imagem. Para cada objeto do domínio, os atributos e funções necessárias serão mantidos para permitir que esse objeto seja identificado e segmentado em uma imagem.

#### **5.4.2 Sistema CIARS**

O sistema CIARS (*Color Image Archival and Retrieval System*) [1] foi projetado com o objetivo de permitir a indexar e recuperar imagens. Esse sistema permite indexação por cor, sendo que os atributos da mesma são extraídos automaticamente através de um algoritmo de *clustering* de cores.

A indexação de atributos é feita de forma automática. Esses atributos correspondem a informações visuais relacionados a cores.

Esse sistema suporta a recuperação baseada em similaridade, sendo que utiliza o conceito de espaço de similaridade, ou seja, os objetos contidos em um mesmo grupo são mais similares comparados com os objetos pertencentes a grupos diferentes.

Um algoritmo de *cluster* é utilizado para busca. Este algoritmo é tratado da mesma forma que nos sistemas tradicionais, isto é, agrupa objetos similares formando clusters.

#### **5.4.3 Análise Comparativa**

Nesta seção é realizada uma comparação entre os sistemas descritos, juntamente com o modelo proposto nesse trabalho, como mostra a Tabela 5.4.

TABELA 5.4 - ANÁLISE COMPARATIVA SISTEMAS

<b>SISTEMAS</b>			
<b>CARACTERÍSTICAS</b>	<b>VIMSYS</b>	<b>CIARS</b>	<b>ESTRUTURA PROPOSTA</b>
Mídia Suportada	Vídeo	Imagem	Vídeo, Som e Imagem.
Forma de Indexação	Manual Automática	Automática	Manual
Forma de Recuperação	Identificador Condicional Similaridade	Similaridade	Identificador
Algoritmo de Busca	Segmentação	Clustering	KMP
Utilização de SGBD Móvel	Não	Não	Sybase Ultralite Oracle Lite DB2 Everyplace SQLServerCE
Padrão de metadados adotados	Não	Não	Dublin Core

As colunas da Tabela 6.1 correspondem aos trabalhos analisados, enquanto que as linhas correspondem às características relevantes ao trabalho proposto. Os modelos analisados tratam de mídias diferentes, sendo que cada um deles possui a sua forma de indexação. O modelo VIMSYS, por exemplo, gerencia informações visuais para reconhecimento de faces humanas, já o modelo CIARS trata da gerencia de imagens. A estrutura de Gestão proposta para dispositivos móveis é permitir ao usuário a indexação de dados não convencionais, ou seja, vídeo, áudio e imagens, não fazendo distinção entre os tipos de mídias.

Quanto à forma de indexação, os modelos, descrevem seus objetos de maneira particular. A forma de indexação mais comum é a indexação manual, ou seja, o usuário é responsável por construir um conjunto mínimo de metadados/valores capazes de descreverem os objetos (imagens, vídeos e áudio). Contudo, existem sistemas que possibilitam outras formas de indexação, como a automática ou a semi-automática. É o caso do modelo VIMSYS que possibilita a indexação de forma manual e automática. Isso se torna possível por meio de um conjunto de atributos básicos, onde a partir desses o usuário interage com o sistema, introduzindo novos atributos. A indexação no modelo CIARS é feita de forma automática. Essa indexação é realizada em um espaço de similaridade, ou seja, os atributos correspondem a atributos de informações visuais relacionados a cores. A gestão proposta permite a indexação manual, ou seja, uma vez definido um conjunto de metadados/valores para descrever um determinado objeto, os mesmos são utilizados. Porém é permitido ao usuário inserir tantos quantos metadados/valores ele julgue necessário para representar o objeto.

O processo de recuperação de informação pode ser através de métodos como: identificador, condicional e por similaridade. O modelo CIARS apresenta a recuperação por meio de similaridade utilizando o conceito de espaço de similaridade, ou seja, os objetos que fazem parte de um mesmo grupo são mais similares do que objetos pertencentes a outros grupos. A gestão proposta permite a recuperação por meio de um identificador, ou seja, é atribuído valor para um determinado metadado e todos os objetos que apresentem o valor do metadado igual ao especificado pelo usuário serão recuperados. O modelo VIMSYS é o único que apresenta os três métodos de recuperação. A recuperação por meio de um identificador permite que o objeto seja recuperado a partir de qualquer um dos metadados/valores disponíveis na base de dados. A recuperação por sentenças condicionais permite a realização da pesquisa em estágios, sendo que a cada estágio o usuário pode ajustar critérios para atingir o resultado desejado. Na recuperação por similaridade as imagens recuperadas serão aquelas que possuem maior grau de similaridade à descrição do usuário.

No que diz respeito aos algoritmos de busca, cada modelo utiliza o que melhor atende as suas necessidades. O modelo VIMSYS utiliza um algoritmo de segmentação. Os atributos e funções necessárias serão mantidos para cada objeto do domínio permitindo a identificação e segmentação do mesmo em uma imagem. Já o modelo CIARS utiliza um algoritmo de *cluster* onde os objetos similares são agrupados formando clusters facilitando com isso a busca. A gestão proposta utiliza o algoritmo KMP (*Knuth-Morris-Prat*), o qual, compara um padrão com um texto, realizando uma varredura completa de texto, visto que o usuário entrará com um padrão (valor) de um determinado metadado e o sistema irá percorrer a tabela metadado até encontrar o valor correspondente ao informado pelo usuário.

Existem, atualmente disponíveis no mercado, quatro SGBDs para dispositivos móveis, os quais podemos citar: o *Sybase Ultralite* desenvolvido pela *Sybase*, o *Oracle Lite* desenvolvido pela *Oracle*, o *DB2 Everyplace* desenvolvido pela *IBM* e o *SQLServerCE* desenvolvido pela *Microsoft*. O *Sybase UltraLite* e o *SQLServerCe* suportam estruturas de dados do tipo BLOB (*Binary Large Object*), o que auxilia e muito o armazenamento de dados não convencionais, um dos objetivos desse trabalho. Todos os quatro SGBDs citados apresentam uma interoperabilidade muito grande no que diz respeito as plataformas a serem utilizadas tais como: *PalmOS*, *Microsoft Windows CE/Pocket PC*, *Microsoft Win32* e *Linux*. Outro aspecto importante diz respeito ao espaço necessário para armazenar o cliente no dispositivo móvel. Nesse aspecto mais uma vez destaca-se o *Sybase Ultralite* por ocupar cerca de 50 kb, contra 150 kb do *DB2* e 800 kb a 3 Mb do *SQLServer Ce*. Tanto o modelo VIMSYS quanto o modelo CIARS foram projetados para serem utilizados em computadores fixos, os quais utilizam bases de dados com grande quantidade de informações, além de requererem capacidades de armazenamento e processamento muito elevadas, tornando sua utilização inviável para dispositivos móveis, uma vez que esses, dispõem atualmente, de grandes limitações nesses quesitos, os quais são muito importantes para o desenvolvimento de aplicações que utilizem SGBDs móveis. O modelo proposto tem como finalidade suportar qualquer SGBD citado anteriormente, porém, sugere-se a utilização do *Sybase Ultralite*, pois, destaca-se dos demais por ocupar pouco espaço na instalação do cliente, além disso,

permite a comunicação com produtos de outros fabricantes, o que o torna o mais versátil dentre os SGBDs estudados. Outro fator importante é com relação à forma de armazenamento de dados não convencionais, sendo que esse SGBD por suportar *Binary Large Object (BLOB)*, facilita o armazenamento desse tipo de dados.

Conforme mostrado no Capítulo 3 existem inúmeros padrões de metadados, sendo que cada um deles possui uma estrutura própria e pode ser aplicado conforme o tipo de dado a ser descrito. O padrão Dublin Core é marcado pela sua simplicidade e interoperabilidade, contém poucos atributos, fator que auxilia na descrição e armazenamento de dados. Os padrões Marc e Gils são utilizados em grandes quantidades de objetos, pois possuem um número muito maior de atributos para descrever os mesmos. Os sistemas VIMSYS e CIARS não seguem nenhum padrão para descrever os metadados. No sistema VIMSYS o usuário é responsável pela escolha dos metadados, ou seja, ele vai inserindo metadados até formar o seu conjunto de metadados, sendo que para cada metadado criado é atribuído um valor. O sistema CIARS possui um conjunto de metadados, os quais dizem respeito às cores do objeto. Os valores para esse conjunto de metadados são obtidos de forma automática através da utilização de um algoritmo de *clusterin*, ou seja, o usuário pode recuperar todas as imagens que tem uma cor particular, como por exemplo, a cor azul. Todas as imagens que tiverem a cor azul serão recuperadas. A Gestão proposta utiliza o padrão Dublin Core para descrição de seus metadados. A escolha pelo DC se deu principalmente pela sua simplicidade e fácil utilização, além do conjunto pequeno, porém significativo de atributos. Essa estrutura disponibiliza um conjunto predefinido de metadados, possibilitando também a inserção de novos metadados pelo usuário, sempre que esse julgar necessário.

## 6. CONCLUSÕES E TRABALHOS FUTUROS

A chegada de novas infra-estruturas de comunicação sem fio abrirá novas oportunidades para serviços de comunicação, entretenimento e gerenciamento, além de ampliar a disponibilização de serviços de conteúdo. O telefone celular e o PDA deverão continuar a evoluir para um terminal móvel com funcionalidades semelhantes entre os dois dispositivos. A quantidade crescente da utilização de dados não convencionais pelos dispositivos móveis acabou gerando desafios novos no que diz respeito à usabilidade

Os problemas ainda encontrados nos dispositivos móveis dizem respeito à capacidade de processamento e armazenamento limitadas, bem como ao tempo de utilização devido à necessidade de recarregar a bateria, os quais devem ser resolvidos brevemente devido ao grande investimento e pesquisa nessa área.

A utilização de metadados já está se tornando um artifício no que diz respeito à descrição e ao armazenamento de objetos digitais facilitando sua busca e recuperação pelos mecanismos de busca disponíveis, o que se aplica também para dados localizados em dispositivos móveis, claro que levando em consideração as restrições apresentadas pelos mesmos.

A discussão sobre utilização de metadados para gerencia de dados não convencionais em dispositivos móveis ainda está no começo, o que se justifica principalmente pela capacidade de armazenamento limitada desses aparelhos, mas tende a evoluir e juntamente com a evolução do mercado.

Como contribuições do trabalho destacamos o estudo sobre metadados e seus padrões, conceitos importantes sobre dispositivos móveis, bem como as tecnologias disponíveis para os mesmos como, SGBDs móveis e linguagens de desenvolvimento e finalmente, uma estrutura de gestão de dados não convencionais baseada em metadados para ambientes de dispositivos com o objetivo de propiciar o usuário um maior controle sobre seus dados..

Como perspectivas de trabalhos futuros a implementação dessa estrutura de gestão proposta é considerada a principal delas, pois, dar-se-ão maiores subsídios para estudos mais completos no futuro. Além da implementação também podemos citar:

- Um estudo sobre a viabilidade da utilização de um SGBD gratuito em dispositivos móveis como, por exemplo, o PostgreSQL;
- Alternativas para gestão de dados em dispositivos móveis;
- Estudos mais amplos sobre as diferentes linguagens de programação utilizada para o desenvolvimento de aplicativos, tendo em vista um maior aproveitamento dos recursos disponíveis por elas;
- Utilização da recuperação através de sentenças condicionais e similaridade;

Um outro aspecto que podemos destacar é no que diz respeito à interface para dispositivos móveis. Pode ser realizado um estudo específico na área de IHC (Interação Homem x Computador) com o intuito de investigar como elaborar um projeto de interface para dispositivos móveis, quais aspectos a considerar, por exemplo, como deve ser feita a entrada de dados, considerações sobre as limitações de tela visível ao usuário propiciando uma maior integração do mesmo com o software para que ambos atinjam seus objetivos.



## REFERÊNCIAS BIBLIOGRÁFICAS

- [1] BABU, G. P; MEHTRE, B. M; KANKANHALLI, M. S. **Color indexing for efficient image retrieval.** *Multimedia Tools and Applications*, p. 327-348, November. 1995.
- [2] BACH, J. R; PAUL, S; JAIN, R. **A visual information management system for the interactive retrieval of faces.** *IEEE Transactions on Knowledge and data Engineering*. P. 619-628. august, 1993.
- [3] BAEZA-YATES, R. and RIBEIRO-NETO, B. **Modern Information Retrieval**, chapter Text and Multimedia Languages and Properties, pages 142,143. ACM press, Addison Wesley, 1999.
- [4] BALDONADO, M. et al. **Metadata for digital libraries: architecture and design rationale.** In Proc. of the 2nd. ACM International Conference on Digital Libraries, pages 47-56, Philadelphia, PA, USA, 1997.
- [5] BRAY, T. et al. **Extensible Markup Language (XML) 1.0 (Second Edition)** W3C Recommendation (outubro 2000). Webreferência Disponível em: <http://www.w3c.org/TR/REC-xml>. Acesso em 10 jun 2005.
- [6] BURNARD, L. and LIGHT, R. **Three SGML metadata formats: TEI, EAD, and CIMI** <[hosted.ukoln.ac.uk/biblink/wp1/sgml](http://hosted.ukoln.ac.uk/biblink/wp1/sgml)> Dez 1996.
- [7] BURNARD, L. **Text Encoding for Information Interchange.** Disponível em <<http://www-tei.uic.edu/orgs/tei/info/teij31/index.html>> Jul 1995. Acesso em: 8 mar 2005.
- [8] CHRISTIAN, E. J. **GILS. What is it? Where's it going.** D-Lib Magazine, dez. 1996. Disponível em <http://www.dlib.org/dlib/> Acesso em: 10 abr 2005.
- [9] COLE, T. W; PAT, A; SCHMITZ, J. **Building an outreach digital library**

- collection.** In Illinois ACRL Spring 2000 Conference, 13 de abril de 2000.
- [10] COSTA, G. **Documento eletrônico e metadados.** Disponível em <http://domain.com.br/clientes/genelice/doceletronico.htm> Acesso em: 18 jan. 2005.
- [11] CSDGM (Content Standard for Digital Geospatial Metadata). Disponível em <http://www.fgdc.gov/metadata/constan.html>. Acesso em: 18 abr 2005.
- [12] DATE, C. J. **Introdução a sistemas de bancos de dados.** Rio de Janeiro: Campus, 1991. 674 p.
- [13] DEPINÉ, F. M. **Protótipo de software para dispositivos móveis utilizando JavaME para cálculo de regularidade em rally.** 2002. 55 f. Trabalho de Conclusão de Curso (Bacharelado em Ciências da Computação) – Centro de Ciências Exatas e Naturais, Universidade Regional de Blumenau, Blumenau.
- [14] DESIO, V; PENNY, P. **Informational Metadata.** Disponível na Internet <http://www.datawarehouse.com/resources/articles/desio.htm>. Sep. 1996.
- [15] DUBLIN CORE Metadata Element Set, V.1.1:Reference Description. Disponível em <http://www.dublincore.org/documents/dces>. Acesso em 12 abr. 2005.
- [16] DUBLIN CORE Metadata Initiative. Disponível em <http://dublincore.org/documents/1998/09/dces/> Acesso em 12 abr. 2005.
- [17] ENDLER, M. **Requisitos e arquiteturas de software para computação móvel.** In I Workshop de Sistemas de Informações Distribuídas de Agentes Móveis, São Paulo - Brasil. USP, 2000.
- [18] ERICSSON, S. (Sony Ericsson) Disponível em <http://www.sonyericsson.com..> Acesso em: 15 jan 2005.
- [19] ESTADOS UNIDOS, Government Printing Office. Superintendent of

- Documents. **What is Gils.** Disponível em [http://www.acess.gpo.gov/su\\_docs/gils/whatgils.html](http://www.acess.gpo.gov/su_docs/gils/whatgils.html). Acesso em: 10 abr 2005.
- [20] FRIDAY, A., DAVIES, N., BLAIR, G., CHEVERST, K. **Developing Adaptive Applications: The MOST Experience.** *Journal of Integrated Computer-Aided Engineering*, 1999, 6(2):143–157
- [21] GARCIA, S. S. **Metadados para documentação e recuperação de imagens.** Master's thesis, Instituto Militar de Engenharia, Rio de Janeiro – BR, 1999.
- [22] GILLILAND-SWETLAND, J. Setting the stage: **Introduction to metadata:** Pathways to digital information. In Baca M. (Ed.), Getty Information Institute, 2002.
- [23] GLOBAL information locator service (GILS). **Making it easier to find all de information.** Disponível em <http://www.gils.net>. Acesso em: 10 abr 2005
- [24] GRADIENTE. Disponível em <http://www.gradiente.com>. Acesso em 15 jan 2005.
- [25] HEARST, M. A. **The use of categories and clusters for organizing retrieval results.** In Natural language information retrieval: Text, Speech and Language Technology. Eds. Tomek Strzalkowski. Dordrecht. Kluwer Academic, 1999. 384 p.
- [26] HTML, HyperText Markup Language Home Page. Webreferência Disponível em <http://www.w3.org/MarkUp>. Acesso em 10 jun 2005.
- [27] IBM. **Db2 everyplace.** Disponível em <http://www3.ibm.com/software/data/db2/everyplace/>>, IBM, 2003. Acesso em 12 jan 2005.
- [28] INFOEXAME. **Crescimento no número de dispositivos móveis.** Revista

- Infoexame. Disponível em <http://www.infoexame.com.br>. Acesso em: 05 mar 2004.
- [29] LEE, W-H. and Buszko, D. **Application Adaptations to the Mobile Environment**. University of Florida Department of Computer and Information Science and Engineering, <http://members.tripod.com/danlee11/mas.htm>.
- [30] MAGAZINE, P. On the road to mobile databases. Pc Magazine. Disponível em <http://www.pcmag.com/article2/0,4149,4180,00.asp>, 2002. Acesso: 18 ago 2004.
- [31] MARC Standards. Disponível em <http://www.loc.gov/marc/>. Acesso em: 18 abr 2005
- [32] MCCRAY, A.; GALLAGHER, M.; FLANNICK, M. Extending the role of metadata in a digital library system. In *Proceedings of the IEEE Forum on Research and Technology Advances in Digital Libraries* (Baltimore, May 19-21). IEEE Computer Society, Los Alamitos, CA, 1999, 190--199.
- [33] MELO, J. C. B. **Diferentes abordagens em comparação de seqüências**. Dissertação de Mestrado, Universidade Federal de Pernambuco, 1996.
- [34] Microsoft. **Sql Server CE everyplace**. Disponível em <http://www.microsoft.com/sql/ce/default.asp/>, Microsoft, 2003 Acesso em : 12 jan 2005.
- [35] MORIATY, T. **What is Metadata? Database Programming and Design**, San Mateo, v. 10, n. 7, p. 57-59, July 1997
- [36] MOTOROLA. Disponível em <http://www.motorola.com.br>- Acesso em: 15 jan 2005.
- [37] NOKIA. Disponível em <http://www.nokia.com.br>. Acesso em: 15 jan 2005.
- [38] ORACLE. **Oracle9i application server in mobile**. Disponível em

<<http://www.oracle.com/ip/deploy/ias/mobile/index.html/>>, Oracle, 2005

- [39] PENG, C. and CHEN, S. **Developing data allocation schemes by incremental mining of user moving patterns in a mobile computing system.** *IEEE Transactions on Knowledge and Data Engineering*, 2003, 15(1):70–85.
- [40] PITTS-MOULTIS, N.; KIRK, C. **XML Black Book - Solução e Poder.** Makron Books, 2000. 627 p.
- [41] PLOTKIN, W. **Text Encoding Initiative.** Disponível em <<http://www-tei.uic.edu/orgs/tei/index.html>> Jun 1999. Acesso em: 8 mar 2005.
- [42] PREGUICA, N. M. et al. **Mobile transaction management in mobisnap.** In *ADBIS-DASFAA*, 2000, pages 379–386.
- [43] SANSUNG. Disponível em <http://www.samsung.com>. Acesso em: 15 jan 2005.
- [44] SCHAEFER, C. **Protótipo de aplicativo para transmissão de dados a partir de dispositivos móveis aplicados a uma empresa de transporte.** 2004. 53f. Trabalho de Conclusão de Curso (Bacharelado em Ciências da Computação) – Centro de Ciências Exatas e Naturais, Universidade Regional de Blumenau, Blumenau.
- [45] SCHAEFER, M. T. **Demystifying metadata:** initiatives for web document description. *Information Retrieval & Library Automation*, vol. 33, n. 11, 1998.
- [46] SCHEINKMAN, A; BRADSHAW, S; HAMMOND, K. **Guiding people to information: providing an interface to a digital library using reference as a basis for indexing.** *Proceedings of the 2000 international conference on Intelligent user interfaces.* Pages 37-43. January 2000.
- [47] SIEMENS. Disponível em <http://www.siemens.com.br>. Acesso em: 15 jan 2005.

- [48] SILBERSCHATZ, A; KORTH, K; Sudarshan, S. **Sistema de Banco de Dados**. Terceira Edição. São Paulo: Makron Books, 1999.
- [49] SYBASE. **Sql anywhere studio8 database**. Disponível em <<http://www.sybase.com/products/anywhere/>>, Sybase, 2003 Acesso em: 12 jan 2005.
- [50] TANAKA, K; OOMOTO, E. **Video database systems – recent trends in research and development activities**. *The Handbook of Multimedia Information Management*, Pretince Hall, 1997.
- [51] VAZ, M. S. M. G. **MetaMídia: Um Modelo de Metadados na Indexação e Recuperação de Objetos Multimídia**. Tese de Doutorado. Centro de Informática. Universidade Federal de Pernambuco, 2000, p. 142.
- [52] VLACH, R. **Mobile database procedures in MDBAS**. In *DEXA Workshop*, 2001, pages 559–563.
- [53] W3C (World Wide Web Consortium). Disponível em <http://www.w3.org>. Acesso em 10 abr. 2005.
- [54] WEAVER, M; DELCAMBRE, L; MAIER, D. **A Superimposed Architecture for Enhanced Metadata**. Proceedings of the Third DELOS Network of Excellence Workshop on Interoperability and Mediation in Heterogeneous Digital Libraries. Darmstadt, Germany, 8-9 September 2001.
- [55] WIGLEY, A; WHEELWRIGHT, S. **Microsoft .NET compact framework: corereference**. Washington: Microsoft Press, 2003.