

Sumário

1. Introdução a Aplicações Não-Convencionais
2. Revisão/Dicas de Modelagem Conceitual
3. BD Orientado a Objetos (BDOO)
4. BD Objeto-Relacional (BDOR)
5. BD Temporal (BDT)
6. BD Geográfico (BDG)
7. BD XML
8. BDs nas Nuvens
- 9. Pesquisa em Gerência de Dados na Web**

Pesquisa na Área de BD

- Foco das pesquisas no século XX
 - ***Core Database*** (desenvolvimento de **SGBDs**)
 - Final da década de 60/década de 70
 - Pesquisa em modelos de BDs e linguagens
 - Exemplos: modelo relacional, álgebra relacional e SQL
 - Década de 80
 - Pesquisa em SGBDs eficientes
 - Gerência de transações, processamento de consultas, controle de integridade, segurança dos dados
 - Década de 90
 - Pesquisa em SGBDs não-convencionais
 - BDOO, BDT, BDG, ...

Pesquisa na Área de BD

- Foco das pesquisas no século XXI
 - **Gerência de Dados** (dados em diversas/heterogêneas fontes de dados – geralmente *“out of SGBDs”*)
 - Dados extremamente úteis para consumo humano
 - Diversas finalidades
 - pesquisa, serviços, análises, ...
 - Desafios
 - Descoberta, extração, organização, integração & interoperabilidade, consulta, escalabilidade, ...
 - Dados de diferentes naturezas
 - Estruturados, semi-estruturados, não-estruturados

Dados em Vários Formatos

Poetas:

- **Carlos Drummond de Andrade** nasceu em Itabira, estado de Minas Gerais, no dia 31 de outubro de 1901, vindo a falecer em 17 de agosto de 1987.

Não estruturado (texto livre, txt, pdf)

```
<TABLE>
  <TR>
    <TH>Nome</TH>
    <TH>Nascimento</TH>
    <TH>Óbito</TH>
    <TH>Formação</TH>
  </TR>
  <TR>
    <TD>Carlos D. de Andrade</TD>
    <TD>Itabira/MG, 31/10/1901</TD>
    <TD>17 de agosto de 1987</TD>
    <TD>Farmácia</TD>
  </TR>
</TABLE>
```

Semi-estruturado (HTML)

```
<autores>
  <autor>
    <nome>Carlos D. de Andrade</nome>
    <nasc>Itabira/MG, 31/10/1901</nasci>
    <obito>17 de agosto de 1987</obito>
    <graduacao>Farmácia</graduacao>
  </autor>
  <autor>
    <nome>Mário Quintana</nome>
    <nasc>Alegrete</nasci>
    <obito>1994</obito>
  </autor>
</autores>
```

Semi-estruturado (XML)

Autor

ID	Nome	CidNatal	AnoNascimento	AnoObito
111	Carlos Drummond de Andrade	Itabira	1902	1987
222	Mário Quintana	Alegrete	1906	1994

Estruturado

Pesquisa na Área de BD

- Principal Fonte de Dados Úteis
 - Dados na *Web*
 - Textos, listas, tabelas, formulários, XML, Wikipedia, redes sociais, ...
 - Dados de natureza ÑE, SE e E
- Utopia da Comunidade de Pesquisa em BD
 - *Web* como um grande BD
 - Organizar o “caos” dos dados
 - Acessar de forma estruturada e declarativa, como em BDR utilizando SQL

Esforços de Pesquisa em Dados na Web

- Novos Modelos de Dados para Organizar e Acessar Dados na Web
 - Exemplo: Modelo vetorial e índices invertidos
 - Fontes de dados modeladas como vetores de termos
 - Indexação por termo

<i>Doc</i>	<i>Termos</i>
D1	Drummond, Itabira, outubro, outubro, fazenda, fazenda, escritor
D2	Quintana, Julho, escritor, escritor, poeta, poeta, poeta, praça
D3	Quintana, Drummond, escritor, poeta, praça,

<i>Termos</i>	<i>Lista Invertida</i>
Drummond	(D1,1) (D3,1)
Quintana	(D2,1) (D3,1)
fazenda	(D1,2)
escritor	(D1,1) (D2,2) (D3,1)
poeta	(D2,3) (D3,1)

Esforços de Pesquisa em Dados na Web

- Similaridade
 - Identificar o mesmo objeto em fontes distintas
 - Principais Objetivos
 - Integração de dados
 - Extração e catalogação de entidades de dados na Web
 - Consultas em diversas fontes por dados específicos
- Web Semântica
 - Dar significado a dados “crus”
 - Iniciativas
 - Suporte de dicionários/ontologias/bases de conhecimento
 - Padrões de anotação de dados
 - Padrões de Metadados

Esforços de Pesquisa em Dados na Web

- *Deep Web*

- Infinitude de BDs com informação útil na Web cujo conteúdo e estrutura está “escondido” (nas profundezas...)
- Acesso ao seu conteúdo através de *Web forms*
 - Acervos bibliográficos, ofertas de emprego, revendas de veículos,
- Objetivos
 - Descobrir e acessar o seu conteúdo (metadados e dados)
 - Facilitar buscas integradas por informação útil de diferentes serviços