



BANCO DE DADOS

Gustavo Aurélio Prieto



Dado Versus Informação

Dado

- Registro de um fato.

Informação

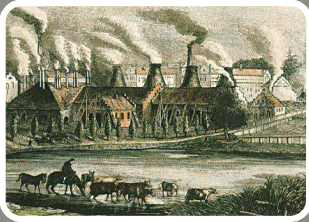
- Dado que foi trabalhado (processado) de forma a obter relevância (utilidade) para uma determinada pessoa ou organização.
- **INFORMAÇÃO = Dado + Conhecimento**

Evolução da Sociedade, Segundo Alvin Toffler



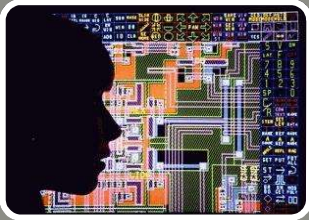
Primeira Onda - Agricultura

- Grande capacidade de produção de alimento, surgem as cidades, estilo de vida cíclico.



Segunda Onda - Revolução Industrial

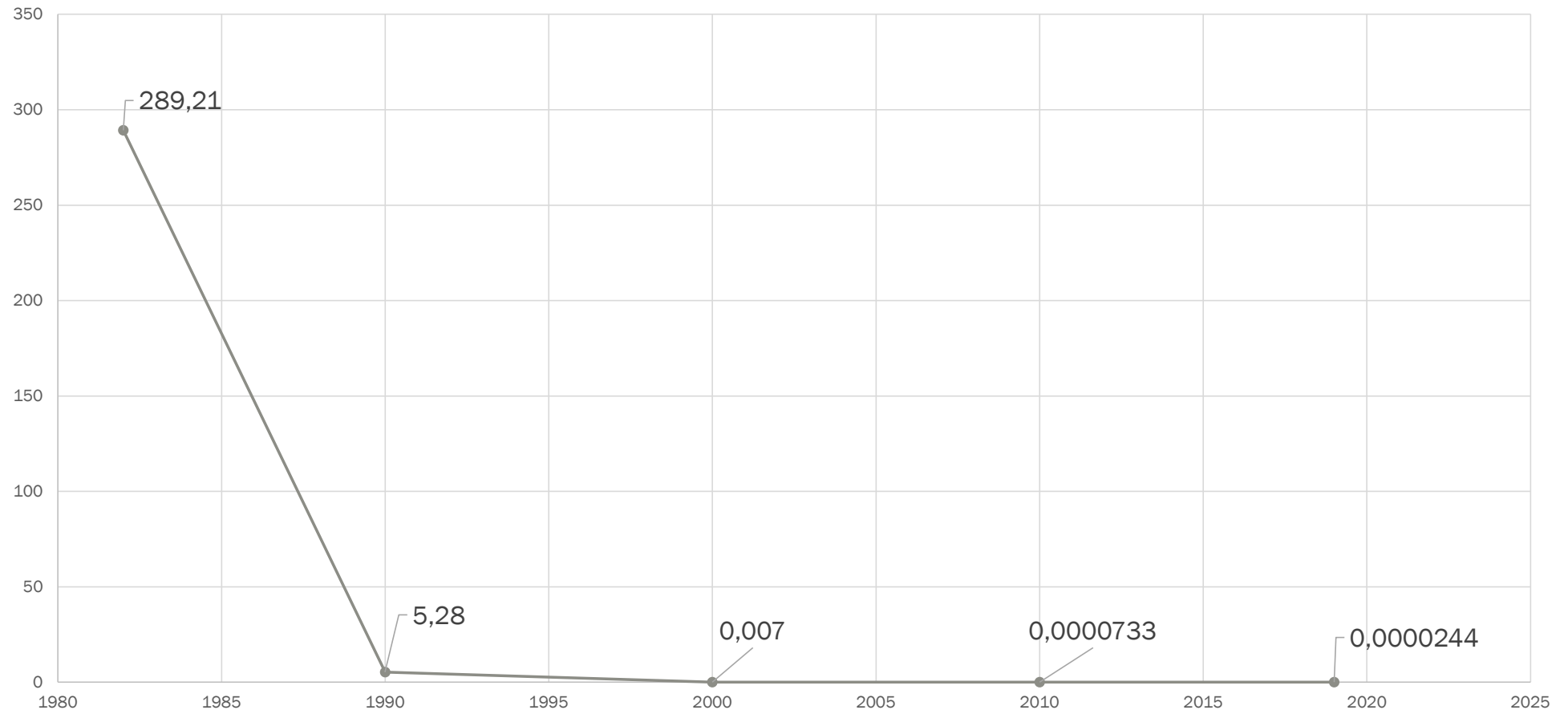
- Substituição do Homem pela Máquina, produção em massa, grandes corporações (Hierarquia e Burocracia)



Terceira Onda

- Sociedade Industrial → Sociedade da Informação
- Alta Tecnologia, informação como “moeda corrente”

\$/MB



Volume de Dados Crescente

Positivo

Negativo

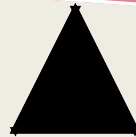
O valor da informação é crescente

A quantidade de dados é cada vez maior

Custo de armazenamento, devido...

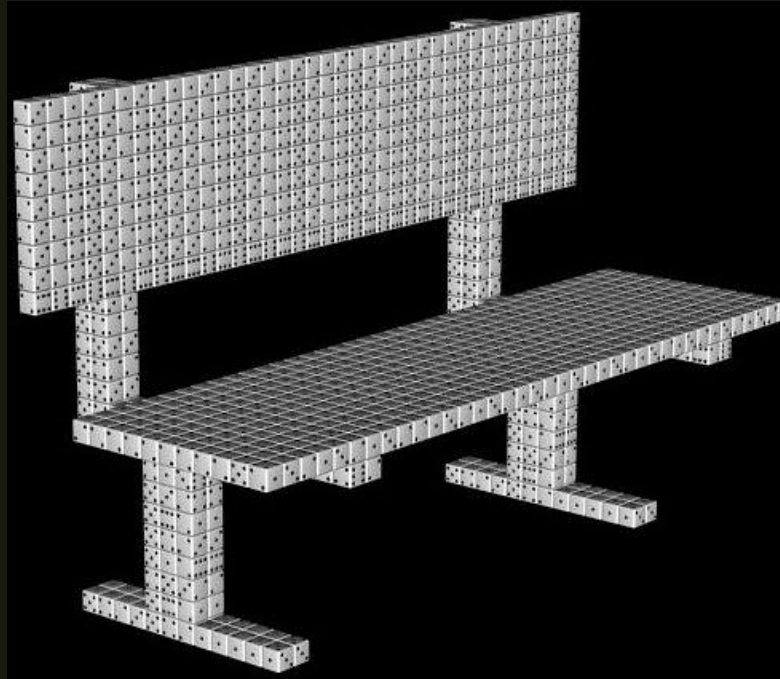
Dificuldade para obter dados relevantes de forma rápida

Alto custo para aquisição e gerenciamento dos dados

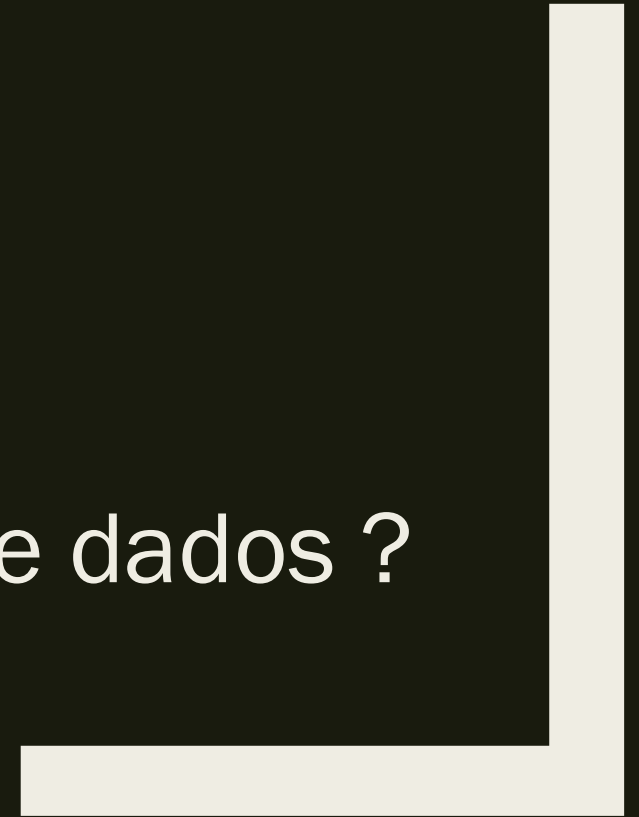




Excesso de
Informação



O que é um banco de dados ?



Banco de Dados



Coleção de dados relacionados.



Propriedades Implícitas:

Representa alguns aspectos do mundo real (minimundo ou Universo de Discurso);

Coleção lógica e coerente de dados com significado inerente;

Possui um grupo de usuários definido e aplicações pré-concebidas, determinadas pelos interesses de seus usuários.

Conceitos Básicos

Banco de Dados: coleção de dados que descrevem as atividades de uma organização. é composto por:

Entidades (cadastros básicos)

Exemplo: estudantes, disciplinas, professores, classes...

Relacionamentos (registro de fatos)

Exemplo: alunos que cursam determinado curso, notas de um estudante...

Exemplo de um Banco de Dados

Cliente

- Nome: char(30) → 30 bytes
- Idade: integer → 6 bytes
- e-mail: char(20) → 20 bytes
- Celular: char(11) → 11 bytes

Tamanho explosivo...

- Um cadastro = 67 Bytes
- Se houverem 100.000 cadastros...
- 6.700.000 B
- Aproximadamente 6,4 MB

Sistema Gerenciador de Banco de Dados (SGBD)

- Coleção de programas que permite aos usuários criar, manter e utilizar um banco de dados.
- Sistema de propósito geral que facilita o processo de definição, construção, manipulação e compartilhamento de um banco de dados entre vários usuários e aplicações.
- Propósito de um SGBD:
 - *Construção;*
 - *Armazenamento;*
 - *Recuperação;*
 - *Manipulação;*
 - *Compartilhamento;*
 - *Segurança.*



ORACLE®



EXEMPLOS DE SGBD

Atores

- Administrador ou DBA (Data Base Administrator)
 - *Acesso, segurança, ocupação de disco, performance, monitoramento.*
- Projetista
 - *Projeto, documentação, análise de requisitos.*
- Desenvolvedor
 - *Desenvolvimento da aplicação, atendimento as necessidades dos usuários.*
- Usuário Final
 - *Pessoas que necessitam dos dados para as suas atividades.*

Processamento de Arquivo VS SGBD

Processamento de Arquivo

- Os dados são armazenados em arquivos, diretamente no HD.
- Cada desenvolvedor cria e mantém um conjunto de arquivos. Todos os aspectos de integridade e segurança devem ser garantidos pelo desenvolvedor.

SGBD

- Natureza auto descritiva do sistema de banco de dados;
- Isolamento entre os programas, os dados e a abstração de dados;
- Suporte para múltiplas visões dos dados;
- Compartilhamento dos dados.

Inconvenientes do Processamento de Arquivos

Redundância e
Inconsistência
dos dados

Dificuldade de
Acesso

Isolamento
dos Dados

Problemas de
Integridade

Problemas de
Atomicidade

Anomalias de
Acesso
Concorrente

Redundância e Inconsistência de Dados

Redundância

- Como os desenvolvedores criam várias soluções no decorrer do tempo, pode ocorrer a duplicação dos mesmos dados em arquivos diferentes.
- Aumento dos custos de armazenamento e processamento.

Inconsistência

- A ocorrência da duplicação dos dados pode levar a inconsistências.
- As cópias de dados serem diferentes entre si.
- Prejuízo incomensurável. Informação incorreta.

Dificuldade de Acesso aos Dados

- Como os dados são acessíveis apenas através de aplicações, caso surja uma necessidade de informação não prevista existem duas opções:
 - *Selecionar uma aplicação já pronta que mais se aproxima da nova solicitação e utilizá-la de maneira incompleta;*
 - *Desenvolver uma nova aplicação que satisfaça a nova necessidade.*
- Ambas insatisfatórias.

Isolamento dos Dados

- Como os dados estão dispersos em arquivos desenvolvidos por diferentes pessoas ocorre:
 - *Diferentes soluções (lógica);*
 - *Diferentes formatos (físico);*
 - *Diferentes localizações (diretórios).*

Problemas de Integridade

- Dependendo dos requisitos, os dados armazenados no banco devem obedecer a certas restrições de consistência.
 - *Restrições de Consistência: regras que os dados precisam obedecer para que façam sentido para a aplicação.*
- Em um sistema mantido por processamento de arquivos, as restrições de consistência são implementadas exclusivamente na aplicação. Isto causa dificuldades pois é difícil manter os diferentes aplicativos em dia. O problema é agravado pelas diferentes cópias de arquivos.

Problemas com Atomicidade

- Atomicidade: uma ação a ser realizada no banco de dados deve ocorrer em sua totalidade ou não ocorrer de forma nenhuma. Ou seja, as ações devem ser levadas ao seu término de forma completa.
- Caso ocorra uma falha o sistema deve se recuperar da mesma revertendo os dados a sua situação antes da ocorrência da mesma.
- Em um sistema de processamento de arquivos a garantia da atomicidade é difícil.

Anomalias de Acesso Concorrente

- Os sistemas modernos em sua maioria são utilizados por diferentes usuários ao mesmo tempo. Isto é chamado de acesso concorrente.
- Este ambiente promove a possibilidade de atualização simultânea do mesmo dado. Para que não ocorram inconsistências é necessário algum tipo de supervisão.
- Prover um sistema de supervisão em um sistema de processamento de arquivos é uma tarefa árdua.

Problemas de Segurança

- Os sistemas de bancos de dados possuem diferentes tipos de usuários que possuem necessidades diferentes.
- Por motivos de segurança não é interessante que usuários tenham acesso a dados que não lhe pertencem ou que não se encaixam na atividade que exerce.
- Também é necessário proibir o acesso de pessoas não autorizadas: espionagem, roubo de dados, sabotagem, etc...
- A manutenção desta segurança é complexa em sistemas de processamento de arquivos.

Abstração de Dados



Nível de Visão

Nível de abstração mais alto

Simplifica a complexidade do banco de dados para o entendimento do usuário final



Nível Lógico

Descreve quais dados estão armazenados e a relação existente entre eles

Utilizado por administradores e projetistas



Nível Físico

Nível de abstração mais baixo

Descreve como os dados estão armazenados

MODELOS DE DADOS EVOLUÇÃO

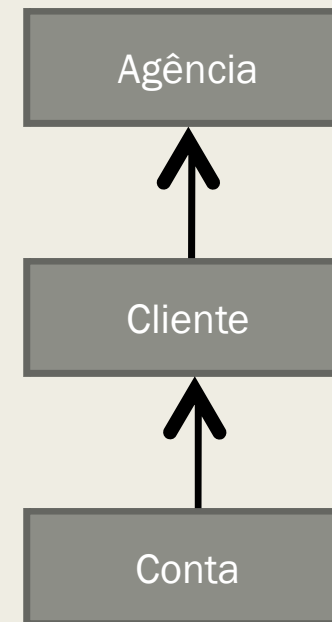


Modelo de Dados

- Coleção de ferramentas conceituais utilizadas para descrever:
 - *Os dados;*
 - *Semântica dos dados;*
 - *As relações existentes entre eles;*
 - *Restrições de consistência.*

Modelos de Dados - Hierárquico

- Relacionamentos pai/filho.
 - *Um-para-um*
 - *Um-para-muitos*
- Características:
 - *Cada pai pode ter vários filhos.*
 - *Cada filho pode ter apenas um pai.*
 - *Duas entidades podem possuir apenas um relacionamento.*
 - *Qualquer recuperação de dados deve obrigatoriamente percorrer a estrutura da árvore. (pai → filho → neto...)*
- Exemplos:
 - *DB2*



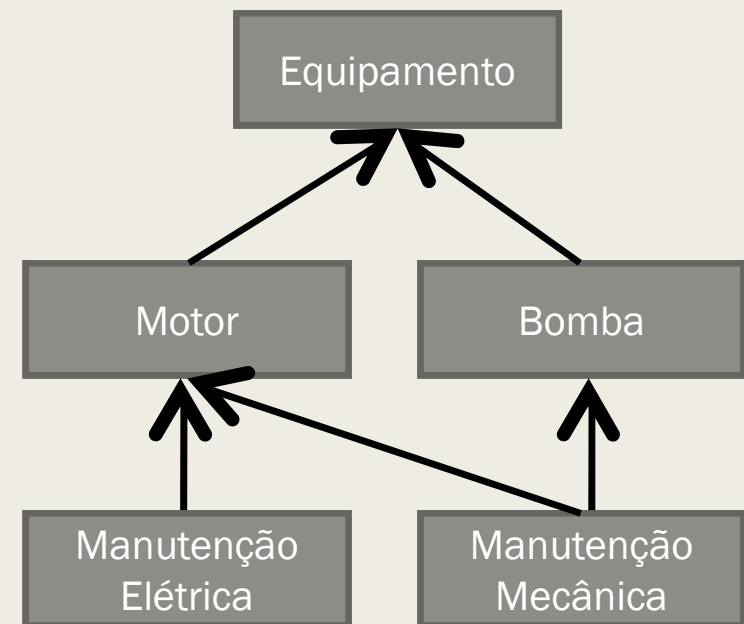
Modelos de Dados - Rede

- Permite a existência de entidades pais com muitos filhos e de entidades filhos com muitos pais.
- O seu diagrama se assemelha a um grafo.
- Foi criado por Charles Bachman:
 - 1969
 - Consórcio CODASYL (*Conference on Data Systems Languages*)



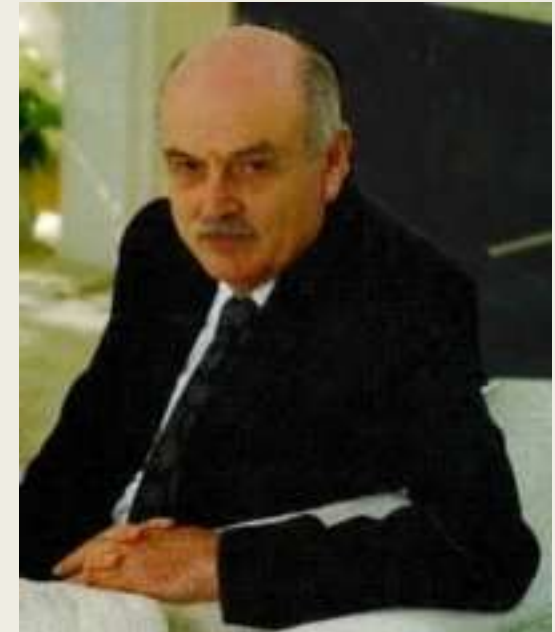
Modelos de Dados - Rede

- Permite que entidades pai possuam vários filhos e vice-versa, ou seja, permite que as entidades possuam mais de um relacionamento.
- Permite modelagens mais naturais e próximas da realidade.
- Porquê não se firmou no mercado...
 - *Percebendo as falhas do modelo hierárquico a IBM lançou um release dos seus produtos (ISM e DL/I) incluindo extensões “semi-rede”*
 - *O surgimento do modelo relacional*
- Exemplos:
 - *IDMS, Turbolmage*



Modelos de Dados - Relacional

- Baseia-se no princípio de que todos os dados estão armazenados em relações (tabelas).
- Toda a sua definição é baseada na lógica de predicados e na teoria dos conjuntos.
- Foi o primeiro modelo de dados descrito em linguagem formal (matemática).
- Inventado por Edward Frank Codd
 - 1970
 - Aprimorado por Chris Date e Hugh Darwen



Edward Frank Codd



Chris Date

Modelos de Dados - Relacional

- Coleção de ferramentas conceituais que descrevem:
 - *Dados*
 - *Relações entre os dados*
 - *Semântica dos dados*
 - *Restrições de consistência*
- Utiliza um conjunto de tabelas para representar as tabelas e as relações entre elas.
- Simplicidade Conceitual
- Maioria dos SGBD do mercado usam o modelo Conceitual
- Exemplos:
 - *MySQL, Oracle, SQL Server, etc...*

