



# Fundamentos de Banco de Dados



Prof. Msc Marlus Dias Silva

marlus.guitar@gmail.com

(64)98133-7874

<https://github.com/marlovi/aulas.git>



# Introdução

- A importância da informação para a tomada de decisões nas organizações tem impulsionado o desenvolvimento dos sistemas de processamento de informações.
- Algumas ferramentas:
  - processadores de texto (editoração eletrônica),
  - planilhas (cálculos com tabelas de valores),
  - Sistemas de Gerenciamento de Bancos de Dados - SGBDs (armazenamento de grandes volumes de dados, estruturados em registros e tabelas, com recursos para acesso e processamento das informações).

# O que é um Banco de Dados?

- Banco de Dados: é uma coleção de **dados** inter-relacionados, representando **informações** sobre um domínio específico
- Exemplos:
  - Lista telefônica
  - Controle do acervo de uma biblioteca
  - Sistema de controle dos recursos humanos de uma empresa

# O que é um Banco de Dados?

- Entretanto, o uso do termo banco de dados é geralmente mais restrito e possui algumas propriedades implícitas:
  - Representa algum aspecto do mundo real (minimundo ou Universo de Discurso);
  - É uma coleção lógica e coerente de dados com algum significado inerente
    - Uma organização aleatória de dados não pode ser considerada um banco de dados;
  - É projetado, construído e povoado por dados, atendendo a uma proposta específica
    - Possui um grupo provável de usuários e algumas aplicações preconcebidas

# O que é um Banco de Dados?

- Em outras palavras, um BD possui alguma fonte da qual os dados são derivados, algum grau de interação com eventos do mundo real e um público que está ativamente interessado no conteúdo do banco de dados
- Outros aspectos dos bancos de dados:
  - Podem ser de **qualquer tamanho** e de **complexidade variável**;
  - Precisam estar **organizados** e **gerenciados** de forma que os usuários possam pesquisar, recuperar e atualizar os dados quando necessário;
  - Podem ser gerados e mantidos manualmente ou de forma automatizada (por um grupo de programas de aplicações específicas ou por um **Sistema Gerenciador de Banco de Dados**);

# O que é um Sistema de gerenciamento Banco de Dados (Sgbd)?

- Um SGBD é constituído por um **conjunto de dados** associados a um **conjunto de programas** para acesso a esses dados
- É um sistema de software de propósito geral que facilita o processo de **definição, construção, manipulação e compartilhamento** de bancos de dados entre vários usuários e aplicações
  - Definição → envolve especificar os tipos de dados, estruturas e restrições para os dados armazenados
  - Construção → é o processo de armazenar os dados em alguma mídia apropriada, controlada pelo SGBD
  - Manipulação → inclui funções para incluir, modificar e apagar dados do Banco de Dados
  - Compartilhamento → permite aos múltiplos usuários e programas acessar, de forma concorrente, o banco de dados.

# O que é um Sistema de gerenciamento Banco de Dados (Sgbd)?

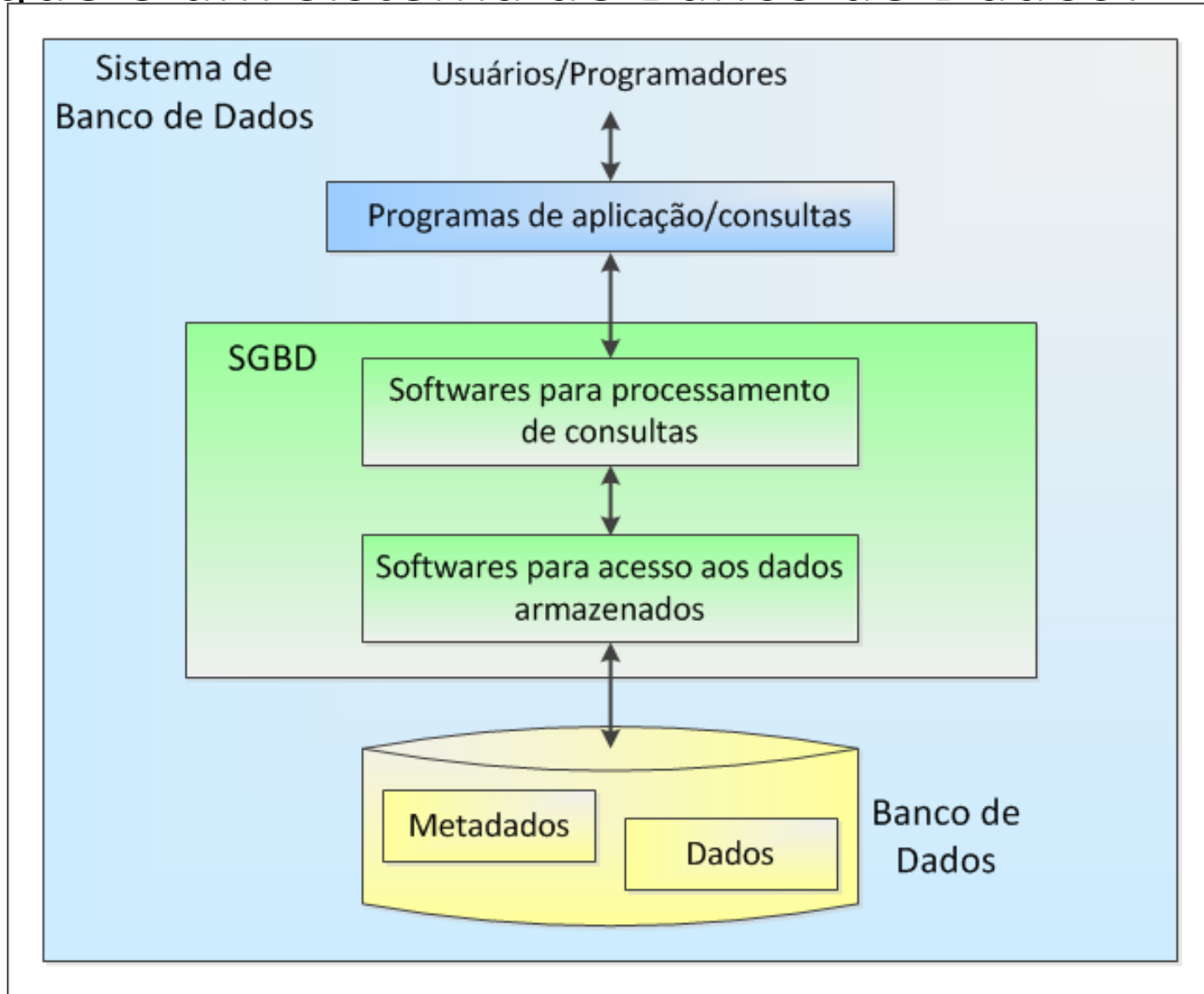
- Objetivo do SGBD:
  - Proporcionar um ambiente **conveniente** e **eficiente** para a **recuperação** e **armazenamento** das informações do BD
- SGBDs são projetados para gerir grandes volumes de informações do BD. Isso implica:
  - Definir as estruturas de armazenamento
  - Definir os mecanismos para a manipulação das informações
  - Garantir a segurança das informações armazenadas contra eventuais problemas do sistema
  - Impedir tentativas de acesso não autorizadas
  - Evitar a ocorrência de resultados anômalos

# O que é um Sistema de Banco de Dados?

- É um sistema de manutenção de registros por computador, envolvendo quatro componentes principais:
  - Dados, hardware, software e usuários
- Exemplo: Sistema Bancário responsável por todas as informações de seus clientes e suas contas
  - Programas para realizar as diversas funcionalidades
- NÃO confundir Sistema de Bancos de Dados com a parte do S.O. chamada Gerenciador de Arquivos



# O que é um Sistema de Banco de Dados?



# Objetivo de um Sistema de Banco de Dados

- Isolar os usuários dos detalhes mais internos do banco de dados
  - Abstração de dados
- Prover independência de dados às aplicações (estrutura física de armazenamento e à estratégia de acesso).

# Sistema de Banco de Dados x Gerenciador de Arquivos

- Desvantagens de um Gerenciador de Arquivos frente ao SGBD:
  - Inconsistência e redundância de dados
    - Vários arquivos e aplicações criadas e mantidas por diversos programadores
      - Formatos dos arquivos e linguagens de programação diferentes
    - Mesma informação repetida em vários lugares
      - Inconsistência nos dados
  - Dificuldade de acesso aos dados
    - Não existência de programas para atender as diversas solicitações.
    - Ex: Lista com o nome dos clientes que moram no CEP 29500-000

# Sistema de Banco de Dados x Gerenciador de Arquivos

- Desvantagens de um Gerenciador de Arquivos frente ao SGBD:
  - Isolamento de dados
    - Dados dispersos em vários arquivos com formatos diferentes
  - Problemas de integridade
    - Dificuldade de implementação de modificações em todos os programas e/ou dados presentes em diferentes arquivos
  - Problemas de atomicidade
    - Todo sistema computacional está sujeito a falhas
    - É preciso garantir que certas operações ocorram de forma atômica, ou seja, ocorre por completo ou não ocorre
    - É difícil manter essa propriedade em sistemas de processamento de arquivos

# Sistema de Banco de Dados x Gerenciador de Arquivos

- Desvantagens de um Gerenciador de Arquivos frente ao SGBD:
  - Anomalias no acesso concorrente
    - Ex: Dois clientes efetuando um saque da mesma conta
  - Problemas de segurança
    - É difícil garantir a efetividade das regras de segurança em sistemas de processamento de arquivos
- Tais desvantagens provocaram o desenvolvimento de Sistemas de Gerenciamento de Banco de Dados

# Visão dos dados

- Um SGBD é um conjunto de arquivos e programas inter-relacionados que permitem ao usuário o acesso para consultas e alterações desses dados
- Um dos maiores benefícios de um BD é fornecer ao usuário uma **visão abstrata dos dados**
  - Ocultar determinados detalhes sobre a forma de armazenamento e manutenção desses dados

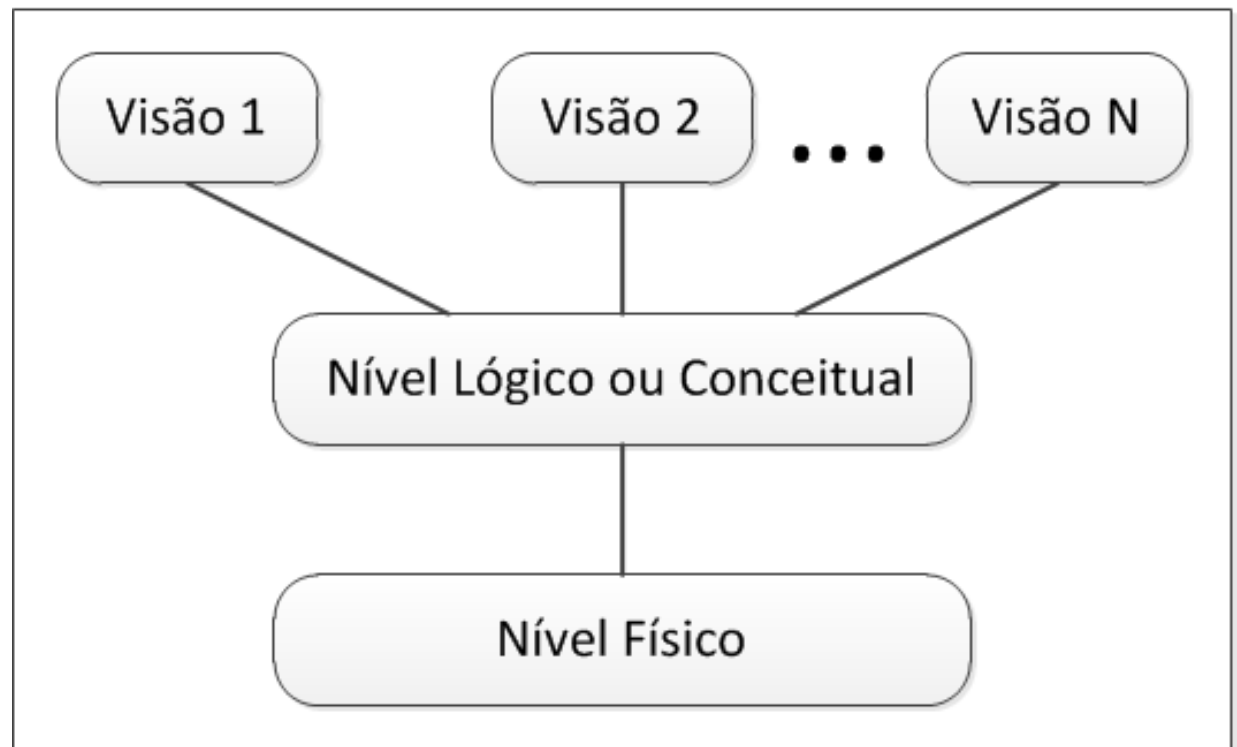
# Visão dos dados

- Abstração de Dados
  - Facilitar a interação dos usuários (muitas vezes leigos) com o sistema de BD

Nível de Visão dos  
Usuários

Nível de  
Programadores e  
Administradores

Nível de  
Armazenamento



# Visão dos dados

- Abstração de Dados

- Existem 03 níveis de abstração:

- Nível Físico → descreve **como** esses dados estão de fato armazenados. As estruturas de dados de armazenamento são descritas em detalhes
    - Nível Lógico ou Conceitual → descreve **quais** dados estão armazenados no BD e **quais** os inter-relacionamentos entre eles
      - É utilizado pelos Administradores do Banco de Dados (DBA) que precisam decidir quais informações devem pertencer ao BD
    - Nível de Visão → descreve apenas **parte** do BD
      - A maioria dos usuários do BD não precisam conhecer todas as informações do BD, apenas parte delas
      - Para cada parte é criada uma visão

- Níveis de Abstração x Linguagens de Programação



# Visão dos dados

- Abstração de Dados
  - Instâncias e Esquemas:
    - Todo BD muda com o tempo. Essas mudanças são ocasionadas por informações inseridas ou excluídas do BD
    - **Instância do Banco de Dados**
      - Conjunto de informações contidas em determinado BD, em um dado momento
    - **Esquema de Banco de Dados**
      - Nome dado ao projeto geral do BD
      - Os esquemas são alterados com pouca frequência
  - Instância e Esquema x Linguagens de Programação

# Visão dos dados

- Abstração de Dados

- Independência de Dados:

- Independência de dados é a capacidade de modificar a definição dos esquemas em determinado nível, sem afetar o esquema do nível superior
    - Existem 02 níveis de independência de dados:
      - **Independência de dados física** → capacidade de modificar o esquema físico sem alterar qualquer programa de aplicação. Tais modificações são feitas para **aprimorar o desempenho**
      - **Independência de dados lógica** → capacidade de modificar o esquema lógico sem alterar qualquer programa de aplicação. Tais modificações são necessárias sempre que **há alteração da estrutura lógica** do BD
    - A independência lógica é mais difícil de conseguir do que a independência física

# Modelo de dados

- O que é um Modelo de Dados?
  - Conjunto de ferramentas conceituais usadas para a **descrição de dados**, **relacionamentos entre dados**, **semântica** de dados e **regras de consistência**
- Existem 03 diferentes grupos:
  - Modelos lógicos com base em objetos, modelos lógicos com base em registros e modelos físicos

# Modelo de dados

- Modelo Lógico com Base em Objetos
  - Usados na descrição de dados no **nível lógico** e de **visões**
  - Existem vários modelos nessa categoria, mas os mais usados são:
    - **Modelo de Entidade-Relacionamento**
    - **Modelo Orientado a Objetos**

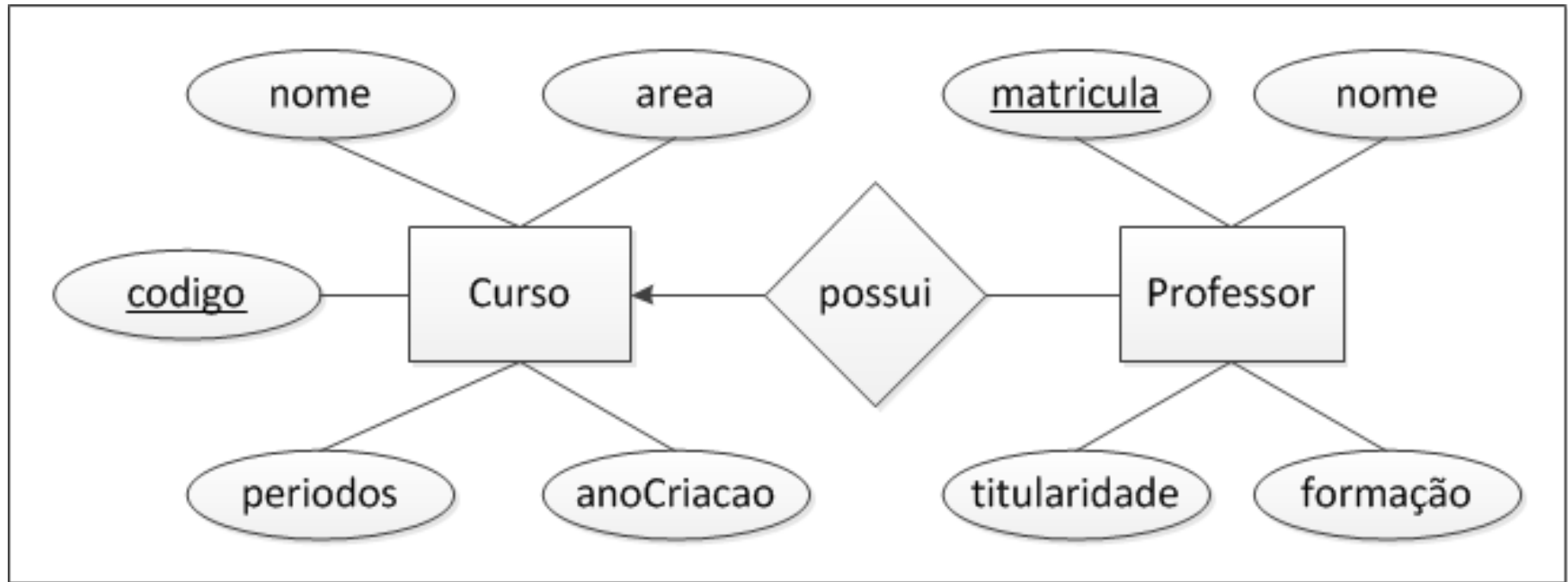
# Modelo de dados

- Modelo de Entidade-Relacionamento (E-R)

- Tem por base a percepção do mundo real como um conjunto de objetos básicos, **entidades**, e do **relacionamento** entre eles
- **Entidade** → é uma “coisa” ou um “objeto” do mundo real que pode ser identificado por outros objetos
- Exemplos:
  - Pessoa, conta bancária, instituição, computador, etc...
- **Relacionamento** → associação entre entidades
- Exemplos:
  - Estuda (estudantes e instituições de ensino), depositante (pessoa e conta bancária)

# Modelo de dados

- Modelo de Entidade-Relacionamento (E-R)
  - Diagrama E-R



# Modelo de dados

- Modelo Orientado a Objetos

- Tem por base um conjunto de objetos
  - Objetos contêm **atributos** que armazenam valores e **métodos** que operam sobre o objeto
  - **Classes** que são conjuntos de objetos
- Existem dois níveis de abstração de dados
  - A parte visível → Métodos
  - A parte invisível → atributos
- Neste modelo o código executável é parte integrante do modelo de dados

# Modelo de dados

- Modelo Lógico com Base em Registros

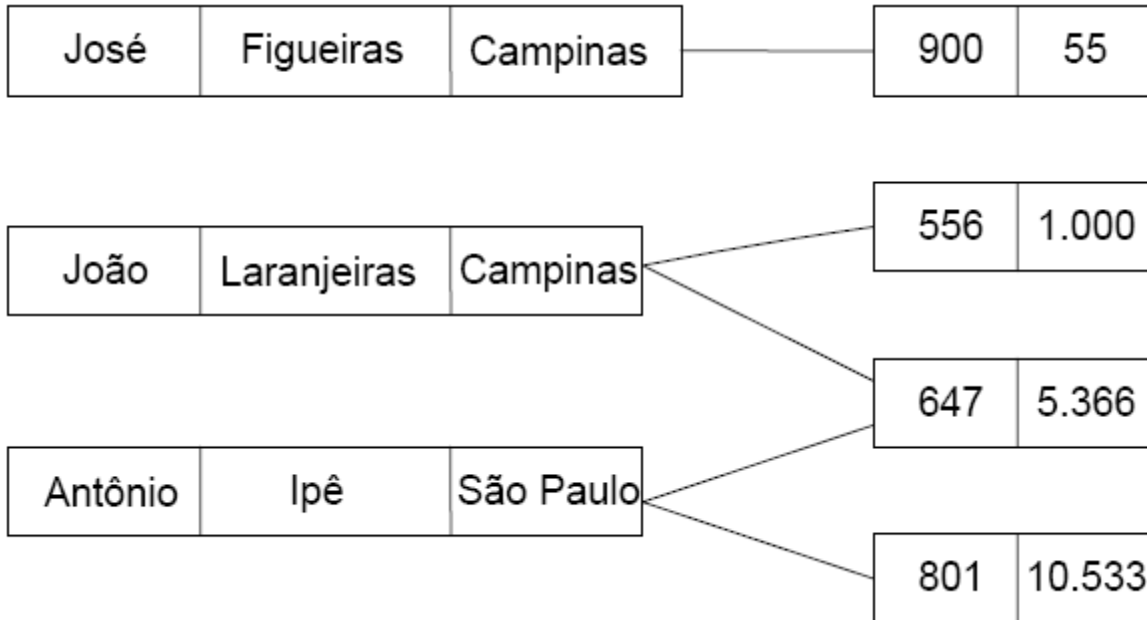
- Usados na descrição de dados no **nível lógico** e de **visões**
- São assim chamados porque o BD é estruturado por meio de registros de formato fixo de todos os tipos
  - Cada registro possui um número fixo de campos/atributos
  - Cada campo possui um tamanho fixo
- Os 03 modelos de dados com base em registros mais conhecidos são:
  - **Modelo de Rede**
  - **Modelo Hierárquico**
  - **Modelo Relacional**



# Modelo de dados

## • Modelo de Rede

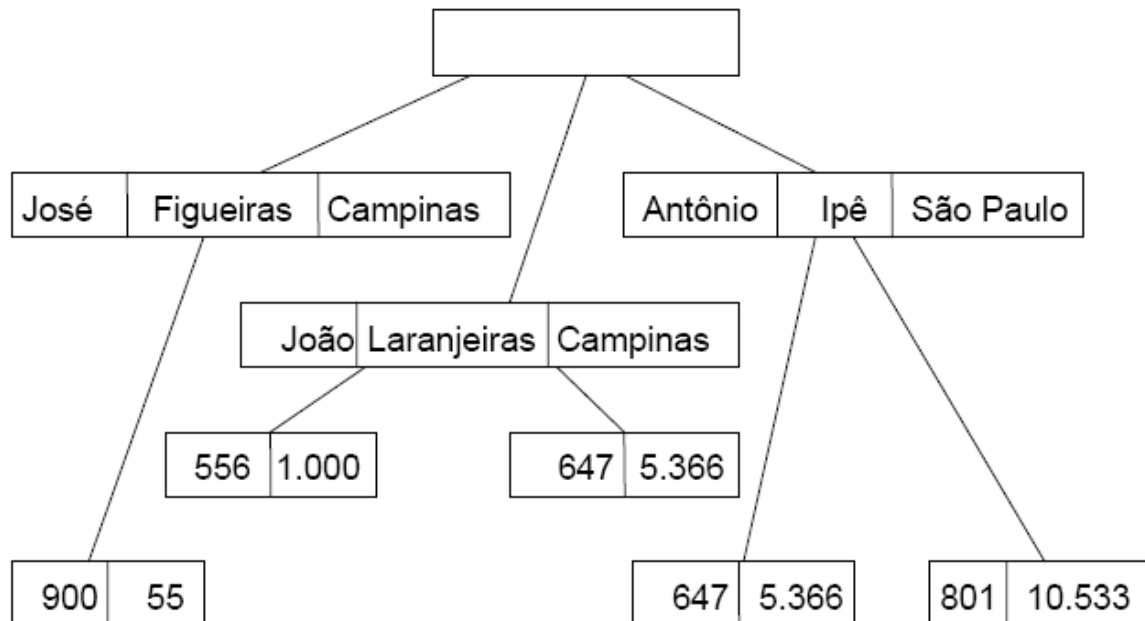
- Usa um conjunto de registros (como em C) e as relações entre os registros são representadas por links (como se fossem ponteiros)



# Modelo de dados

- Modelo Hierárquico

- Similar ao modelo de rede
  - Dados e suas relações são representados por registros e links
- A diferença é que os registros estão organizados em árvores ao invés de gráficos arbitrários



# Modelo de dados

## • Modelo Relacional

- Usa um conjunto de tabelas para representar tanto os dados quanto a relação entre eles

Tabela Cliente (dados)

cód-cliente	nome	rua	cidade
015	José	Figueiras	Campinas
021	João	Laranjeiras	Campinas
037	Antônio	Ipê	São Paulo

Tabela Cliente-Conta  
(relacionamento)

cód-cliente	nro-conta
015	900
021	556
021	647
037	647
037	801

Tabela Conta (dados)

nro-conta	saldo
900	55
556	1.000
647	5.366
801	10.533

# Modelo de dados

- **Modelo Relacional**
  - Este modelo não utiliza ponteiros e nem links. Logo possibilitou-se o desenvolvimento de fundamentos matemáticos para a sua definição
    - Álgebra Relacional

# Modelo de dados

- Modelo Físico de Dados

- Usados na descrição de dados no nível físico
- Existem apenas dois modelos amplamente conhecidos:
  - Modelo Unificado (Unifying model)
  - Modelo de Partição de Memória (Frame-memory model)
- Captam os aspectos de implementação do sistema de banco de dados
  - Não serão abordados nesta disciplina

# Linguagens de banco de dados

- O sistema de BD provê dois tipos de linguagens:
  - Específica para os esquemas do BD
    - Linguagem de Definição de Dados ou Data-Definition Language (DDL)
  - Expressar consultas e atualizações
    - Linguagem de Manipulação de Dados ou Data-Modeling Language (DML)

# Linguagens de Definição de dados

- Define um conjunto de expressões usadas para especificar o esquema de BD
  - Linguagem de Definição de Dados ou Data-Definition Language (DDL)
- Os elementos que formam o esquema de BD são armazenados em um conjunto de tabelas que constituem o **dicionário de dados** ou **diretório de dados**
  - É um arquivo de **metadados** → dados a respeito de dados

# Linguagens de manipulação de dados

- Por manipulação pode-se dizer:
  - Recuperação, inserção, remoção e modificação
- O objetivo desta linguagem é proporcionar uma interação eficiente entre homens e sistema
- A Linguagem de Manipulação de Dados (**DML**) viabiliza o acesso ou a manipulação dos dados de forma compatível ao modelo de dados apropriado
  - **DMLs Procedurais** → exigem que o usuário especifique **quais** dados são necessários e **como** obtê-los
  - **DMLs Não-procedurais** → exigem que o usuário especifique **quais** dados são necessários **sem** especificar como obtê-los



# Linguagens de manipulação de dados

- DMLs Não-procedurais
  - Vantagem
    - São normalmente mais fáceis de aprender e de usar
  - Desvantagem
    - Podem gerar código menos eficiente que os gerados por linguagens procedurais
- A parte da DML responsável pela recuperação de informações é chamada **linguagem de consulta** (query language)

# Gerenciamento de transações

- O que é uma transação?
  - É uma **coleção de operações** que desempenha uma **função lógica única** dentro de uma aplicação do sistema de banco de dados
  - Cada transação é 01 unidade de **atomicidade** e **consistência**
    - Mas durante a execução da transação podem ocorrer estados temporariamente inconsistentes do banco de dados
      - Essa inconsistência pode gerar problemas caso ocorra uma falha

# Gerenciamento de transações

- Frequentemente, muitas operações em um BD constituem uma única unidade lógica de trabalho
- Exemplo: Transferência de fundos eletrônica
  - Essa transferência deve acontecer como um todo ou nada deve ser feito
    - Esse tudo ou nada é chamado **Atomicidade**
  - Além do mais a soma do saldo da conta A com a conta B deve ser preservada
    - Essa exigência de “corretude” é chamada **Consistência**
  - Após a transferência, os novos valores de A e B devem persistir
    - Essa persistência é chamada **Durabilidade**

# Gerenciamento de transações

- Assegurar as propriedades de atomicidade e durabilidade é responsabilidade do sistema de banco de dados, em especial o componente de **gerenciamento de transações**
- Por causa dos vários tipos de falhas possíveis, uma transação pode não se completar com sucesso e o BD precisa retornar ao estado anterior em que se encontrava antes do início desta transação

# Falhas e concorrência

- O **sistema de BD** é responsável por **detectar as falhas** e **recuperar o BD**, garantindo seu retorno ao último estado consistente
- É responsabilidade do **gerenciador de controle de concorrência** controlar a interação entre transações concorrentes de modo a garantir a consistência do BD

# Administrador de banco de dados

- O **Administrador do Banco de Dados (DBA)** é a pessoa que centraliza o **controle** tanto dos **dados** armazenados quanto dos **programas** que acessam esses dados
- Dentre as funções do DBA estão:
  - Definição do esquema
  - Definição da estrutura de dados e métodos de acesso
  - Esquema e modificações na organização física
  - Fornecer autorização de acesso ao sistema
  - Especificações das regras de integridade

# Usuários de banco de dados

- Há 04 tipos de usuários de sistema de BD
- O que diferencia é o modo de interação com o sistema
  - Programadores de Aplicação
    - Profissionais em computação que interagem com o sistema através de programas escritos em alguma linguagem de programação e que fazem chamadas DML ao sistema
  - Usuários Sofisticados
    - Formulam suas solicitações ao BD por meio de linguagens de consultas

# Usuários de banco de dados

- Há 04 tipos de usuários de sistema de BD
  - **Usuários Especialistas**
    - Escrevem aplicações especializadas de BD
    - Ex.: Sistemas baseados em Conhecimento, Sistemas de armazenagem de dados gráficos e de áudio
  - **Usuários Navegantes**
    - Usuários comuns que interagem com o sistema chamando um dos programas aplicativos permanentes
    - Ex.: Usuário que pede a transferência de R\$ 50,00 da conta A para a conta B



# Visão geral da estrutura do sistema

- Um sistema de BD está dividido em módulos específicos de modo a atender a todas as funções do sistema
- Os componentes funcionais podem ser divididos pelos **componentes de processamento de consultas** e pelos **componentes de administração de memória**
- Os componentes de processamento de consultas incluem:
  - **Compilador DML** → traduz comandos DML em instruções inteligíveis ao componente de tratamento de consultas
  - **Pré-compilador para comandos DML** → convertem comandos DML em chamadas de procedimentos da linguagem hospedeira

# Visão geral da estrutura do sistema

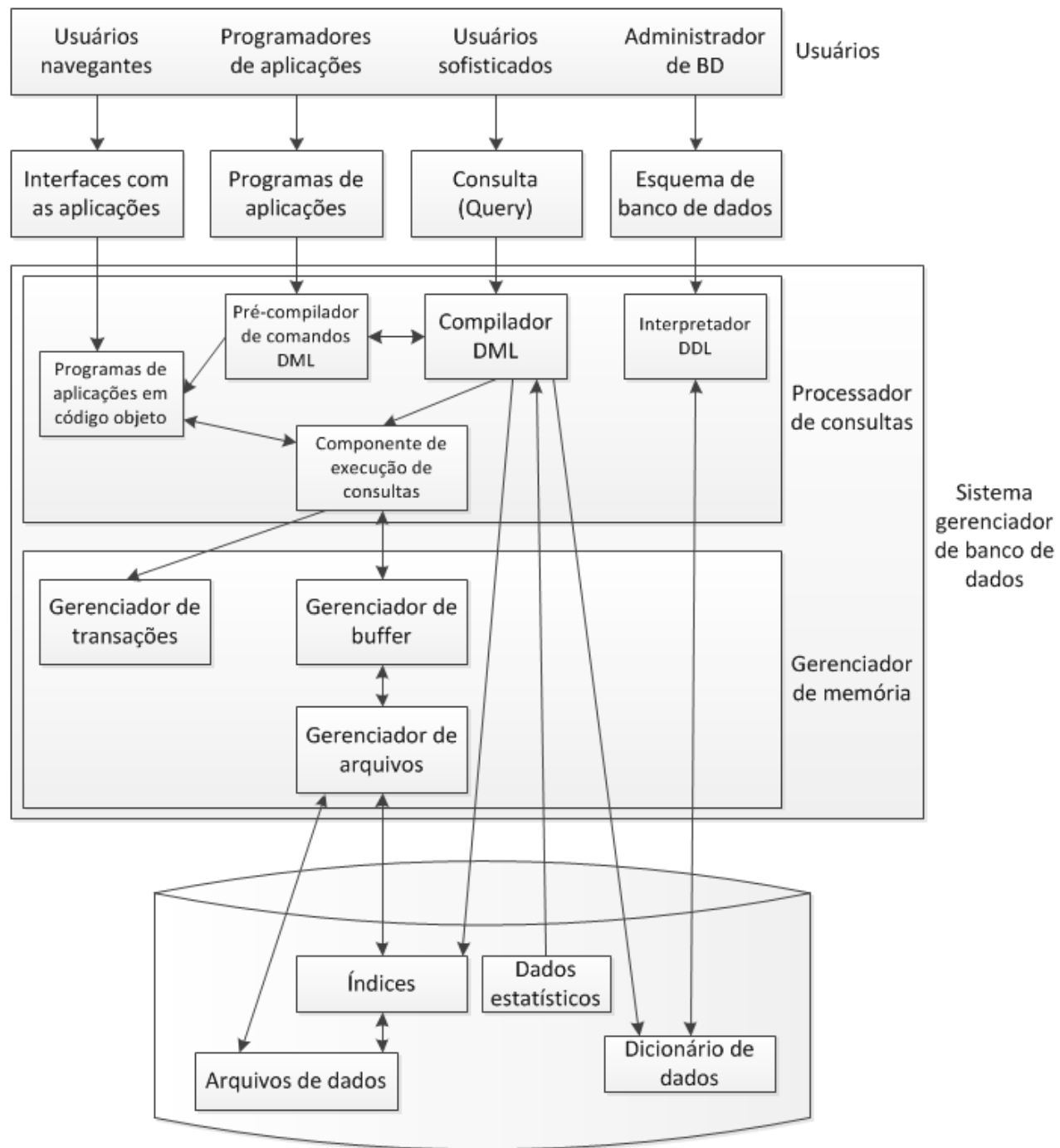
- Os componentes de processamento de consultas incluem:
  - **Interpretador DDL** → interpreta os comandos DDL e registra-os em um conjunto de tabelas que contêm metadados
  - **Componentes de tratamento de consultas** → executam instruções de baixo nível geradas pelo compilador DML

# Visão geral da estrutura do sistema

- Os componentes de administração de memória incluem:
  - **Gerenciamento de autorizações e integridade** → testam o cumprimento das regras de integridade e a permissão ao usuário no acesso ao dado
  - **Gerenciamento de transações** → garante que o BD estará em estado consistente e que transações concorrentes executem sem conflitos
  - **Administração de arquivos** → gerencia a alocação de espaço no armazenamento em disco e as estruturas de dados usadas
  - **Administração de buffer** → gerencia o envio de dados do disco para a memória principal

# Visão geral da estrutura do sistema

- Ainda existem algumas estruturas de dados exigidas como parte da implementação física do sistema
  - **Arquivos de dados** → armazena o próprio BD
  - **Dicionário de dados** → armazena os metadados relativos à estrutura do BD
  - **Índices** → proporcionam acesso rápido aos itens de dados
  - **Estatísticas de dados** → armazenam informações estatísticas relativas aos dados contidos no BD
    - Usadas pelo processador de consultas para otimizar a execução de uma consulta



# Bibliografia

- SILBERSCHATZ, Abraham; KORTH, Henry F.; SUDARSHAN, S.. Sistema de banco de dados. 3 ed. São Paulo: Pearson Makron Books, 2005, Cap. 1