

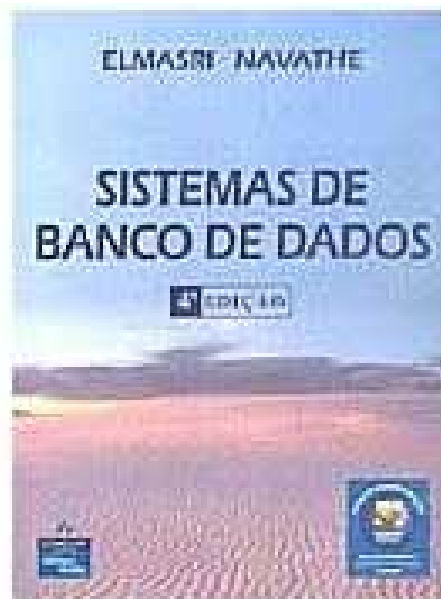
Bancos de Dados II

Rodrigo Salvador Monteiro
salvador@ic.uff.br

Bibliografia

- ▶ Livro texto

- ▶ Elmasri, R., Navathe, S., Sistemas de Bancos de Dados, 4^a edição, Editora Pearson, 2005





Conceitos e Arquitetura

Tópicos

- ▶ Definições
- ▶ Arquitetura
- ▶ Modelo/Esquema

Definições



Definições

- ▶ **Dados**

- ▶ conjunto de valores brutos

- ▶ **Informação**

- ▶ dados associados a seu significado, interpretação, semântica

- ▶ **Conhecimento**

- ▶ informações aplicadas em algum contexto (uso, descoberta, ...)



Exemplo

Dados

Material Bruto
Pouco ou nenhum significado
para o leitor ou a audiência

70 – 78 – 84
63 – 59 – 58
13 – 14 – 19

Informação

Dados processados, interpretados
Dados com contexto
Dados + metadados

*Crimes violentos letais intencionais
(2008-2009-2010)**

Alemão: 70 – 78 – 84
Manguinhos: 63 – 59 – 58
Cantagalo: 13 – 14 – 19

Conhecimento

Estabelece relações entre pedaços de
informação
Conjunto organizado e compreensível
de informações
Pode ser aplicado em situações reais
base para tomar decisões...



Qual o local
estrategicamente mais
favorável para
implantarmos a próxima
UPP?

*dados fictícios

Definições

- ▶ **Base de Dados**

- ▶ Representação é livre
- ▶ Arquivos texto (flat files)
- ▶ Nem sempre provê descrições dos dados armazenados

- ▶ **Banco de Dados**

- ▶ Obrigatoriamente provê informação (dados descritos)
- ▶ Dados são representados segundo um padrão

- ▶ **Gerência de Dados**

- ▶ Conjunto de funcionalidades para armazenamento e recuperação dos dados
 - ▶ garantindo consistência, desempenho e integridade dos dados
- ▶ SGBDs: Sistemas de Gerência de Bancos de Dados



Base de Dados - GenBank

LOCUS	SCU49845	5028 bp	DNA	PLN	21-JUN-1999
DEFINITION	Saccharomyces cerevisiae TCP1-beta gene, partial cds, and Axl2p (AXL2) and Rev7p (REV7) genes, complete cds.				
ACCESSION	U49845				
VERSION	U49845.1 GI:1293613				
KEYWORDS	.				
SOURCE	Saccharomyces cerevisiae (baker's yeast)				
ORGANISM	Saccharomyces cerevisiae Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.				
REFERENCE	1 (bases 1 to 5028)				
AUTHORS	Torpey,L.E., Gibbs,P.E., Nelson,J. and Lawrence,C.W.				
TITLE	Cloning and sequence of REV7, a gene whose function is required for DNA damage-induced mutagenesis in Saccharomyces cerevisiae				
JOURNAL	Yeast 10 (11), 1503-1509 (1994)				
MEDLINE	95176709				
PUBMED	7871890				
FEATURES	Location/Qualifiers				
CDS	<1..206 /codon_start=3 /product="TCP1-beta" /protein_id="AAA98665.1" /db_xref="GI:1293614" /translation="SSIYNGISTSGLDLNNGTIADMRQLGIVESYKLKRAVVSSASEA AEVLLRVDNIIIRARPRTANRQHM"				
gene	687..3158 /gene="AXL2"				

Banco de Dados - GUS DB

id	sequence	length	description
21	gtaggagttctggagctgctggctggagaggaggggtggacgaagctctc...	2322	Homo sapiens hypothetical protein LOC12387
25	aggaagctcctataaagggctgcatctctggactcacagctatcagac...	2258	Homo sapiens oviductal glycoprotein 1, 120kD
26	aagcttgatcaagcgctccagtgatgcccggcagcgcgggcagggcct...	101016	Gordonia westfalica plasmid pKB1, complete
27	gtacggaattccgtacggaattccgtcagcctctggccaagatggtgaa...	241	TcTR-1354 TcTR Trypanosoma cruzi cDNA cl
28	gtacggaatttttaactggtttaaaattaataataaagagaaaaga...	253	TcTR-1339 TcTR Trypanosoma cruzi cDNA cl
29	gttttttttttttttttctcggttttaatttttatgttttt...	279	TcTR-1102 TcTR Trypanosoma cruzi cDNA cl
30	gtaagatcgctgccactgcattctagcctgagcaccaaagtgagactctg...	88	TcTR-878 TcTR Trypanosoma cruzi cDNA clo
31	gtacggcttccgtttaattttcagacagggcctgtctttgttactcaga...	156	TcTR-713 TcTR Trypanosoma cruzi cDNA clo
32	ggacggaattccgtaggaattgcttgcgccaggaggtcaagaacagtct...	91	TcTR-536 TcTR Trypanosoma cruzi cDNA clo
33	gtactttcaatagatcgacgagggagctgctctgctacgtacgaaac...	128	TcTR-496 TcTR Trypanosoma cruzi cDNA clo
34	gtacggaattccgtctgtactccagcctgggtgacagagaaagactctg...	84	TcTR-288 TcTR Trypanosoma cruzi cDNA clo
35	gtacggaattccgtacggaattccgtttgtaatttcttttaagtagtt...	481	TcTR-61 TcTR Trypanosoma cruzi cDNA clon
36	gtgatacattttctgtagatttaggacttgcttctattgatcctgattc...	313	TcTR-6 TcTR Trypanosoma cruzi cDNA clone
37	atccaattccggacggaattcctgatacagtttctgtactatattgcct...	193	TcTR-1507 TcTR Trypanosoma cruzi cDNA cl
38	gcggacggaattccggatacagtttctgtactatattgggcctccaaag...	311	TcTR-1506 TcTR Trypanosoma cruzi cDNA cl
39	attccggacggaattccgtacggattccgtttgaattccgttttttc...	146	TcTR-1505 TcTR Trypanosoma cruzi cDNA cl
40	ggcggctcgttttttttttttgggtgtggttcctccgccaactccgaaa...	154	TcTR-1504 TcTR Trypanosoma cruzi cDNA cl
41	tttagtgggcgatctctcttctctccctgtgtgtgtgttggcttta...	155	TcTR-1503 TcTR Trypanosoma cruzi cDNA cl



Banco de Dados

- ▶ Coleção auto-descritiva de registros integrados
- ▶ Inclui a definição de sua própria estrutura
 - ▶ Dados
 - ▶ Metadados
- ▶ Contém
 - ▶ dados (!)
 - ▶ registro das operações realizadas (log)
 - ▶ índices
 - ▶ dados sobre sua utilização
 - ▶ Usuários
 - ▶ Aplicativos
 - ▶ Estatísticas
 - ▶ ...

Dados

9713025	José Maria da Silva	487 2233	Arnaldo Vieira
9713038	Maria Cláudia Araujo	671 4245	Pedro Nascimento
9823007	Carlos Manoel Soares	265 3896	Maria Luiza Campos
9823158	Armando Gomes	287 6508	Maria Luiza Campos
9912250	Carla Ferrentini	274 3208	Arnaldo Vieira

Metadados

ESTUDANTE

Matrícula	Nome	Telefone	Orientador
9713025	José Maria da Silva	487 2233	Arnaldo Vieira
9713038	Maria Cláudia Araujo	671 4245	Pedro Nascimento
9823007	Carlos Manoel Soares	265 3896	Maria Luiza Campos
9823158	Armando Gomes	287 6508	Maria Luiza Campos
9912250	Carla Ferrentini	274 3208	Arnaldo Vieira

Metadados

- ▶ Dados sobre dados
- ▶ Permitem a manutenção, representação e identificação dos dados
- ▶ Garantem a consistência e a persistência das definições das unidades de dados
- ▶ Proporcionam aos usuários um entendimento comum das unidades de dados

Metadados

SysTables

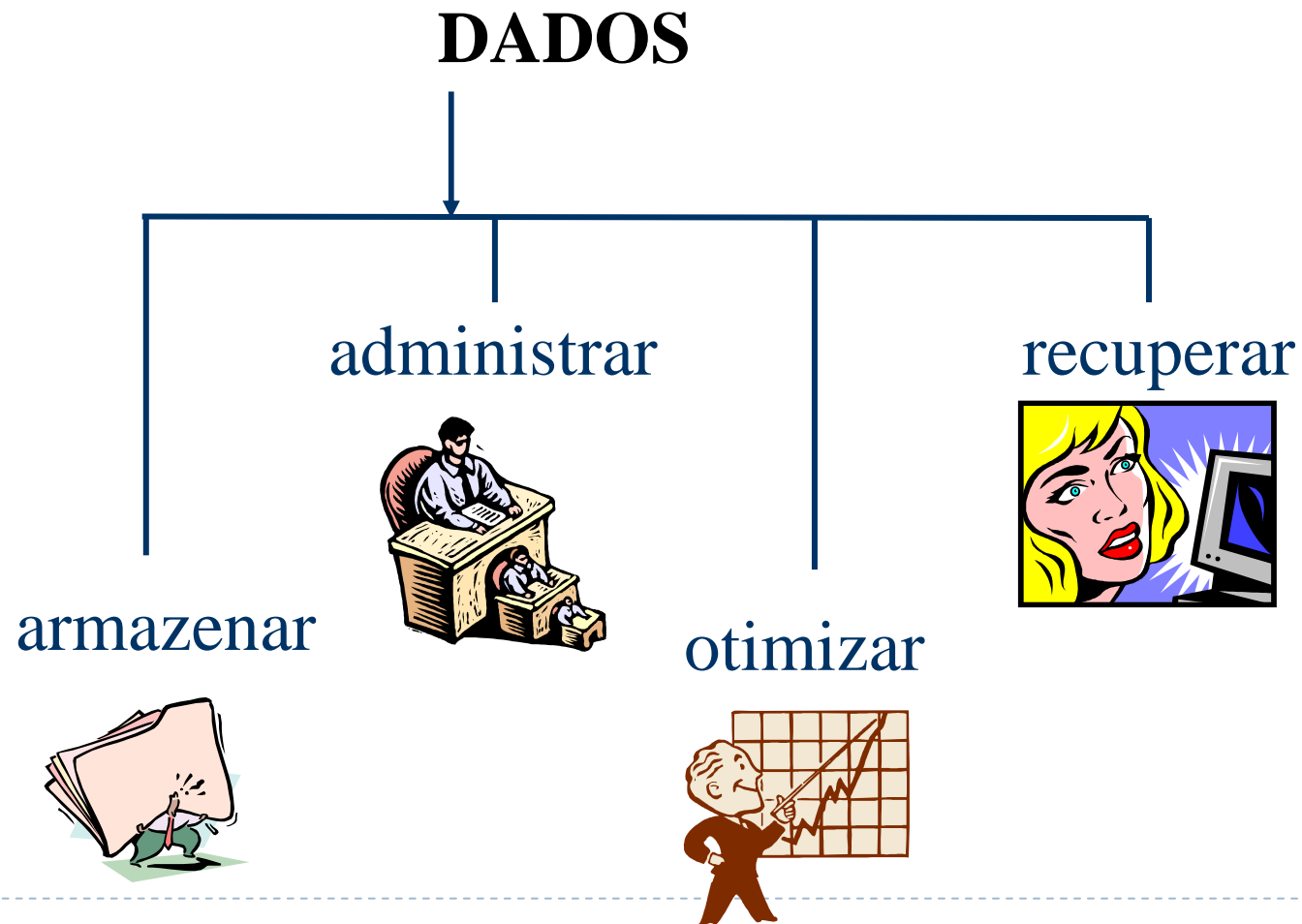
Table Name	Number of Columns	Primary Key
Estudante	4	Matrícula
Disciplina	5	Código
Matrícula	3	Nome + Código
Orientador	5	Identidade

SysColumns

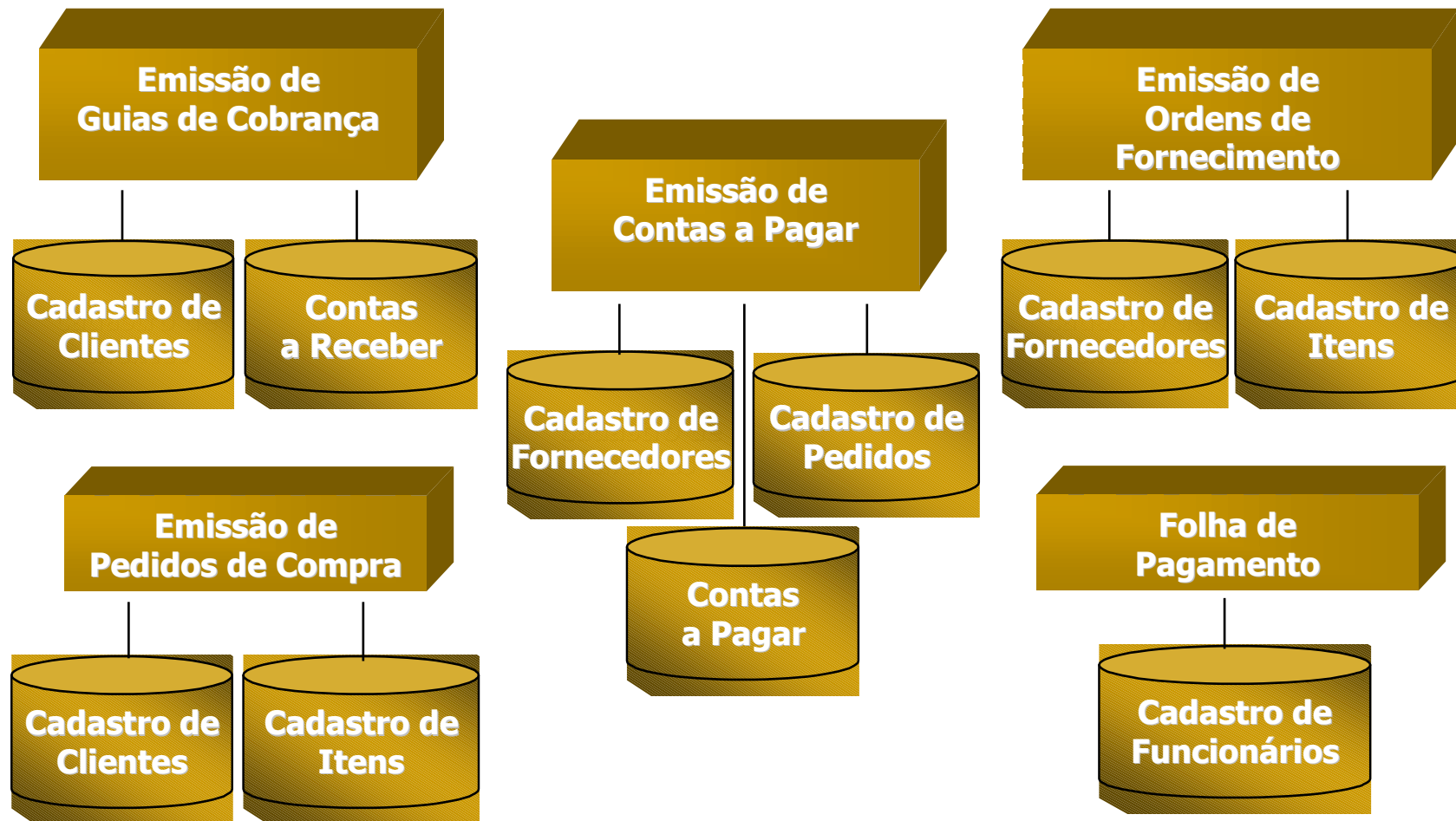
Column Name	Table Name	Data Type	Length
Matrícula	Estudante	Integer	4
Nome	Estudante	Text	30
Telefone	Estudante	Text	10
Orientador	Estudante	Text	30

Alguns autores/implementações de SGBD chamam de catálogo.

Gerência de dados



Sistema Tradicional de Arquivos



Limitações

- ▶ Dados isolados e em diferentes arquivos
- ▶ Registros incompatíveis
- ▶ Duplicação de dados
- ▶ Programas dependentes dos dados
- ▶ Dificuldade de representar o todo (a realidade do usuário)

Sistema Gerenciador de Banco de Dados



SGBD

- Sistema de software de finalidade genérica
 - facilita o processo de definição, construção e manipulação de bancos de dados para várias aplicações.
- Definir um BD envolve
 - especificar os tipos de dados, as estruturas e as restrições para os dados que serão armazenados no BD.
- Construir o BD
 - processo de armazenar os referidos dados em algum meio de armazenamento que seja controlado pelo SGBD.
- Manipular o BD
 - inclui funções como
 - fazer consultas ao BD para recuperar dados específicos,
 - atualizar o BD para refletir alterações no mini-mundo e
 - gerar relatórios a partir dos dados.



Características da abordagem de BDs

- ▶ **Natureza auto-contida**
 - ▶ metadados armazenados num catálogo (dicionário de dados)
- ▶ **Independência entre programas e dados**
 - ▶ estrutura dos arquivos de dados armazenada no catálogo separadamente dos programas de acesso
- ▶ **Abstração de dados**
 - ▶ representação lógica dos dados através de modelos de dados
- ▶ **Suporte a múltiplas visões de dados**
 - ▶ diferentes perspectivas (visões) de usuários
- ▶ **Suporte a transações**
 - ▶ compartilhamento de dados e processamento multi-usuário



Vantagens de Utilizar um SGBD

- ▶ Controle de Redundâncias
- ▶ Restrição do Acesso Não-Autorizado
- ▶ Armazenamento Persistente para Objetos de Programas e Estruturas de Dados
- ▶ Inferência e Ações Utilizando Regras
- ▶ Múltiplas Interfaces de Usuário
- ▶ Representação de Relacionamentos Complexos entre Dados
- ▶ Imposição de Restrições de Integridade
- ▶ Cópia e Recuperação (backup e recovery)



Funcionalidades de um SGBD

- ▶ Compartilhamento de dados
- ▶ Controle de redundância
- ▶ Controle de acesso
- ▶ Múltiplas interfaces
- ▶ Representação de relacionamentos
- ▶ Gerência das restrições de integridade
- ▶ Capacidade de “backup” e restauração



Implicações do Enfoque de BD

- ▶ Potencial para Impor Padrões
- ▶ Tempo Reduzido para Desenvolvimento de Aplicações
- ▶ Flexibilidade
- ▶ Disponibilidade de Informações Atualizadas
- ▶ Economias de Escala



Fatores Inibidores do Uso de SGBD

▶ Custos

- ▶ Investimentos iniciais elevados em hardware, software e treinamento
- ▶ A generalidade que um SGBD fornece para definir e processar dados
- ▶ Custos adicionais para fornecer segurança, controle de concorrência, recuperação e funções de integridade

▶ Dados estáticos

- ▶ BD e aplicações simples, bem definidos e não se espera que sejam alterados

▶ Tempo real

- ▶ Overhead do SGBD

▶ Acesso monousuário



Papéis em um Ambiente de BD

- ▶ **Administrador do SGBD**
 - ▶ Responsável pela instalação e manutenção do SGBD
 - ▶ Possui os mais altos privilégios
 - ▶ Detém profundos conhecimentos de SO
- ▶ **Administrador de Banco de Dados**
 - ▶ Responsável por criar e manter BDs
 - ▶ Possui altos privilégios nos BDs que administra
- ▶ **Administrador de Dados**
 - ▶ Responsável pelo ambiente de dados corporativo
- ▶ **Usuários**
 - ▶ Programadores e analistas, usuários finais

Administrador do SGBD

Funções

- ▶ **Instalação**
- ▶ **Autorização**
 - ▶ coordena com DBAs
- ▶ **Tuning**
 - ▶ configura servidores, sistemas de locking e logging
- ▶ **Manutenção**
 - ▶ põe/tira o servidor do ar
 - ▶ gerencia uso de disco
 - ▶ DBA do BD master
 - ▶ responsável pelos upgrades do sistema



Administrador de BD (ABD)

Funções

▶ Pré-produção

- ▶ responsável pelas estruturas de armazenamento e índices
- ▶ controle da transição entre fases do desenvolvimento
- ▶ cargas de dados entre ambientes

▶ Produção

- ▶ criar e destruir BDs
- ▶ criar objetos compartilhados no BD
- ▶ garantir segurança e integridade
- ▶ back-up e recovery
- ▶ monitoramento de desempenho
- ▶ previsão de necessidades futuras



Administrador de Dados (AD)

Funções

- ▶ modelagem de dados
- ▶ definição de políticas e responsabilidades sobre os recursos de dados
- ▶ estabelecimento de padrões
- ▶ integração de conceitos
- ▶ Documentação
- ▶ Responsável pela Arquitetura de Informações da organização



AD versus ABD

▶ Administração de Dados

- ▶ planejamento conceitual e organização dos recursos de dados
- ▶ Tarefas
 - ▶ modelagem de dados, desenvolvimento de políticas e padrões para uso e gerência dos dados, treinamento de usuários, coordenação de atividades com projetistas de sistemas e usuários

▶ Administração de Banco de Dados

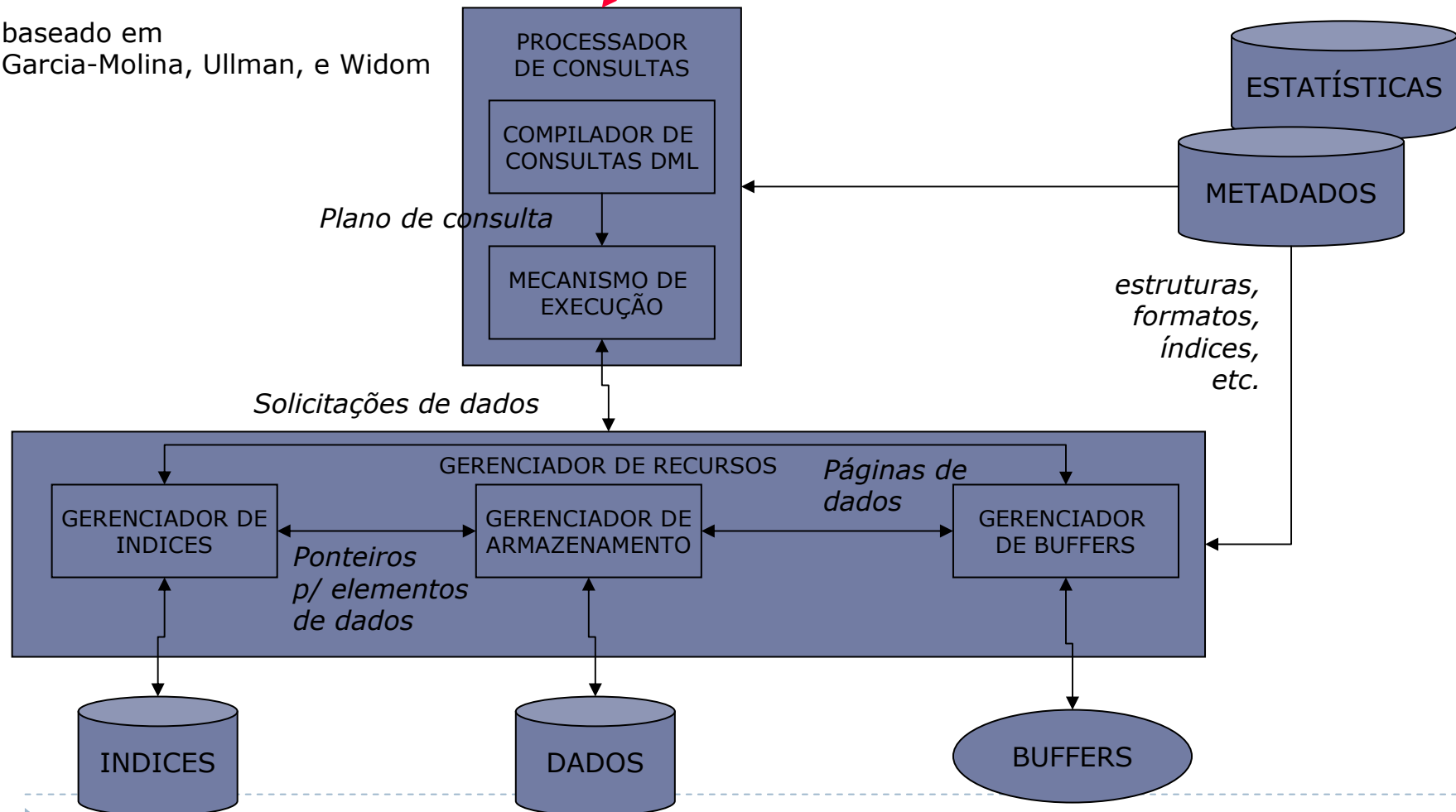
- ▶ voltada para aspectos técnicos associados a um ambiente específico de um sistema de gerenciamento de banco de dados
- ▶ operação e manutenção física de um ou mais BDs.



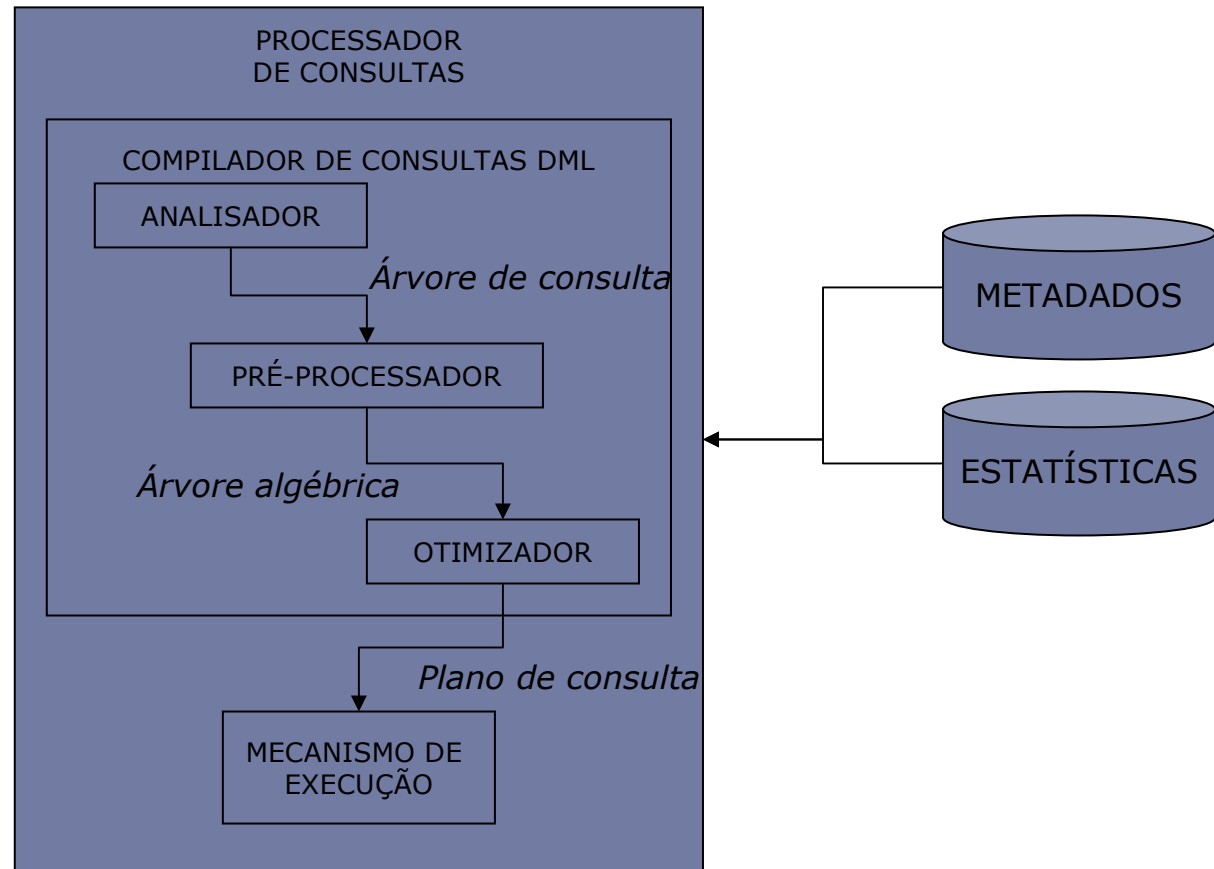
Componentes de um SGBD

Usuários

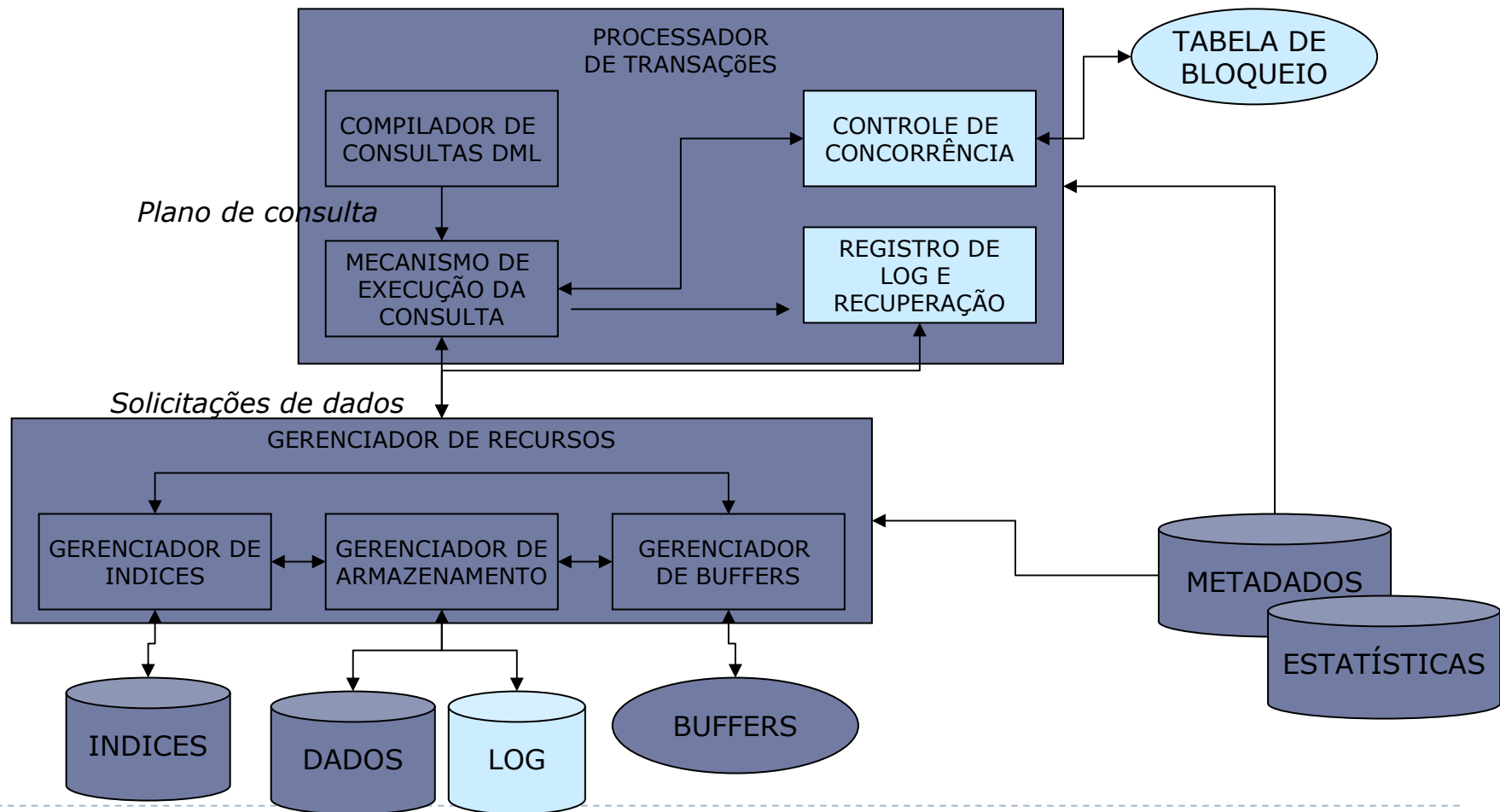
baseado em
Garcia-Molina, Ullman, e Widom



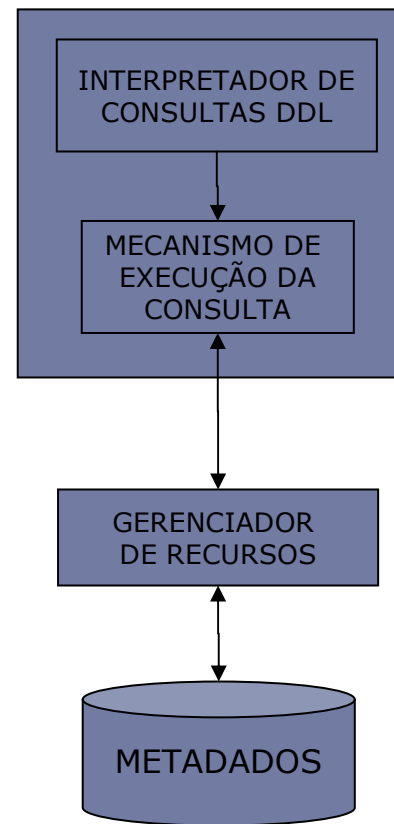
Componentes de um SGBD



Componentes de um SGBD



Componentes de um SGBD



Revisão de Conceitos

- ▶ Modelo de Dados
- ▶ Esquema de Dados
- ▶ Instâncias



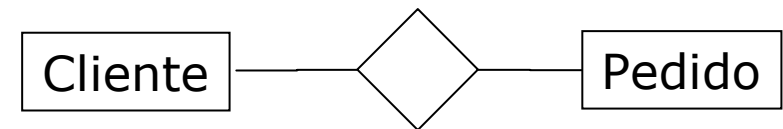
Modelo de Dados

- ▶ Utilizados para representar uma realidade
- ▶ Coleção de conceitos para descrever um conjunto de dados e as operações que os manipulam
- ▶ Conceitos definidos em termos verbais ou gráficos
 - ▶ Em geral, possui uma notação gráfica

Modelo de Dados - exemplos

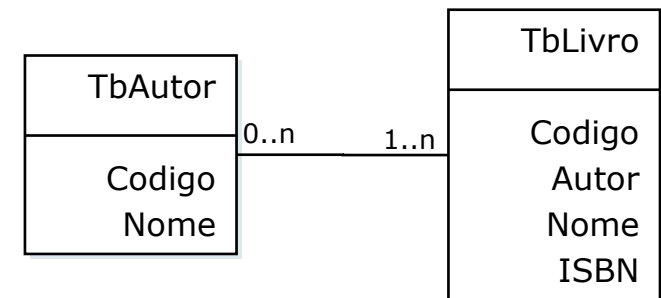
▶ Modelo entidade-relacionamento

- ▶ Entidades, relacionamentos



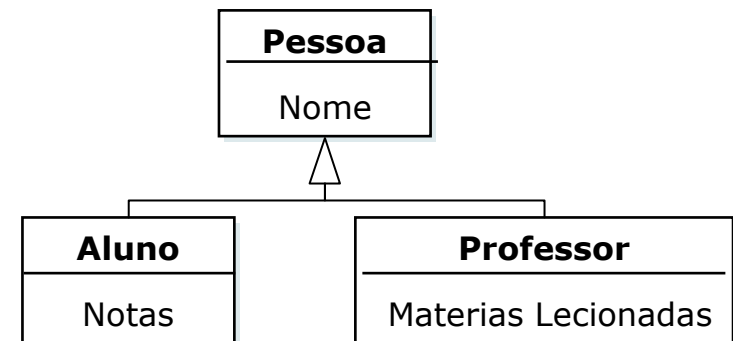
▶ Modelo relacional

- ▶ Tabelas, registros, campos



▶ Modelo de classes

- ▶ Classes, objetos, atributos



Diferentes níveis de abstração

▶ Modelo Conceitual

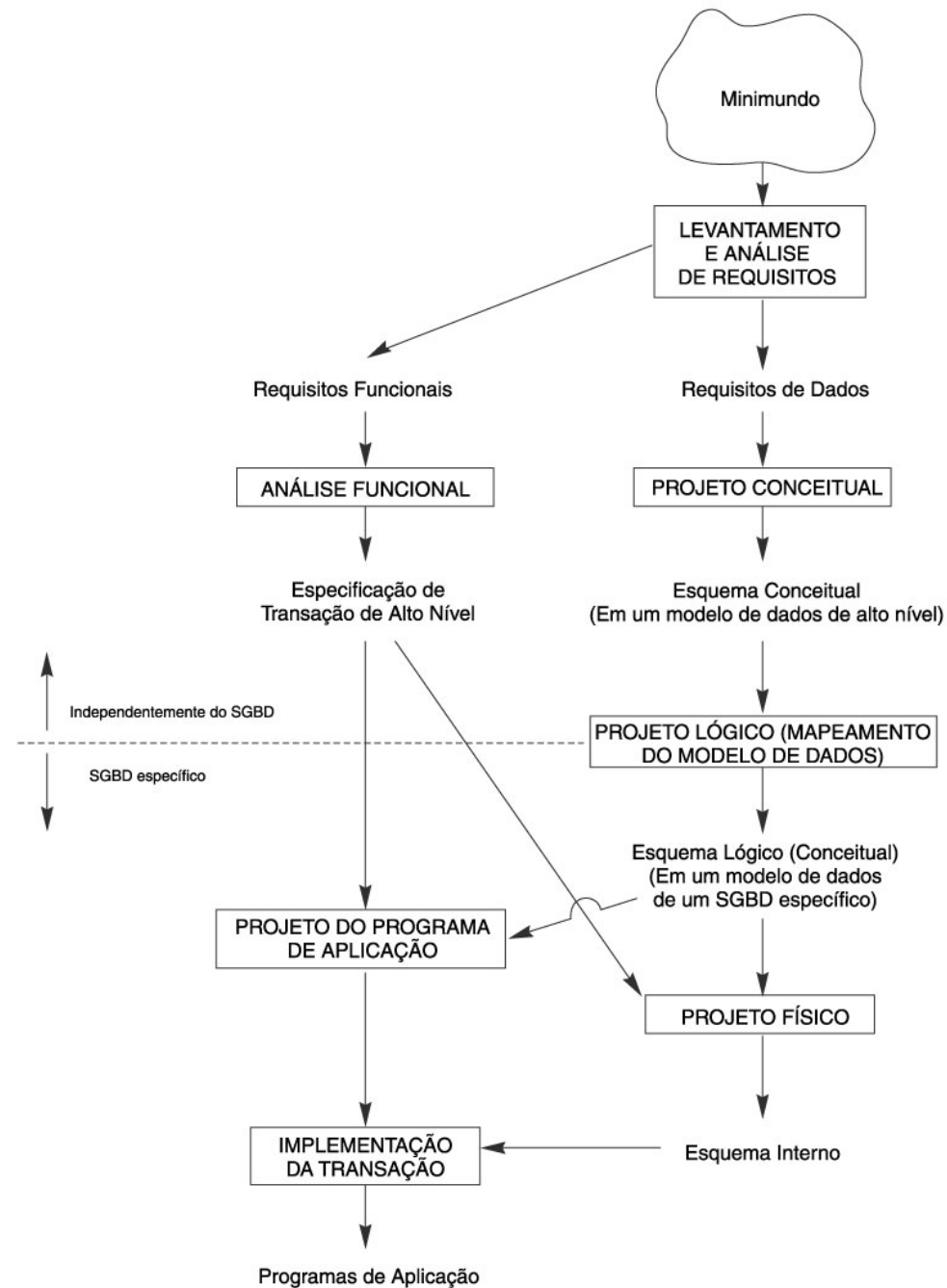
- ▶ representação (o mais fiel possível) da realidade
- ▶ alto-nível de abstração (mas não sem detalhamento...)

▶ Modelo Lógico

- ▶ descrição das estruturas lógicas dos dados
 - ▶ Esquema das tabelas
- ▶ forma passível de ser processada por computador

▶ Modelo Físico

- ▶ estruturas físicas de armazenamento
 - ▶ tabelas, partições do disco
- ▶ métodos de acesso aos dados
 - ▶ índices



Classificação de SGBDs

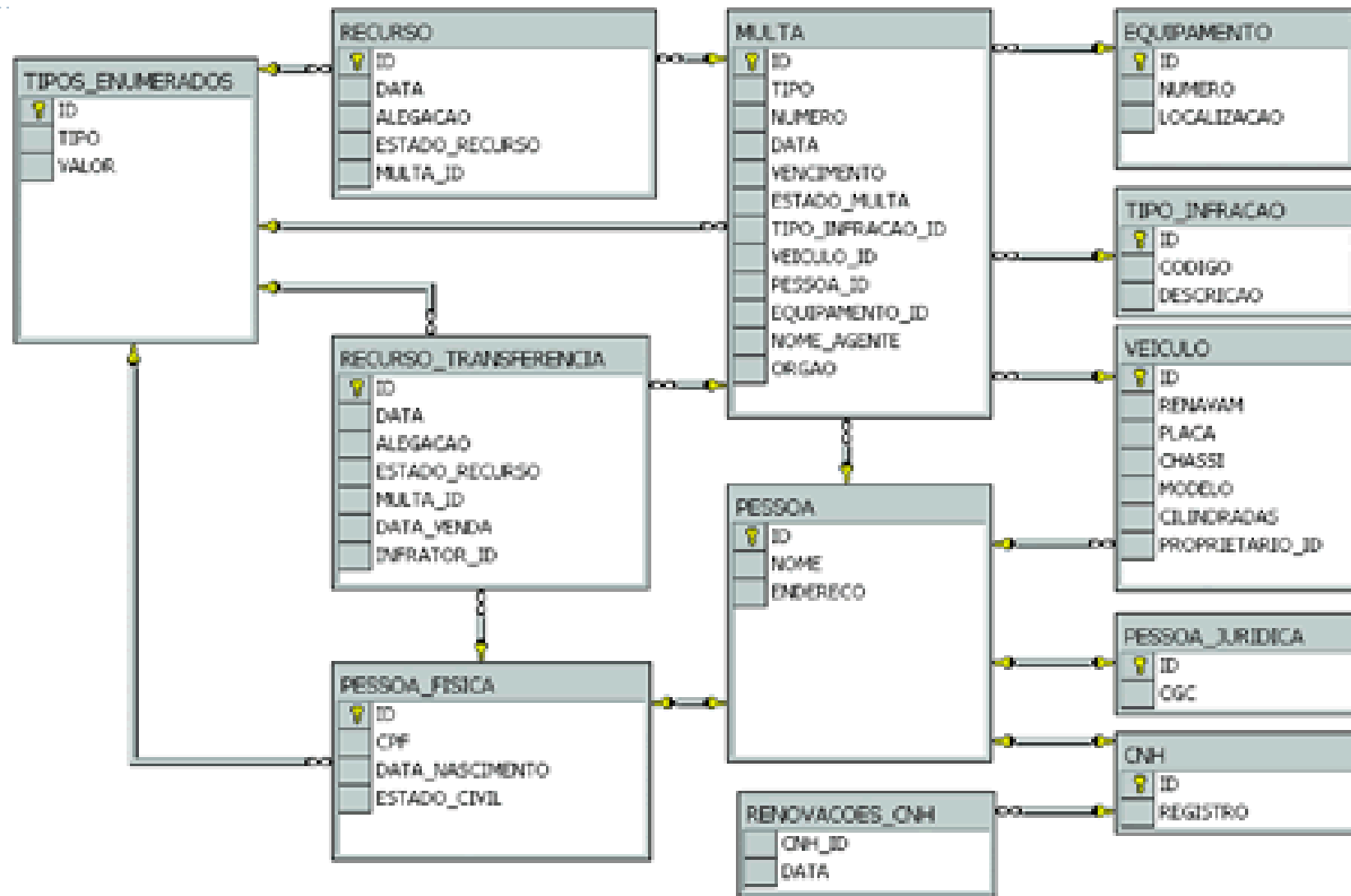
- ▶ Segundo o modelo (lógico e físico) adotado para estruturação e armazenamento dos dados
 - ▶ Hierárquico
 - ▶ Redes
 - ▶ Relacional (mais popular)
 - ▶ Relacional-objeto
 - ▶ Orientado a Objetos
 - ▶ XML



Esquema

- ▶ Representação de uma porção específica da realidade, a partir de um modelo de dados particular
 - ▶ “instância de um modelo”
- ▶ Coleção estática de estruturas sintáticas e/ou gráficas que descrevem um certo conjunto de dados
- ▶ Representa uma restrição às possíveis coleções de dados que podem descrever uma dada realidade

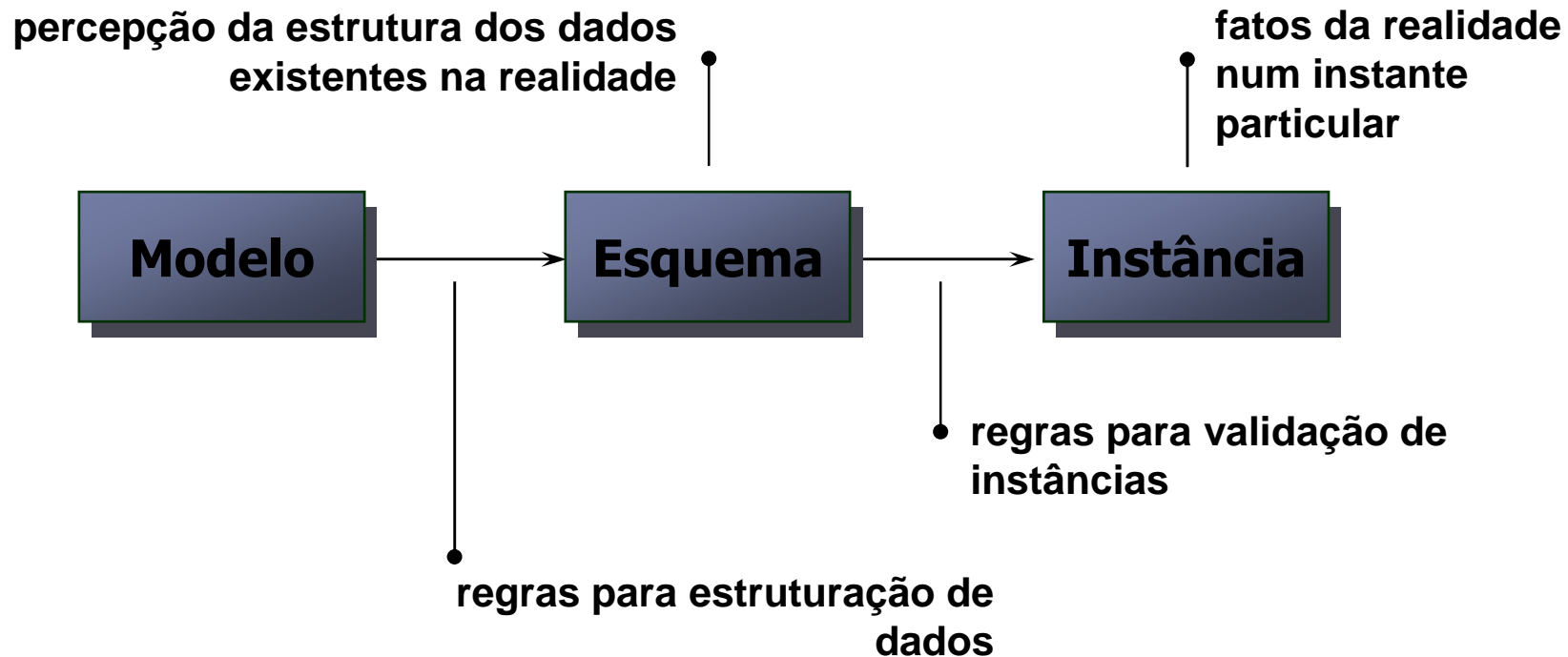
Esquema - exemplo



Instância de Esquema

- ▶ Coleção dinâmica de dados que se conforma à estrutura definida pelo esquema
- ▶ Representa o estado do banco de dados num instante particular

Modelo, Esquema e Instância

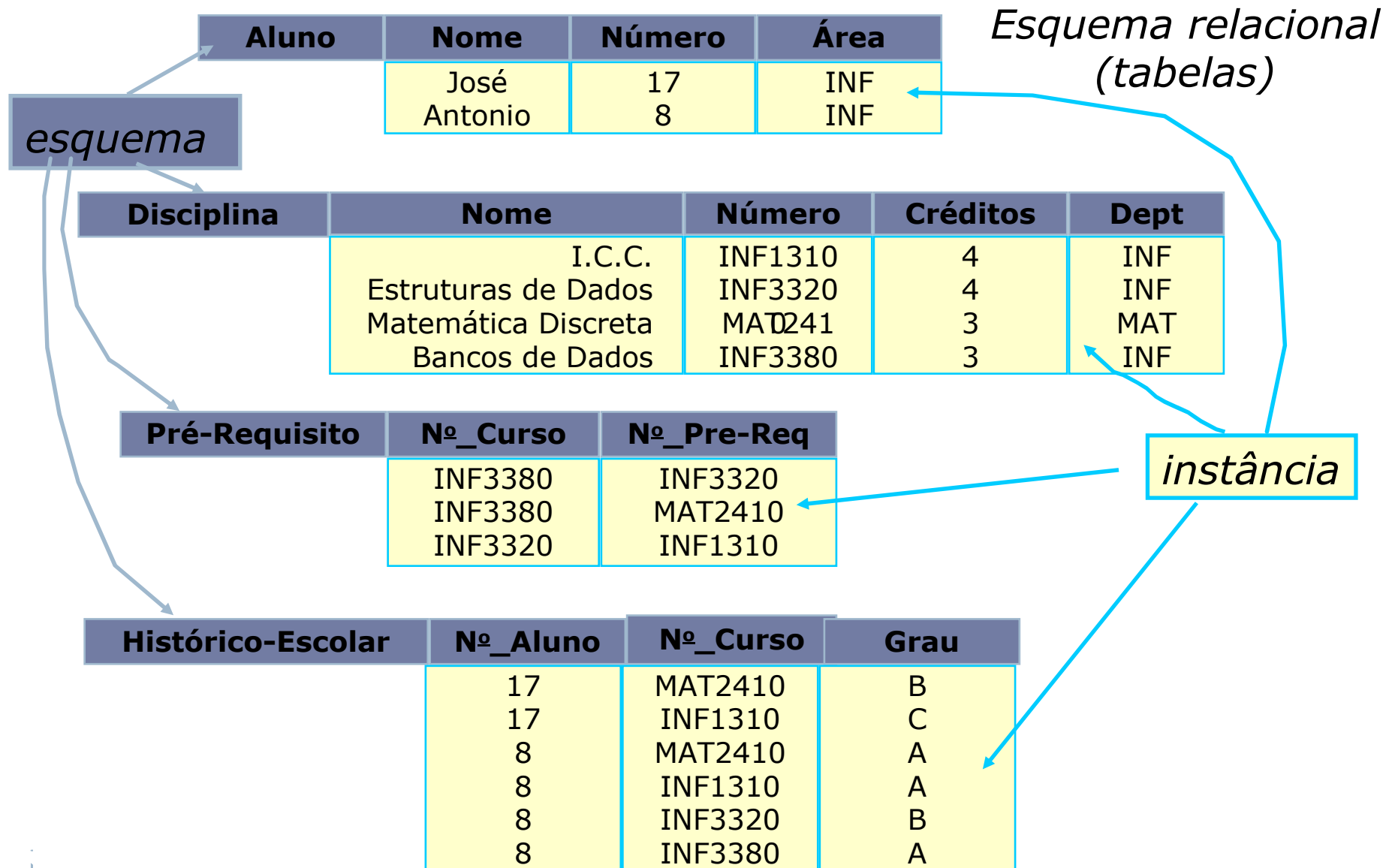


Exemplo

- ❖ Esquema
 - ❖ descreve uma realidade na qual existem pessoas e automóveis
- ❖ Registros: Pessoa e Automóvel
 - ❖ representam uma abstração sobre dois conjuntos de objetos de uma realidade
 - ❖ Descritos por características (atributos): Nome, sexo, endereço, ...
- ❖ É possível criar instâncias de pessoas e de automóveis
 - ❖ “pessoas que estão nesta sala”
 - ❖ “carros que têm autorização para estacionar na UFF”

CONCEITOS DO MODELO DE DADOS
<ul style="list-style-type: none">• Campo• Registro

ESQUEMA				
Pessoa	Nome	Sexo	Endereço	Identidade
Automóvel	Placa	Marca	Modelo	Cor



Arquitetura de 3 Esquemas (padrão ANSI)

