

Banco de Dados

Prof. Dr. Ronaldo Celso Messias Correia

ronaldo.correia@unesp.br

Objetivo e Ementa da Disciplina

➤ Objetivo:

- Introduzir os fundamentos que permitam ao aluno adquirir o domínio básico da tecnologia de banco de dados

➤ Ementa:

- Características das Abordagens no suporte ao armazenamento de dados
- Sistemas de Banco de Dados: Conceitos Básicos, Arquitetura e Classificação Modelos de Dados
- Modelagem Conceitual de Dados e Projeto Lógico de Banco de Dados
- Aspectos de implementação de Sistemas de Banco de Dados
- Sistemas de Banco de Dados não Convencionais
 - Banco de Dados Não Relacionais

Bibliografia

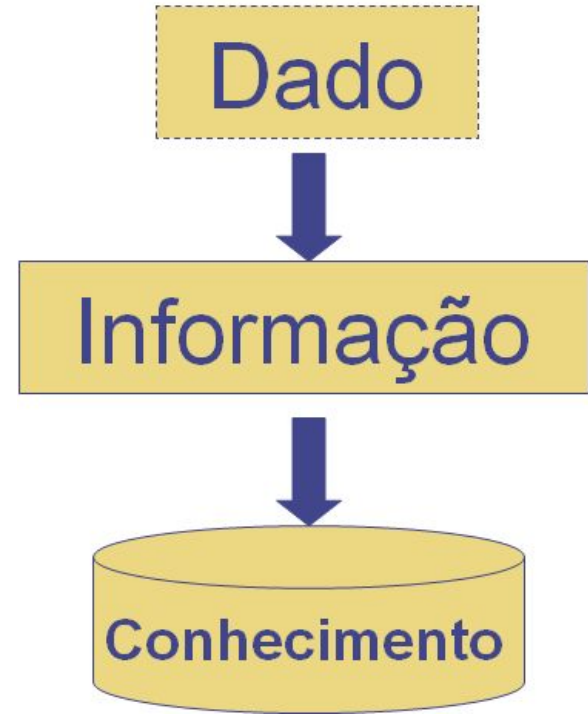
- SILBERSCHATZ, A., e KORTH, S. e SUDARSHAN, S. Sistema de Banco de Dados, 6a edição, Editora Campus, 2012.
- NAVATHE, S. e ELMASRI, R. Sistemas de Banco de Dados, 6a edição, Pearson, 2011.
- DATE, C. Introdução a Sistemas de Banco de Dados, Campus, 2003.
- Artigos Científicos

Avaliações - Critérios

- As notas de todas as atividades irão variar de 0 (zero) a 10,0 (dez)
- As notas das atividades em grupo serão atribuídas individualmente
- A média final será calculada da seguinte maneira:
$$MP = (Pb1 + Pb2)/2$$
$$MF = (MP * 0,8) + (MT * 0,2)$$
- Onde:
 - MF = Média Final
 - MP = Média das Provas Bimestral
 - Pb = Notas Provas Bimestrais
 - MT = Média das notas das atividades práticas do bimestre
- Caso o aluno não obtenha a nota mínima para aprovação, será oferecida uma terceira avaliação (EXAME) e a Média Final passa a ser obtida pela expressão:
$$MFnova = (MFanterior + EXAME)/2$$

Conceitos: Dados, Informação e Conhecimento

- **Dados** : é a representação da informação, que pode estar registrado em papel, num quadro de aviso ou no meio digital
- Dados são códigos que constituem a matéria prima da informação, ou seja, é a informação não tratada. Conjunto de letras, números ou dígitos que colocado isoladamente, não contém significado claro
- Pode estar registrado em papel, num quadro de aviso ou no disco rígido do computador
- O computador armazena e processa dados



Conceitos: Dados, Informação e Conhecimento

- **Informação** : é qualquer fato ou conhecimento do mundo real e que pode ou não ser registrado /armazenado
- O conceito de informação vem ser o dado trabalhado ou tratado agregado com sentido natural e lógico para quem usa a informação
- As informações têm significado, podem ser tomadas decisões ou fazer afirmações considerando as informações.
- Exemplo:
 - Informação: Está muito quente hoje
 - Dado: A temperatura hoje é de 38 graus Celsius

Conceitos: Dados, Informação e Conhecimento

- **Conhecimento** : Quando a informação é “trabalhada” por pessoas e pelos recursos computacionais, possibilitando geração de cenários, simulações e oportunidades, pode ser chamada de conhecimento
- Produz ideias e experiências que as informações por si só não será capaz de mostrar
- Exemplo:
 - Percepção da dificuldade
 - Uso de experiências semelhantes
 - Concepção de equipamentos, pessoas, materiais e pessoas, que são vitais para um serviço
 - Entendimento de contratos que podem ser negociados, visando à adequação à realidade de uma atividade

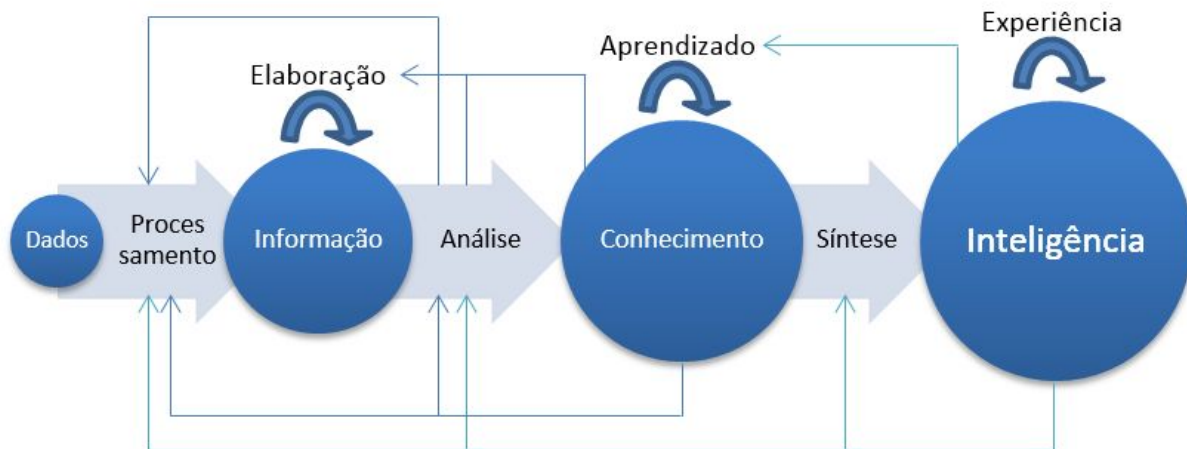
Conceitos: Dados, Informação e Conhecimento

- O Conhecimento pode ser classificado em:
- **Tácito** - aquele que é pessoal, que está internalizado e que é difícil de ser formulado e comunicado (Percepções, ideias, experiências, *Know-how*)
- **Explícito** - aquele que já está codificado e, portanto, passível de ser transmitido (melhores práticas, políticas, documentos).



Conceitos: Dados, Informação e Conhecimento

Dados	Informação	Conhecimento
Simple observações sobre o estado do mundo	Dados dotados de relevância e propósito	Informação valiosa da mente humana. Inclui reflexão, síntese, contexto
<ul style="list-style-type: none">▪ Facilmente estruturado▪ Facilmente obtido por máquinas▪ Frequentemente quantificado▪ Facilmente transferível	<ul style="list-style-type: none">▪ Requer unidade de análise▪ Exige consenso em relação ao significado▪ Exige necessariamente a mediação humana	<ul style="list-style-type: none">▪ De difícil estruturação▪ De difícil captura em máquinas▪ Frequentemente tácito▪ De difícil transferência



Conceitos: Banco de Dados

- É uma coleção de dados relacionados
- O uso do termo é mais restrito em virtude das seguintes características:
 - Um BD representa algum aspecto do mundo real, o qual chamamos de Minimundo ou Universo de Discurso
 - É um conjunto lógico e ordenado de dados que possuem algum significado inerente
 - Um BD é projetado, construído e povoado com dados que possuem objetivos específicos
- Ingredientes necessários em um BD:
 - Uma fonte de dados da qual derivamos os dados
 - A interação com o mundo real
 - Público que demonstra interesse nos dados contidos no Banco

Conceitos: Sistema Gerenciador de Banco de Dados (SGBD)

- É uma coleção de programas que permite aos usuários criar e manter um banco de dados. Facilita os processos de definição, construção, manipulação e compartilhamento de bancos de dados entre vários usuários e aplicações
 - **Definição** : Especificação dos tipos de dados, das estruturas das tabelas e das restrições que devem ser impostas aos dados que serão armazenados
 - **Construção** : Processo de acumular os dados num meio de armazenamento controlado pelo SGBD
 - **Manipulação** : Operações como atualização do banco de dados (inclusão, exclusão e alteração de registros) e extração de dados, como consultas e relatórios impressos
 - **Compartilhamento** : Permite aos múltiplos usuários e programas acessar, de forma concorrente, o banco de dados

➤ Exemplos:



Comparação com o Processamento Tradicional

- No processamento tradicional de arquivos os usuários definem os arquivos necessários para cada aplicação específica, resultando em redundância e desperdício de espaço de armazenamento
- No enfoque de BD não é armazenado somente o banco em si, mas sim a estrutura do banco de dados e uma descrição completa (Catálogo do Sistema)
- Acesso não requer conhecimento destas estruturas (Independência dos Dados)
 - Quando houver alteração na estrutura de dados os programas não precisam ser alterados
 - As informações do catálogo são chamadas de **Metadados**

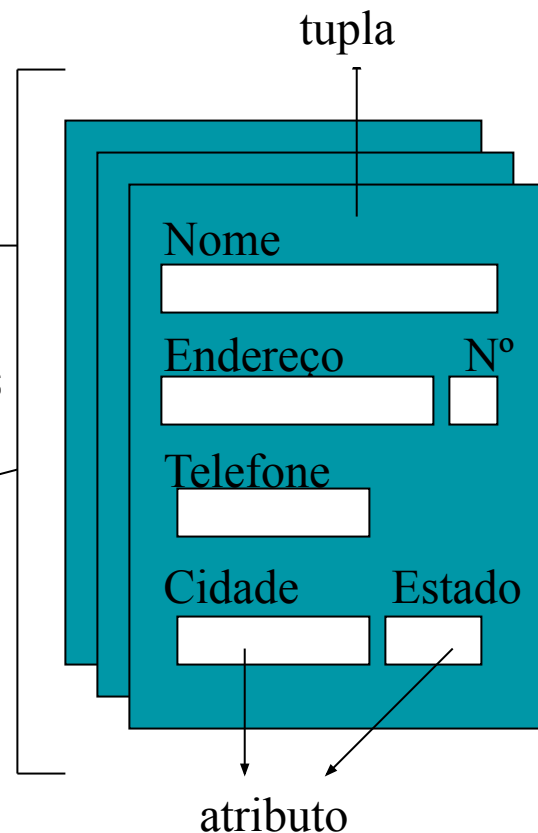
Conceitos: Terminologia

➤ Terminologia Básica

- Campo: unidade básica de informação mínima com significado
- Registro: conjunto de campos
- Arquivo: conjunto de registros
- Banco de Dados (BD): conjunto de arquivos e as formas de manipulação

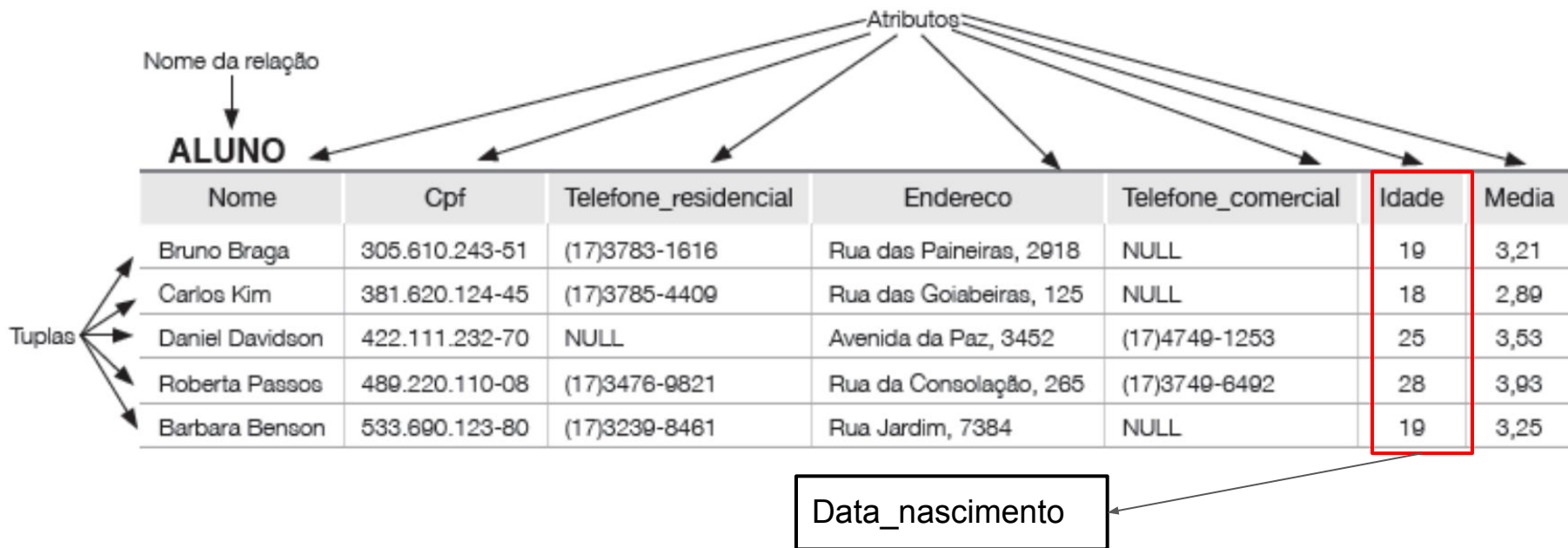
➤ Terminologia Relacional

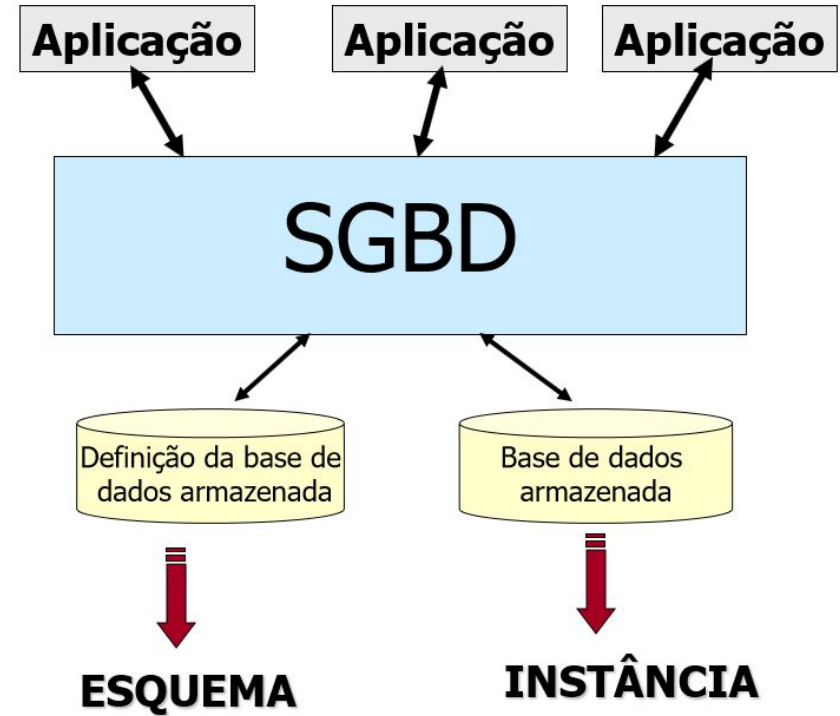
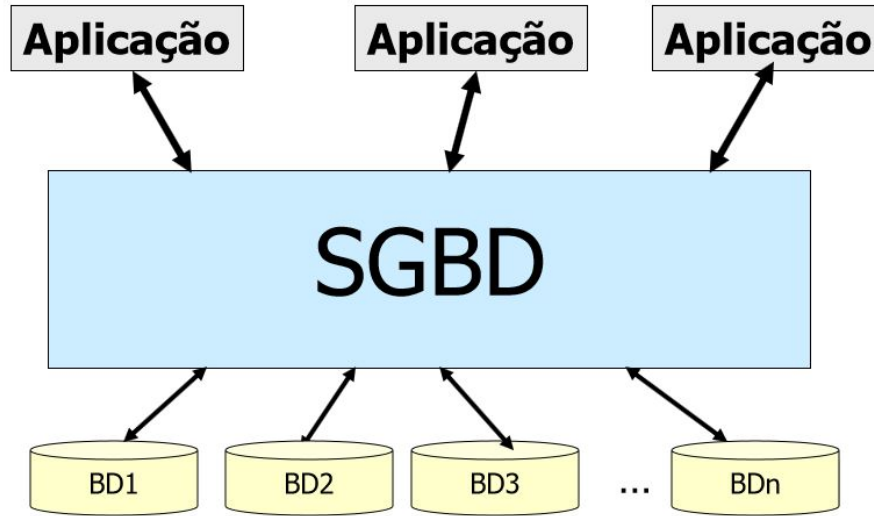
- Campo: atributo
- Registro: tupla
- Arquivo: tabela ou relação
- Domínio: conteúdo da relação



Conceitos: Terminologia

- Exemplo de uma relação (tabela):





Vantagens de Utilizar um SGBD

- Independência entre programas e dados
 - Catálogo que consiste de metadados – dados sobre os dados
- Independência entre operações e programas
 - Funções/procedimentos de manipulação dos dados armazenados também fazem parte do BD
- Segurança: Controle de acesso mais especializado
- Acesso compartilhado aos dados
- Suporte a Visões
 - Mesmo conjunto de dados pode ser apresentado a usuários diferentes de forma distinta
- Facilidades de Backup e Restauração
- Fornecimento de Múltiplas Interfaces aos Usuários
 - Baseadas em Menus e formulários (GUIs – Graphical USer Interface)

Linguagens de consulta e interfaces de linguagem de programação

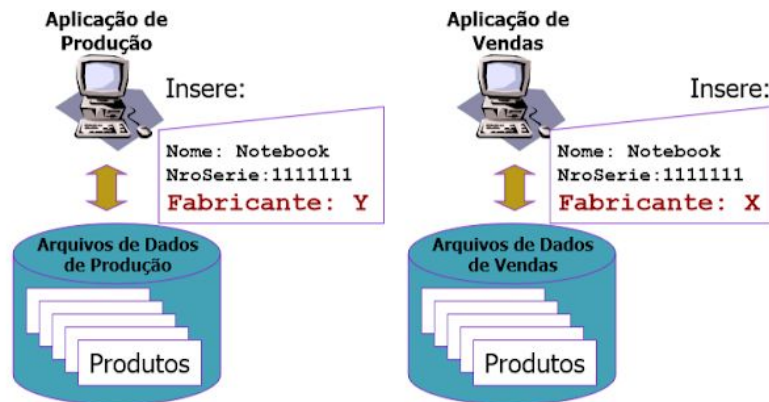
Profissionais envolvidos

- DBA (Administrador de Banco de Dados)
 - Responsável por autorizar o acesso, coordenar e monitorar seu uso, e por adquirir recursos de software e hardware
 - Avaliar o desempenho e problemas de segurança do sistema
- Projetista do BD:
 - Responsável por escolher as estruturas para representar e armazenar dados e visões
- Usuário Final: podem integrar com o BD através de:
 - Aplicações de BD: usuários leigos
 - Linguagens de BD: usuários casuais a sofisticados
- Analistas de Sistemas e programadores:
 - Determinam as necessidades dos usuários finais e implementam as aplicações
- Projetistas do SGBD
 - Projetam e implementam os módulos e interfaces do SGBD, como um pacote
- Desenvolvedores de Ferramentas:
 - Programas que facilitam o projeto e uso de um sistema de banco de dados, aprimorando seu desempenho

Consistência de Dados

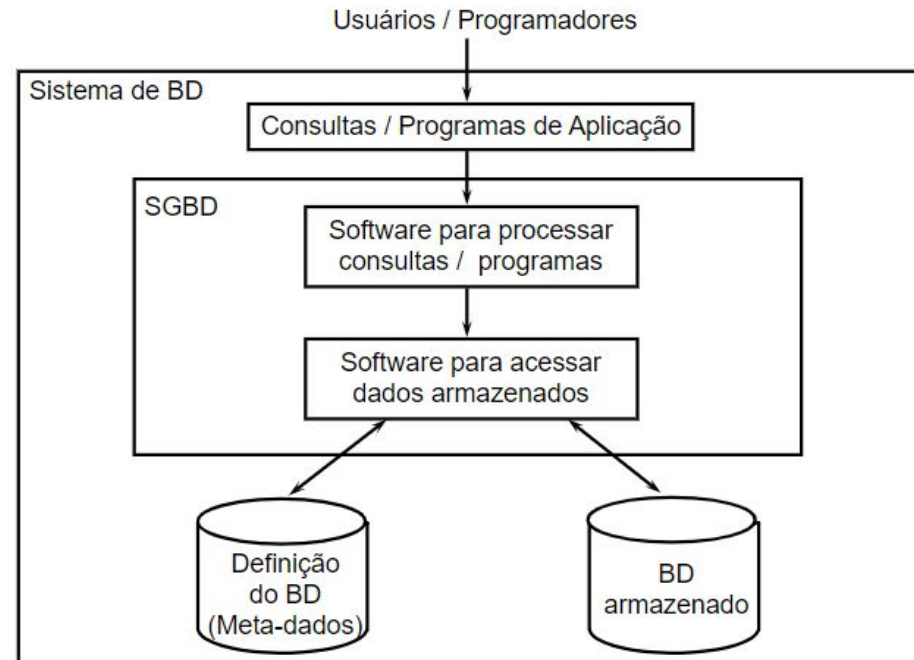
- É o estado ou caráter do que é coerente, do que tem solidez, veracidade, credibilidade, estabilidade, realidade
- Consistência: sempre que a mesma informação é armazenada, mesmo que em locais diferentes (redundância), ela tem o mesmo valor
- Quando os dados se encontram em um estado inconsistente, informações incorretas ou contraditórias podem ser fornecidas aos usuários

REDUNDÂNCIA ➡ **INCONSISTÊNCIA**



Sistema de Banco de Dados

- É um ambiente de hardware e de software composto por dados armazenados em um banco de dados (BD), pelo software de gerência do banco de dados (SGBD) e os programas de aplicação



Abstração de Dados / Modelo de Dados

- Abstração de Dados é a supressão de detalhes da organização e armazenamento dos dados, destacando recursos essenciais para um melhor conhecimento desses dados
- Modelo de Dados é uma coletânea de conceitos que podem ser utilizados para descrever a estrutura de um banco de dados, oferecendo os meios necessários para alcançar a abstração
- Modelo de Dados - Categorias:
 - Modelos de dados conceituais (alto nível)
 - Possuem conceitos que descrevem os dados como os **usuários os percebem** : entidades, atributos e relacionamentos
 - Entidade-Relacionamento; Orientado a objetos
 - Modelos de dados representacionais (ou de implementação)
 - Descrevem a estrutura de um BD da forma que será manipulado pelo SGBD
 - Modelo Relacional; Modelo de rede; Modelo Hierárquico
 - Modelos de dados físicos (baixo nível)
 - Descrevem detalhes de como os dados estão armazenados no computador

Sistema de Banco de Dados

- Em qualquer modelo de dados é importante distinguir entre a descrição do banco de dados e o banco de dados de fato
- **Esquema**
 - É a descrição do banco de dados
 - É definido durante o projeto do banco de dados e não se espera que seja alterado frequentemente
 - Em linguagem de programação é equivalente a definição de um tipo de dados
- **Instância**
 - É o banco de dados em si
 - Em uma linguagem de programação, isto é equivalente a uma declaração de uma variável do tipo definido e o seu valor

Arquitetura de Três Esquemas (ou níveis)

➤ Proposta para auxiliar na realização e visualização das seguintes características:

- Independência de dados e operação de programas
- Suporte a múltiplas visões
- Uso do catálogo para armazenar a descrição do BD

➤ O objetivo é separar o usuário da aplicação do BD físico

■ 1. Nível Interno – esquema interno

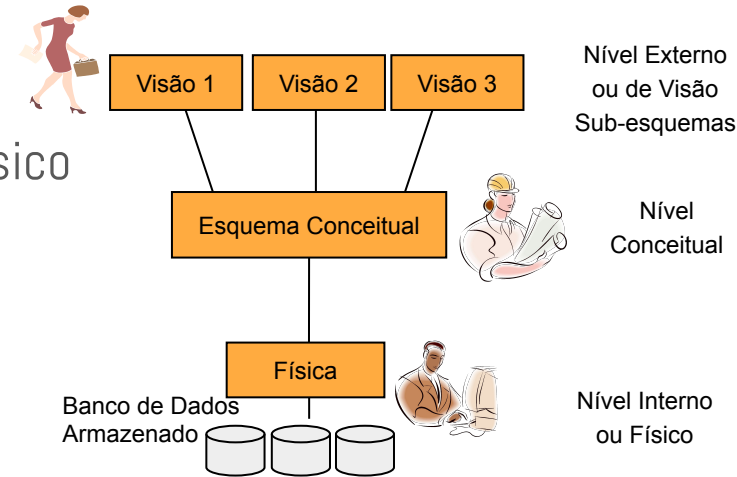
- Descreve a estrutura física de armazenamento do banco de dados
- Utiliza um modelo de dados físico

■ 2. Nível Conceitual – esquema conceitual ou lógico

- Descreve a estrutura da base de dados sem detalhes de estrutura de armazenamento físico
- Que dados estão armazenados e como estão relacionados

■ 3. Nível Externo – esquema externo (visões dos usuários)

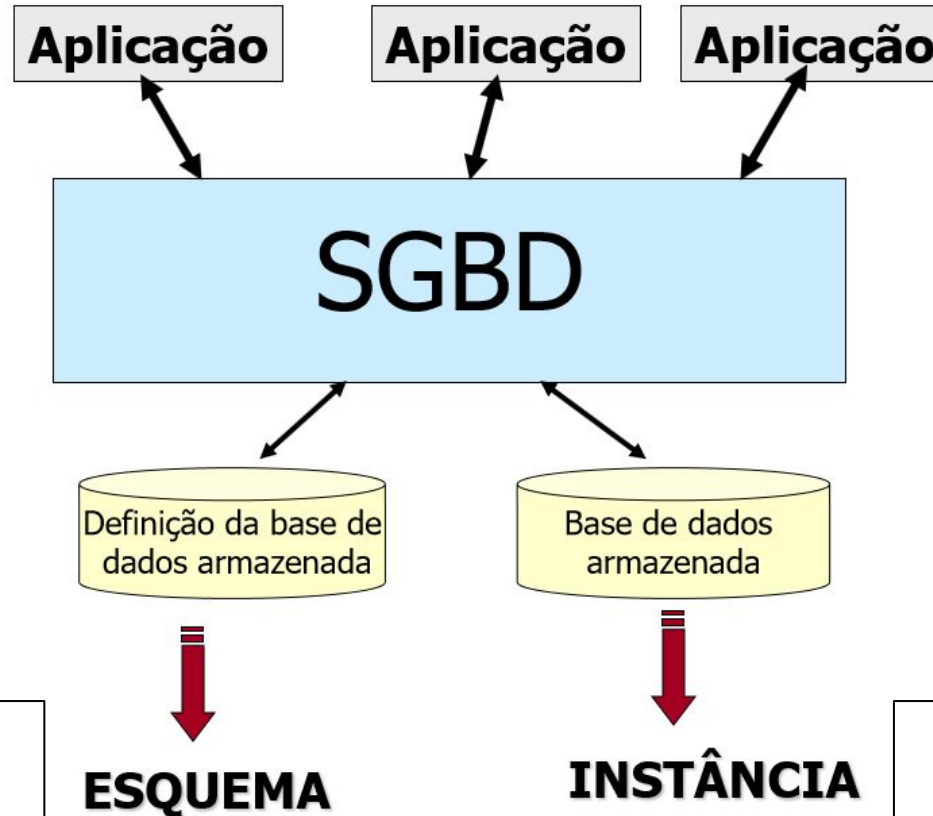
- Descreve as visões dos usuários: a parte da base de dados em que cada grupo de usuários tem interesse
- Descrição de sub-esquemas (Modelo conceitual: MER; Modelo Implementação: Relacional)



Arquitetura de Três Esquemas (ou níveis)

- É a capacidade de mudar o esquema em um nível do sistema de banco de dados sem que ocorram alterações do esquema no próximo nível mais alto
- Independência de dados lógica
 - Refere-se a capacidade de modificar o esquema lógico sem que, com isso, qualquer programa de aplicação precise ser reescrito
- Independência de dados física
 - Refere-se a capacidade de modificar o esquema físico sem que, com isso, qualquer programa de aplicação precise ser reescrito
- O conceito de independência de dados é de várias formas similar ao conceito de tipo abstrato de dados empregado nas linguagens de programação

Sistema de Banco de Dados



- Definição
- Estático (ou quase!)

- Manipulação
- Dinâmica



Sistema de Banco de Dados



Sistema de Banco de Dados

Desenvolvimento de Sistemas de Banco de Dados

Ciclo de Vida



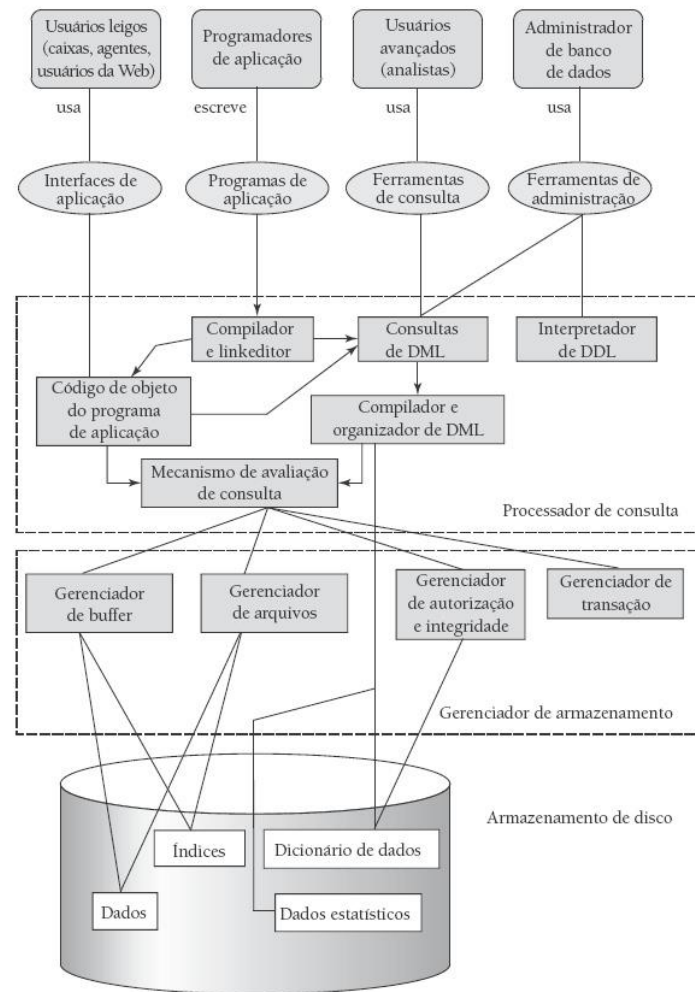
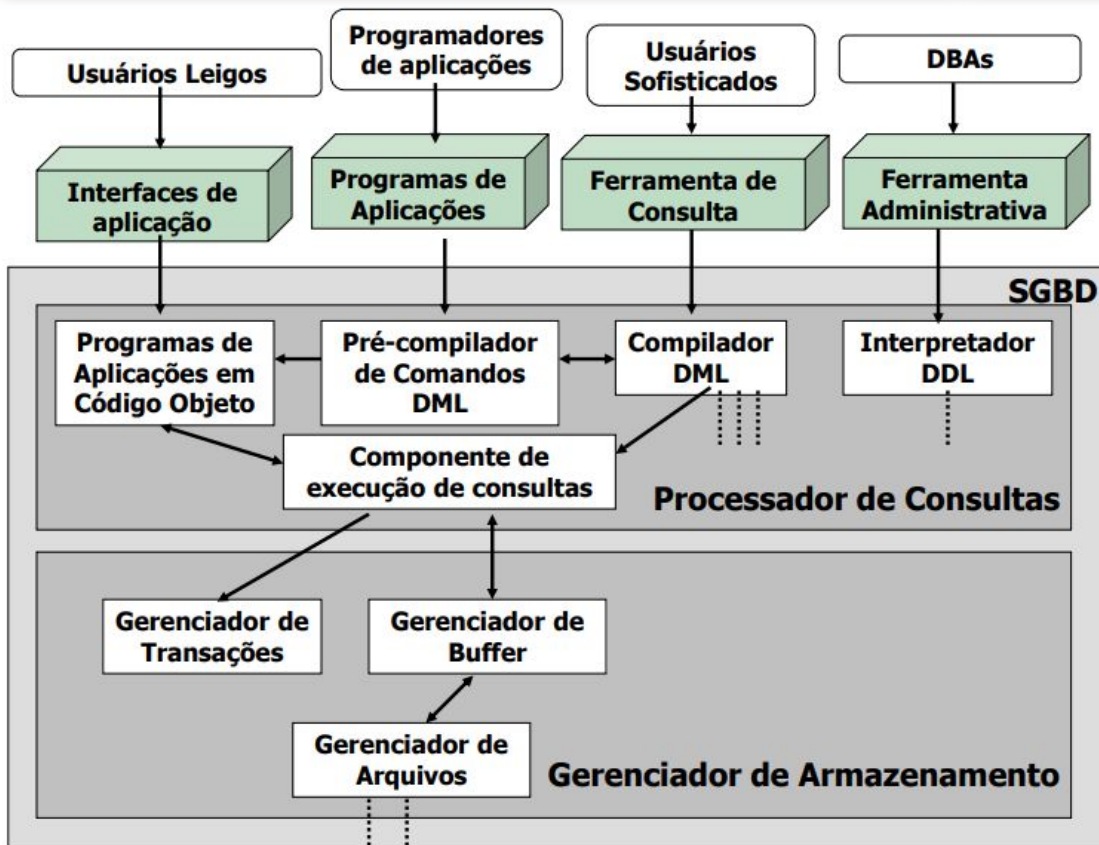
Sistema de Banco de Dados

- Projeto Conceitual
 - Independente de SGBD
 - Modelo Conceitual – MER
- Projeto Lógico
 - Esquema Lógico - Mapeamento do Modelo Conceitual para modelo do SGBD
 - Ex: Modelo Relacional
- Projeto Físico
 - Estruturas Físicas de Armazenamento
 - Organização de registros físicos - Índices
 - Critérios
 - Tempo de resposta; Espaço utilizado; Número de transações

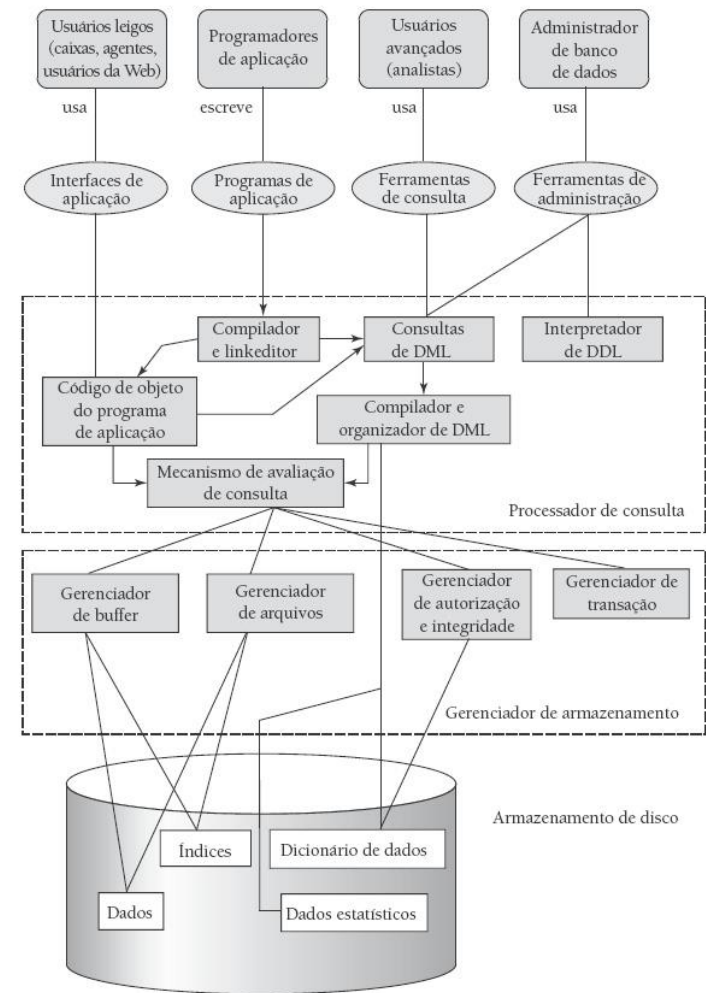
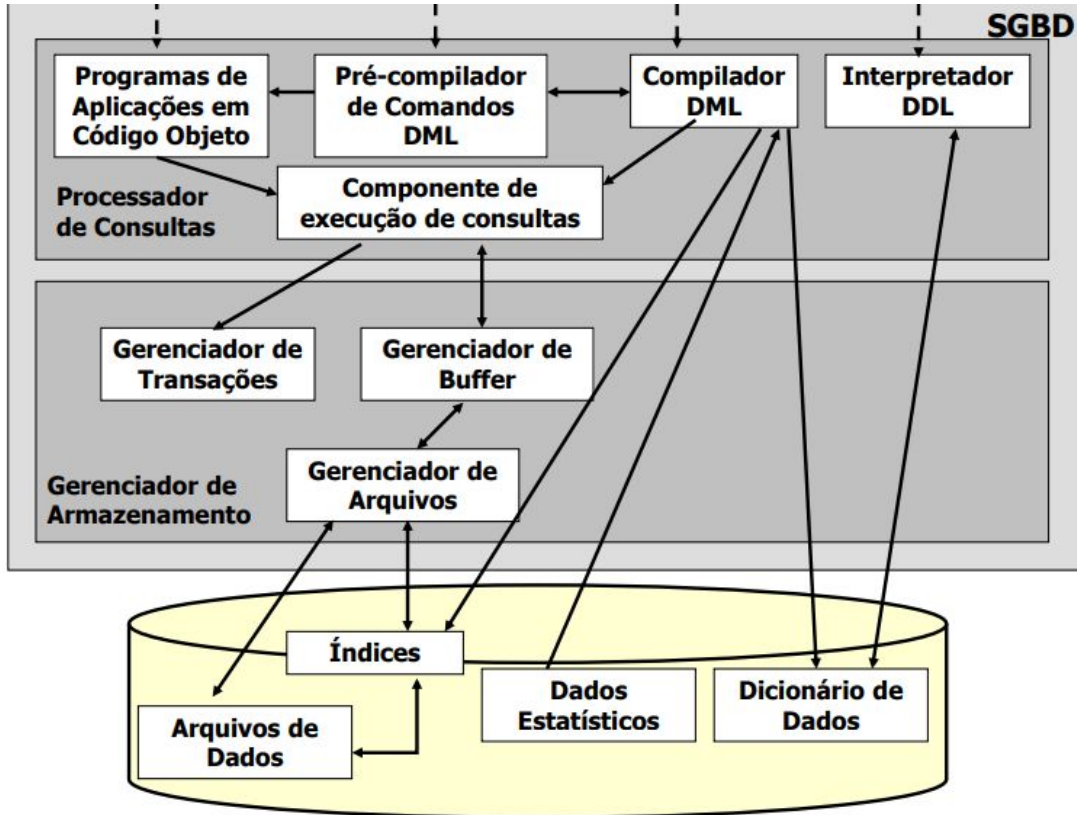
Linguagens de Banco de Dados

- **Linguagem de Definição de Dados (DDL)** : Permite especificar o esquema do banco de dados, através de um conjunto de definições de dados
 - O compilador DDL do SGBD processa e identifica os comandos, que são armazenados no catálogo (metadados)
- **Linguagem de Manipulação de Dados (DML)** : Permite ao usuário acessar ou manipular os dados.
 - Uma consulta ("query") é um comando que requisita uma recuperação de informação
 - A parte de uma DML que envolve recuperação de informação é chamada linguagem consulta
 - DMLs procedurais: exigem que o usuário especifique quais dados são necessários, e como obtê-los: Ex: Álgebra Relacional e SQL
 - DMLs não procedurais: exigem que o usuário especifique quais dados são necessários, sem especificar como obtê-los: Ex: Cálculo Relacional

Componentes de um SGDB



Componentes de um SGDB



Componentes: Processador de Consultas



Compilador DML

- Analisa sintaticamente e semanticamente comandos DML expressos em uma linguagem de consulta (ex. SQL)
- Traduz estes comandos para uma das formas de representação interna de consultas (ex. álgebra relacional)



Pré-Compilador DML

- Inseridos em programas de aplicação, traduz comandos DML em chamadas a procedimentos (rotinas) na linguagem hospedeira



Interpretador DDL

- Interpreta comandos DDL e os armazena no catálogo
- Tabelas contendo meta-dados
- Descrição do banco de dados – Esquema



Mecanismo de Consultas

- Responsável pela otimização e geração de planos de execução de consultas
- Executam instruções geradas pelo compilador DML

Componentes: Sistema de Armazenamento

➤ Gerenciador de transações

- Controle de concorrência
- Recuperação do banco de dados após falha (estado consistente)

➤ Gerenciador de arquivos (File System)

- Responsável pelo armazenamento físico em disco
- Gerencia a alocação de espaço em disco

➤ Gerenciador de buffer

- Responsável para recuperar objetos em disco e carregá-los na memória principal em forma de páginas
- SGBD possui uma área de buffer em memória principal
- Mapeamento: **Bloco** **Página**
(disco) (buffer do SGBD)

Componentes: Arquivos de dados + Índices + Catálogo

➤ Arquivo de dados

- Armazena o próprio banco de dados

➤ Índices

- Estruturas de índices para os arquivos de dados
- Proporcionam acesso rápido aos itens de dados

➤ Catálogo: Armazena esquema do banco de dados (meta-dados)

- Nomes das tabelas
- Atributos de cada tabela
- Definição de índice para uma tabela, etc...

➤ Dados Estatísticos

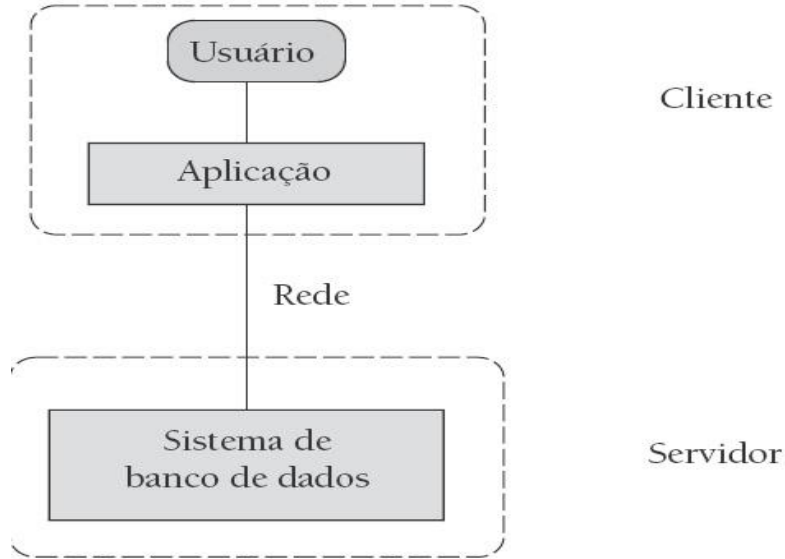
- Informações utilizadas pelo processador de consultas para seleção de meios eficientes para execução de uma consulta
- Exemplo: Cardinalidade de uma tabela

Arquiteturas do Banco de Dados

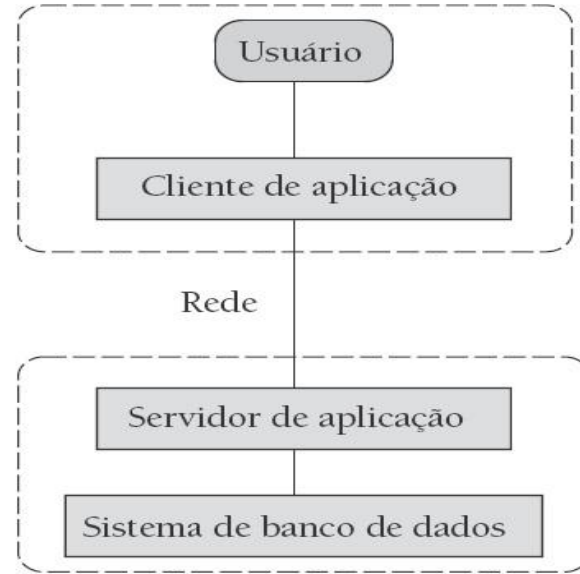
- A arquitetura de um sistema de banco de dados é bastante influenciada pelo sistema de computador subjacente em que o sistema de banco de dados é executado:
 - Centralizada
 - Cliente-servidor
 - Paralelo (multiprocessador)
 - Distribuído

Arquiteturas do Banco de Dados

- As aplicações são particionadas em duas ou três partes



(a) Arquitetura de duas camadas



(b) Arquitetura de três camadas

Modelos Lógicos baseados em registros

- Descrição dos dados nos níveis conceitual e externo
- O banco de dados é estruturado em registros de formatos fixos, de diversos tipos
- Cada tipo de registro tem sua coleção de campos ou atributos, de tamanho fixo
- Há linguagens para expressar consultas e atualizações no banco de dados
- Os três modelos mais comumente utilizados são:
 - Relacional
 - Rede
 - Hierárquico
- O modelo relacional é o mais utilizado atualmente

Modelo Relacional

- Utiliza um conjunto de tabelas para representar tanto os dados como a relação entre eles
- Cada tabela possui múltiplas colunas e cada uma possui um nome único
- Conjunto de operadores
 - Álgebra Relacional e Cálculo Relacional
- Restrições de Integridade
 - Integridade de chave primária
 - Integridade Referencial

Nome	CPF	Rua	Cidade	Nr_Conta
José	015425446	Rua das Flores	São Paulo	5418-7
Maria	154879984	Rua Linda	Bauru	4876-9
José	015425446	Rua das Flores	São Paulo	8745-6

Nr_Conta	Saldo
5418-7	541,20
4876-9	145,00
8745-6	235,90

Modelo Rede

- Os dados são representados por um conjunto de registros
- Cada registro é uma coleção de campos (atributos), cada qual contendo somente um valor
- As relações entre os registros são representados por links (ligações)
- Um link é uma associação entre dois registros

Nome	CPF	Rua	Cidade
José	015425446	Rua das Flores	São Paulo

Maria	154879984	Rua Linda	Bauru
-------	-----------	-----------	-------

Nr_Conta	Saldo
----------	-------

5418-7	541,20
--------	--------

4876-9	145,00
--------	--------

8745-6	235,90
--------	--------

Modelo Hierárquico

- É similar ao modelo em rede, pois os dados e suas relações são representados, respectivamente, por registros e links
- A diferença é que os registros estão organizados em árvores



História dos Sistemas de Banco de Dados

- Década de 1950 e início da década de 1960:
 - Processamento de dados usando fitas magnéticas para armazenamento
 - Fitas fornecem apenas acesso sequencial
 - Cartões perfurados para entrada
- Final da década de 1960 e década de 1970:
 - Discos rígidos permitem acesso direto aos dados
 - Modelos de dados de rede e hierárquico em largo uso
 - Ted Codd define o modelo de dados relacional
 - Ganharia o ACM Turing Award por este trabalho
 - IBM Research inicia o protótipo do System
 - UC Berkeley inicia o protótipo do Ingres
 - Processamento de transação de alto desempenho (para a época)

História dos Sistemas de Banco de Dados

- Década de 1980:
 - Protótipos relacionais de pesquisa evoluem para sistemas comerciais
 - **SQL se torna o padrão**
 - Sistemas de banco de dados paralelos e distribuídos
 - Sistemas de banco de dados orientados a objeto
- Década de 1990:
 - Grandes aplicações de suporte a decisão e exploração de dados
 - Grandes data warehouses de vários terabytes
 - Surgimento do comércio Web
- Década de 2000:
 - Padrões XML e XQuery
 - Administração de banco de dados automatizada
 - NoSQL
 - Big Data

Exercícios

- O que você entende por Banco de Dados? E SGBD?
- Explique cada um dos níveis de abstração de um BD definido pelo ANSI/SPARC.
- Comente a seguinte afirmação: o esquema de um banco em um modelo relacional é armazenado como se fosse um dado convencional.
- Relacione três funcionalidades providas por um SGBD.
- Escolha dois modelos de dados lógicos e discorra sobre eles.
- Quais as principais diferenças entre os modelos de rede e os modelos hierárquico?
- Defina os seguintes termos pertinentes a linguagens de consultas em BD: DDL, DML e SQL.

Atividade

- Pesquisar sobre os SGBDs existentes no mercado
 - Custos por licença de uso
 - Aplicações