

ARQUITETURA DE BANCOS DE DADOS

ANTONIO DAVID VINISKI
antonio.david@pucpr.br
PUCPR

AGENDA

ARQUITETURA DE BANCOS DE DADOS

- APRESENTAÇÃO
- ARQUITETURA DE BANCOS DE DADOS
 - História / Conceitos
 - SQL x NoSQL
 - Projeto e Arquitetura de Banco de Dados
 - Introdução ao Modelo Relacional
- PLANO DE ENSINO
 - Ementa / Temas
 - Mapa Mental
 - Processo de Avaliação
- Atividades



Who am I?

Quem sou eu?



APRESENTAÇÃO

Quem sou eu?

- Engenheiro de Computação – UEPG (2011 - 2015)
- Mestre em Computação Aplicada – UEPG (2016 - 2018)
- Professor Substituto – IFPR Telêmaco Borba (2018 - 2019)
- Doutorado em Informática – PUCPR (2019 - atual)
- Cientista de Dados – HIMARKET (2019 - atual)
- Professor Auxiliar – PUCPR – (2022 - atual)



CONTATOS

Quem sou eu?

- EMAIL:

- antonio.david@pucpr.br

- LINKEDIN - <https://www.linkedin.com/in/antonio-david-viniski-a18812b8/>

- DISCORD – Antônio David Viniski#1457



Banco de Dados

O que é? Onde vive? De o que se alimenta?

DADOS



ORDENADOS



ORGANIZADOS



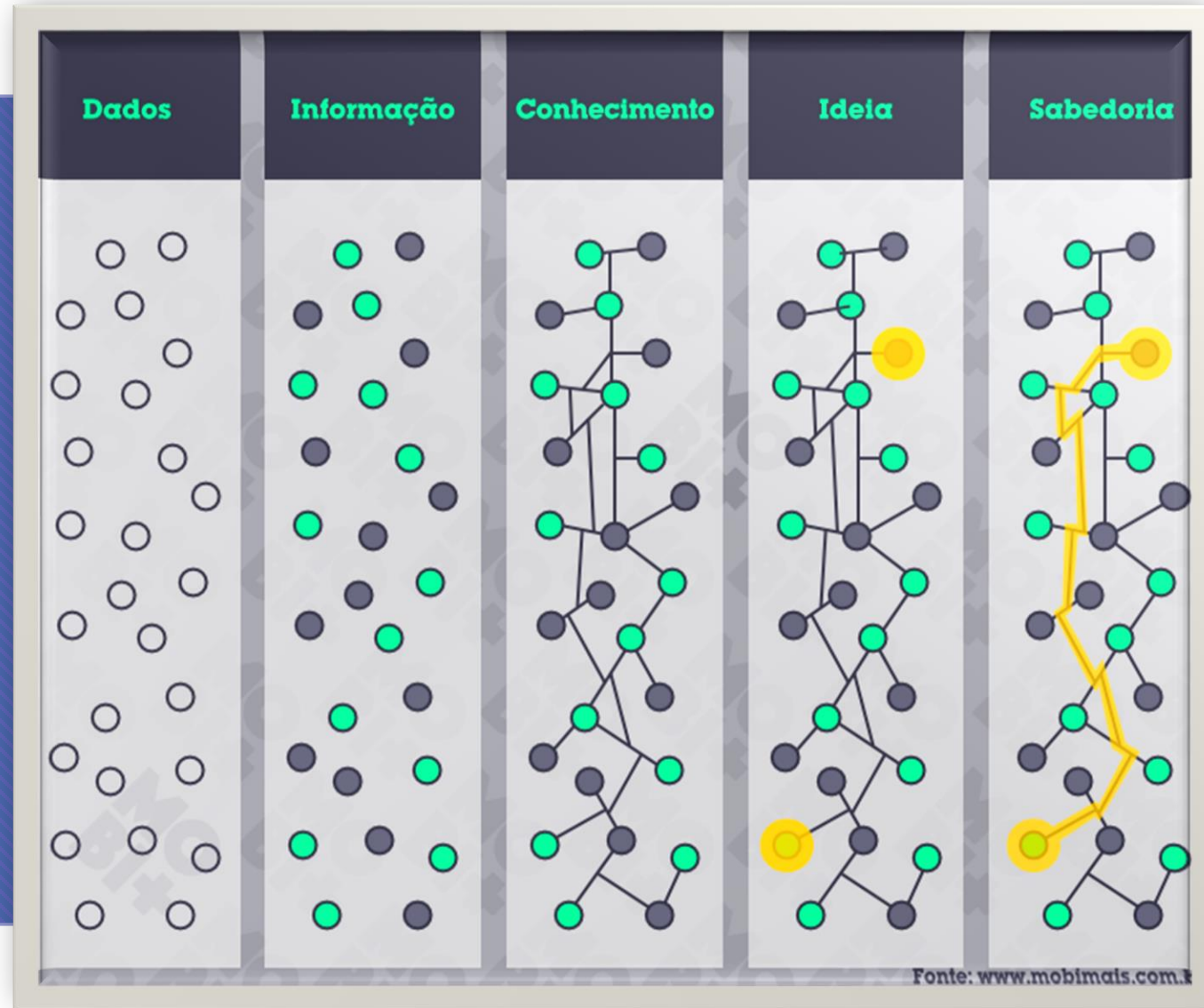
**Apresentação /
Visualização**



É justamente isso que um banco de dados faz (com um bom analista de sistemas, administrador de bancos de dados, arquiteto de informação, engenheiro de dados, cientista de dados, etc).

**DATA IS THE
NEW OIL**

**INFORMATION
IS POWER**



Polícia de Xangai usava infraestrutura do Alibaba em banco de dados invadido

Por **Gabrielly Bento**, editado por **Karol Albuquerque** | 18/07/2022 11h19

COMPARTILHAR



LEIA A SEGUIR

Recentemente, um banco de dados que armazena informações de **identificação pessoal** de um bilhão de cidadãos chineses foi roubado por um grupo de **cibercriminosos**. Segundo um relatório, a plataforma, que usava infraestrutura do Alibaba, estava online, sem **proteção de credenciais** e livre para quem soubesse onde buscar.

De acordo com o *Wall Street Journal*, a investigação para saber como conseguiram essa violação está em andamento. Foi alegado que o banco de dados do Alibaba, utilizado pela polícia de Xangai, não estava atualizado. Assim, até a configuração de uma senha para a plataforma não era uma opção.

REDES SOCIAIS

<https://olhardigital.com.br/2022/07/18/seguranca/policia-xangai-infraestrutura-alibaba-banco-de-dados/>

Troca de banco de dados pela MTU, em Cuiabá (MT), interrompe recargas dos Cartões TEM Transporte nesta sexta-feira (22)

Publicado em: 19 de julho de 2022



<https://diariodotransporte.com.br/2022/07/19/troca-de-banco-de-dados-pela-mtu-em-cuiaba-mt-interrompe-recargas-dos-cartoes-tem-transporte-nesta-sexta-feira-22/>

Banco de dados da Steam pode ter deixado escapar a data de Sonic Frontiers

A primeira data registrada indicava que a estreia aconteceria no dia 8 de novembro, mas agora ela foi atualizada para 3 de dezembro

20 de julho de 2022 por [Valdecir Emboava](#)



<https://meups.com.br/noticias/steamdb-vazado-a-data-de-sonic-frontiers/>

Prazos de atualização de dados do CadÚnico são prorrogados



A Prefeitura de Juiz de Fora (PJF) informa que o prazo para a atualização dos dados das **famílias já cadastradas** no CadÚnico, beneficiárias do Auxílio Brasil, foi prorrogado até 14 de outubro. Os usuários podem conferir se estão em revisão cadastral ou se estão com as informações atualizadas através do aplicativo do Cadastramento Único. A ferramenta informa, ainda, como prosseguir para a regularização dos registros. Para aqueles que estão em **situação de averiguação cadastral**, o prazo foi estendido até 12 de agosto.

A atualização cadastral é de extrema importância para que as informações presentes no banco de dados do CadÚnico estejam sempre condizentes com a realidade das famílias brasileiras. Além disso, a falta de atualizações pode levar a uma suspensão do benefício, e um possível cancelamento.

<https://www.pjf.mg.gov.br/noticias/view.php?modo=link2&idnoticia2=76117>

Área de TI no Brasil: os melhores cargos e salários

Nos últimos tempos, a área de TI no Brasil vem se aquecendo cada vez mais e novas oportunidades surgem no mercado de trabalho. Entenda o porquê.



Digital House

11 de Ago de 2021 • 5 min de leitura



Segurança da Informação

Este profissional possui a responsabilidade de garantir a confiabilidade, a integridade e a disponibilidade de todos os dados que fazem parte de uma infraestrutura de TI ou da administração de redes.

É importante ter um perfil analítico para lidar com as demandas do dia a dia, além de sempre buscar por novas atualizações e saber quais são as ameaças que existem no mercado. Sendo assim, é possível estruturar políticas de prevenção, além de ações para detectar e diminuir riscos à organização.

Dados

São milhares de dados que circulam no universo digital e é fundamental que as empresas sejam capazes de coletar, analisar e identificar tudo o que for relevante. Com isso, elas passam a entender, de fato, o que está acontecendo em seu nicho de atuação e a necessidade de seus clientes.

Essa área é responsável pela administração de banco de dados das organizações, garantindo a confiabilidade e a veracidade das informações contidas. A partir delas, as empresas tomam iniciativas, otimizam suas demandas internas, aplicam melhorias e tomam decisões estratégicas.

[Área de TI: os melhores cargos e salários do mercado \(digitalhouse.com\)](https://digitalhouse.com)

[Guia 2022] Conheça as áreas de TI mais bem pagas

Análise e Desenvolvimento de Sistemas, Cibersegurança, Ciência de Dados, Dev Ops, Dicas, Engenharia de Software, Gerenciamento de Projetos de TI, Graduação, Guia de Profissões, Redes de Computadores, Tecnologia da Informação

Engenheiro de Dados

Atualmente, estamos cercados de dados e informações de todos os tipos. Com a popularização da internet, a qualidade de dados aumenta exponencialmente. E para transformar essas informações em conhecimento que possam gerar resultados para as empresas, é possível utilizar a tecnologia. Para gerenciar essa quantidade astronômica de informações, engenheiros de dados são as pessoas mais indicadas.

Eles sabem como lidar, processar e usar os dados a favor das empresas. Há tecnologias e técnicas específicas para tornar o trabalho mais preciso e que trazem informações importantes que farão a diferença para as instituições. **O salário médio deste profissional é de R\$ 8.884,00.**

Data Base Administrator (DBA)

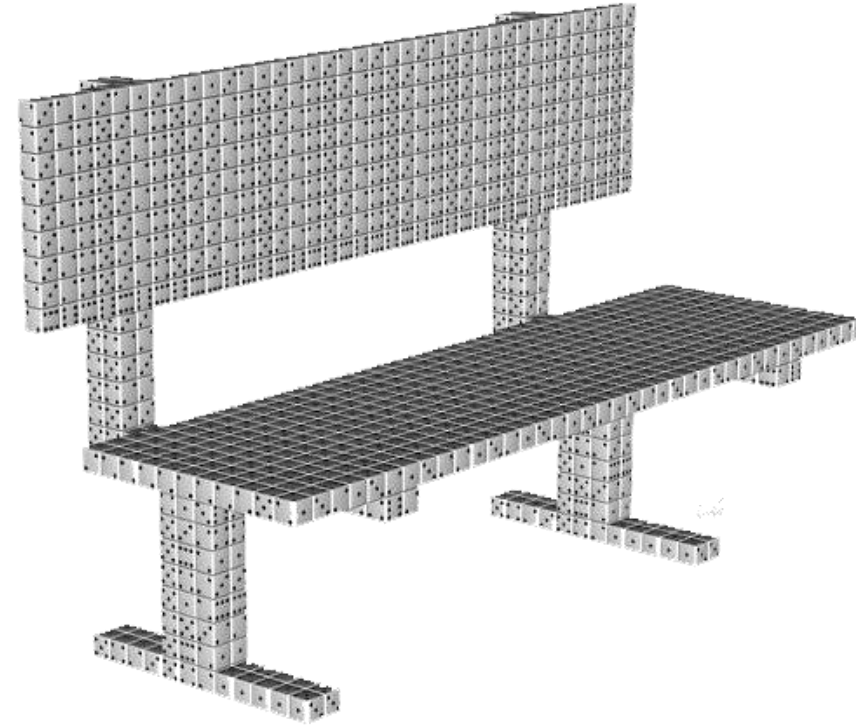
O profissional de DBA é o administrador de banco de dados. Fica sob sua responsabilidade o cuidado com as informações que a empresa possui. Aqui falamos de dados pessoais de associados, clientes, prontuários de pacientes, dados bancários, governamentais e de pesquisas, só para ficar em alguns exemplos.

O administrador de banco de dados é como uma espécie de guardião também. Ele deve saber acessar esses dados, quando solicitados, e cuidar para que esses não sejam indevidamente acessados. **O salário deste profissional é de R\$ 6.732,00.**

[\[Guia 2022\] Conheça as áreas de TI mais bem pagas - Blog da Ampli](#)

IMPORTÂNCIA DE UM BANCO DE DADOS

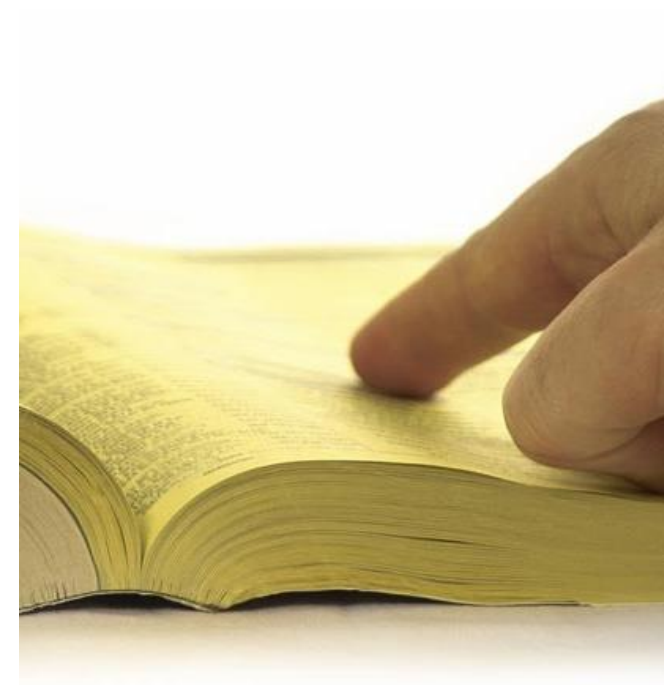
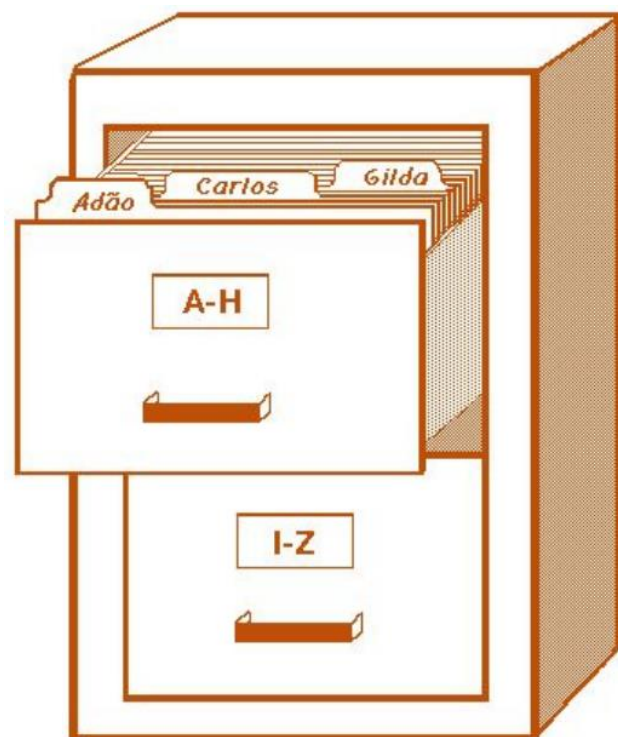
- Encontre uma notícia recente que mostre o impacto que o banco de dados pode ter em uma organização.
 - A. Informe a data da notícia.
 - B. Elencar as principais características da notícia (Vazamento de informações, Atualizações Cadastrais, Perda de Dados [erros de programação], Atualizações na Estrutura, etc).
 - C. Qual o impacto do acontecimento para a organização?



Mas o que seria mesmo um banco de dados?
Vamos voltar no tempo...

BANCO DE DADOS

História I



BANCO DE DADOS

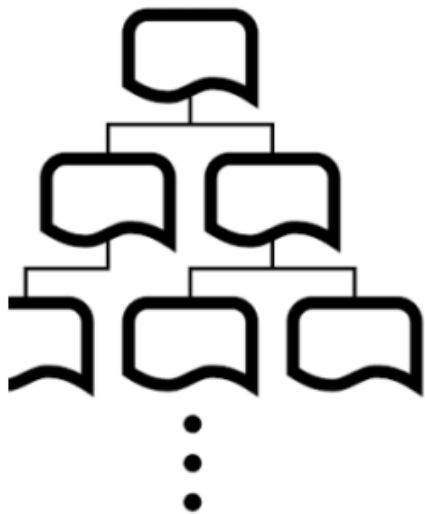
História II

GERAÇÃO	DATA	CARACTERÍSTICAS
Primeira	1950	Sistema de arquivos; Acesso Seqüencial; Memória secundária: fitas magnéticas e cartões perfurados.
Segunda	1960	Sistema de arquivos; Acesso Randômico; Memória secundária: discos magnéticos.
Terceira	1970	Geração pré-relacional; Banco de dados; Informação lógica e física: pouca distinção; Modelos emergentes: redes e hierárquico.
Quarta	1980	Geração relacional; Sistemas de banco de dados; Informação lógica e física: clara distinção; Modelo emergente: relacional.
Quinta	1990	Geração pós-relacional; Modelos: estendido, orientado a objetos e orientado a regras.

BANCO DE DADOS

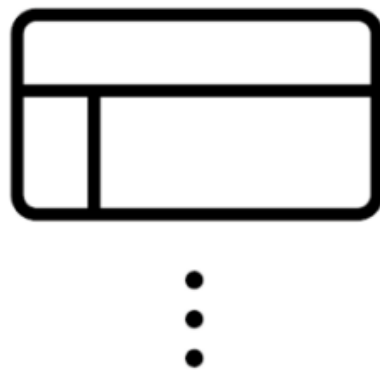
História III

Hierarchical



1960s

Relational



1980s

NoSQL



2000s

Graph Thinking



2020s

Dinâmica em Grupo

Pesquisa e Discuta em Grupo

- O que é um banco de dados?
- Quais tipos de dados podem ser armazenados em um BD?
- Quais os bancos de dados SQL mais utilizados atualmente?
- Quais os principais componentes de um banco de dados relacional?
- Como a tecnologia de bancos de dados está presente no seu dia-a-dia?

INÍCIO

SGBD: Modelagem,
estrutura, consulta.
Aplicações tradicionais

AGORA

NoSQL: documentos, texto,
grafos e afins
Conexão: Web, streams,
móveis, paralelo, ...
Contexto: nuvem, hardware,
privacidade...



Recuperação de Informação

Coleta, extração, indexação,
processamento, entrega
E.g., sistemas de
recomendação e
classificadores

Mineração de Dados

Descoberta de Conhecimento
Mineração de Dados
E.g., aprendizado de máquina
e big data analytics

Bancos de Dados é tudo isso e mais

Absolutamente qualquer aplicação
de Computação precisa de dados
(I/O)

Gerenciar os dados é o primeiro
passo para qualquer delas

Estrutura dos dados

Dados **estruturados**

Estrutura **rígida** e previamente **planejada**

Organização em **blocos semânticos (relações)** e definição de mesmas descrições para dados de um mesmo grupo (**atributos**).

Ex.: bancos de dados

Dados **semi**estruturados

Representação estrutural heterogênea (não completamente desestruturados e nem fortemente tipados)

Auto-descritivo (esquema de representação junto com os dados)

Ex.: XML, RDF, OWL.

Dados **não** estruturados

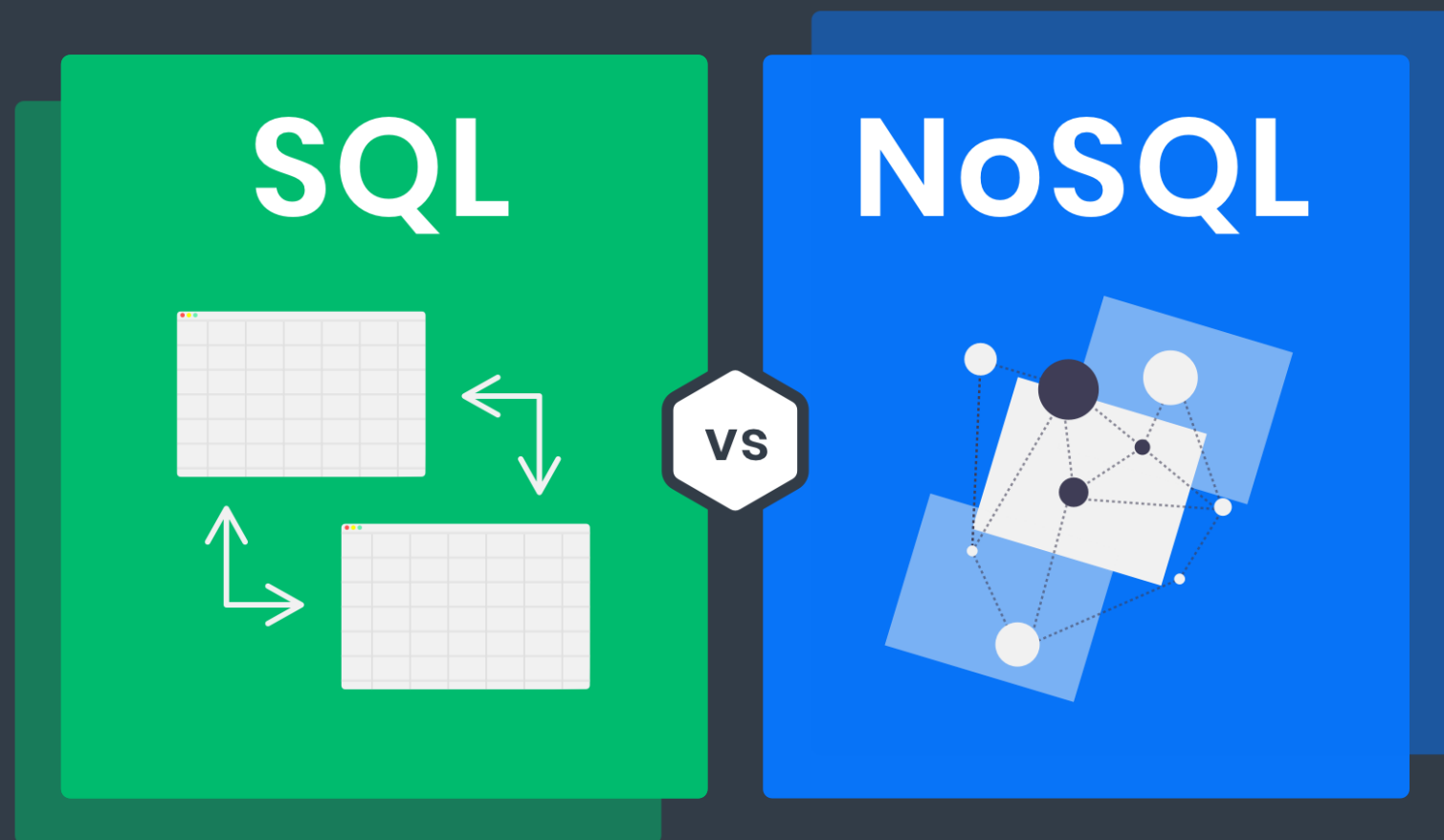
Sem estrutura pré-definida

Constituem a **maioria dos dados corporativos** (cerca de 90%)

Ex.: relatórios, documentos, imagens, áudios e vídeos.

SQL vs NoSQL

Definição



SQL vs NoSQL

Definição

SQL

- Structured Query Language, ou Linguagem de Consulta Estruturada ou SQL, é a **linguagem** de pesquisa declarativa **padrão** para banco de dados **relacional**.

NoSQL

- NoSQL (originalmente se referindo a "no SQL": "não SQL" ou "**não relacional**", posteriormente estendido para Not Only SQL - Não Somente SQL) é um termo genérico que representa os bancos de dados não relacionais.

SQL

Vantagens e Desvantagens

Vantagens

- **Sintaxe amigável**: fácil de ler e entender.
- **Integridade referencial**: relacionamento entre as tabelas, compatível com ACID (Atomicity, Consistency, Isolation, Durability).
- **Eficiência de recursos**: redução da duplicação de dados, uso eficiente dos recursos disponíveis e ao desempenho mais rápido.
- **Linguagem estável**: padrões bem definidos e muitos usuários experientes.

Desvantagens

- Requer **dimensionamento vertical**, o que não sai barato. A escala horizontal é muito desafiadora com SQL.
- Requer que os desenvolvedores usem uma **estrutura predefinida** e definam o esquema. Isso exige um **design inicial cuidadoso**.
- As consultas em bancos de dados SQL podem se tornar **lentas**, pois podem exigir a **junção** de dados de várias tabelas.

NoSQL

Vantagens e Desvantagens

Vantagens

- **Alta escalabilidade**: escalabilidade horizontal perfeita, alta disponibilidade e sem pontos únicos de falha.
- **Flexibilidade**: sem estrutura específica desde o início. Adição de recursos aos seus dados em tempo real.
- **Facilidade de uso**: inserir dados é fácil e compatível com linguagens de programação populares.
- **Alto desempenho em escala**: podem atingir níveis extremamente altos de desempenho à medida que os dados crescem.
- **Consultas rápidas**: As consultas não exigem a junção de dados de tabelas diferentes para que os dados possam ser recuperados rapidamente.

Desvantagens

- A maioria dos bancos de dados NoSQL **não** oferece suporte a transações **ACID** em vários documentos.
- Com exceção dos bancos de dados de uso geral, muitos bancos de dados NoSQL são otimizados apenas para casos de **uso específicos**.
- NoSQL é uma tecnologia relativamente **jovem e imatura**. Não é tão bem desenvolvido e estável quanto o SQL.

Conceitos BD

Definições

- **[Chu, 1985]**

Um banco de dados é um conjunto de arquivos relacionados entre si.

- **[Date, 1975]**

Um banco de dados é uma coleção de dados operacionais usados pelas aplicações de uma determinada organização.

Conceitos BD

Propriedades

- **[Elmasri & Navathe, 2000]**

Banco de dados é uma coleção de dados relacionados.

- Entretanto, um BD possui as seguintes propriedades implícitas:

- Representa aspectos do mundo real (mini-mundo ou universo de discurso).
- É uma coleção de dados logicamente coerentes com algum significado inerente.
- É projetado, construído e instanciado (“povoado”) para uma aplicação específica.

Conceitos

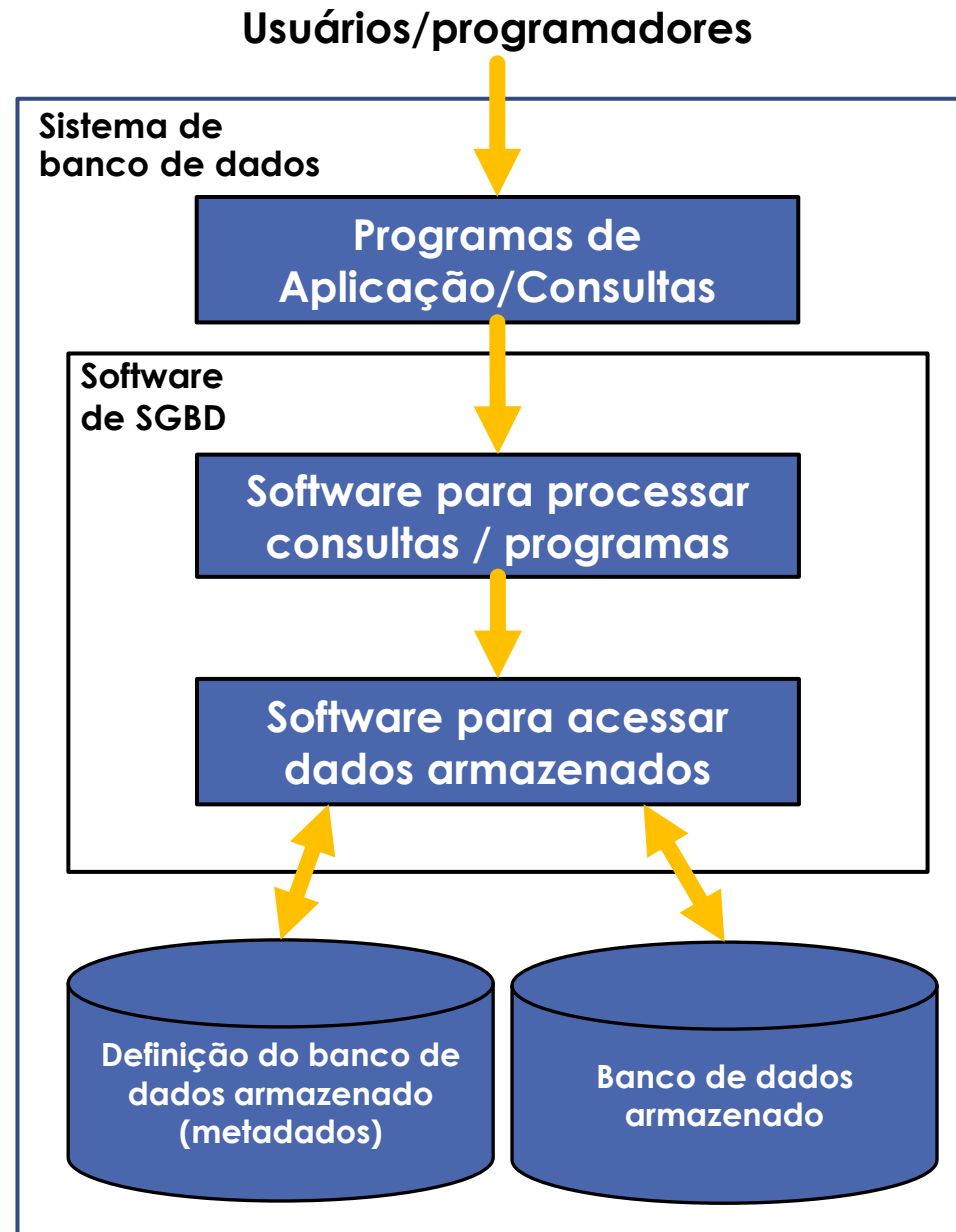
SGBD

Um Sistema Gerenciador de Banco de Dados (SGBD) é um conjunto de programas que permite criar e manter (manipular) um banco de dados.

- Por meio desse sistema, um usuário pode:
 - Adicionar, alterar ou remover arquivos (tabelas/documentos) no banco de dados;
 - Inserir, alterar, remover ou recuperar dados dos arquivos (tabelas/documentos).

SGBD + BD = sistema de banco de dados

Sistema de Banco de Dados



SGBD - Vantagens

Controle de Redundância
dos Dados

Controle de Acesso
(Segurança)

Armazenamento
Persistente dos Dados

Existência de Múltiplas
Interfaces para os Usuários

Manutenção de Restrições
de Integridade

Acesso Multiusuário
(Controle de Concorrência
e recuperação de Falhas)

Linguagens de Bancos de Dados

Sistema de Banco de Dados

- O sistema de Banco de Dados fornece duas linguagens:
 - Linguagem de Definição de Dados (DDL).
 - Linguagem de Manipulação de Dados (DML).

Linguagem de Definição de Dados (DDL)

Linguagens de Bancos de Dados

- A DDL é responsável pela especificação:
 - Esquema de banco de dados.
 - Propriedades adicionais dos dados.
 - Estrutura de armazenamento.
 - Métodos de acesso.
- Especifica as restrições de integridade:
 - Restrições de domínio (Tipos de dados).
 - Integridade referencial (Relacionamentos).
 - Autorização (Controle de Acesso).



```
create table departamento  
  (nome_dept   char (20),  
   edificio    char (15),  
   orçamento   numeric (12,2));
```

Linguagem de Manipulação de Dados (DML)

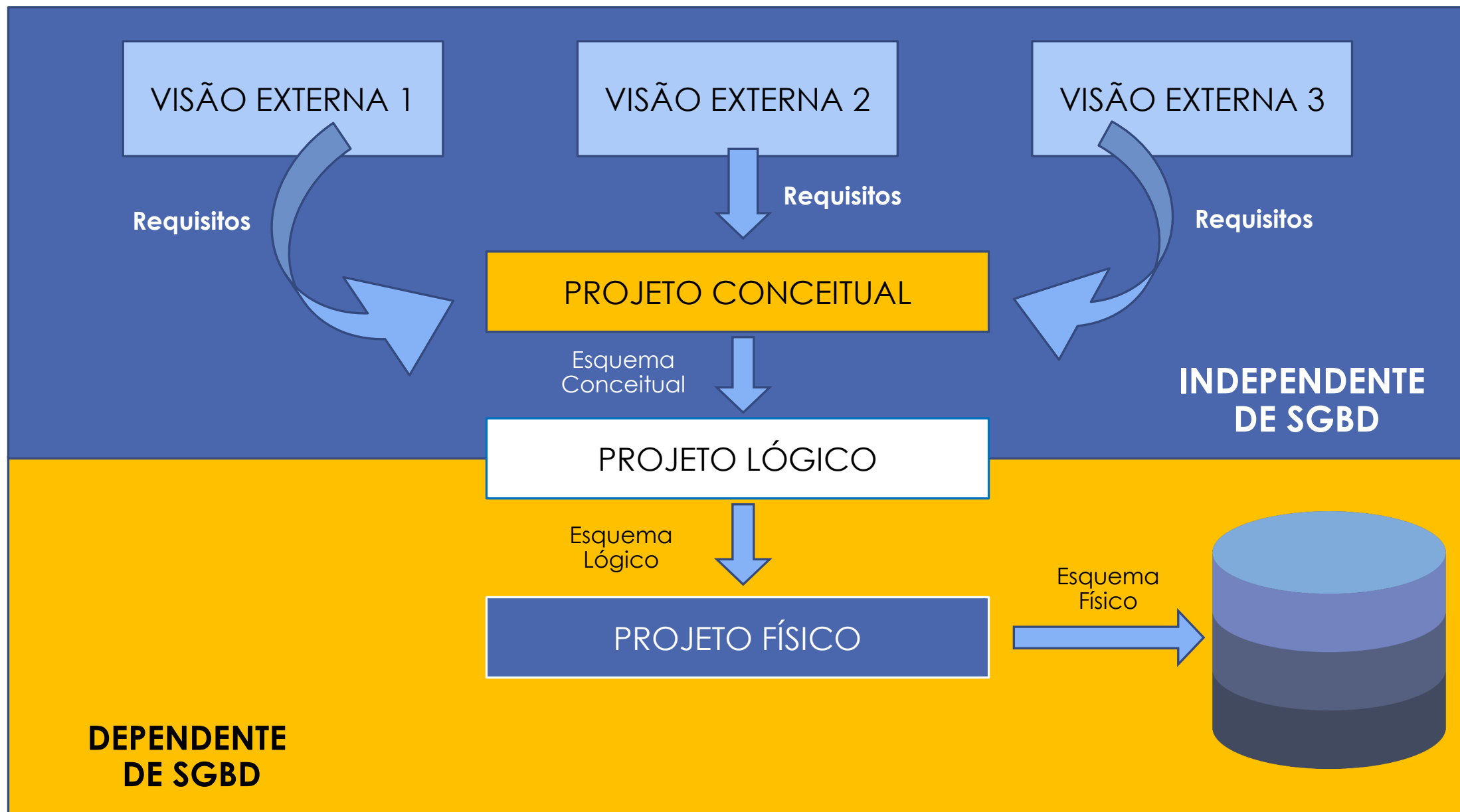
Linguagens de Bancos de Dados

- Permite aos usuários:
 - Acessar ou manipular dados conforme organizados.
- Tipos de acesso:
 - Recuperação de informações armazenadas no banco de dados.
 - • Inserção de novas informações no banco de dados.
 - • Exclusão de informações do banco de dados.
 - • Modificação de informações armazenadas no banco de dados.

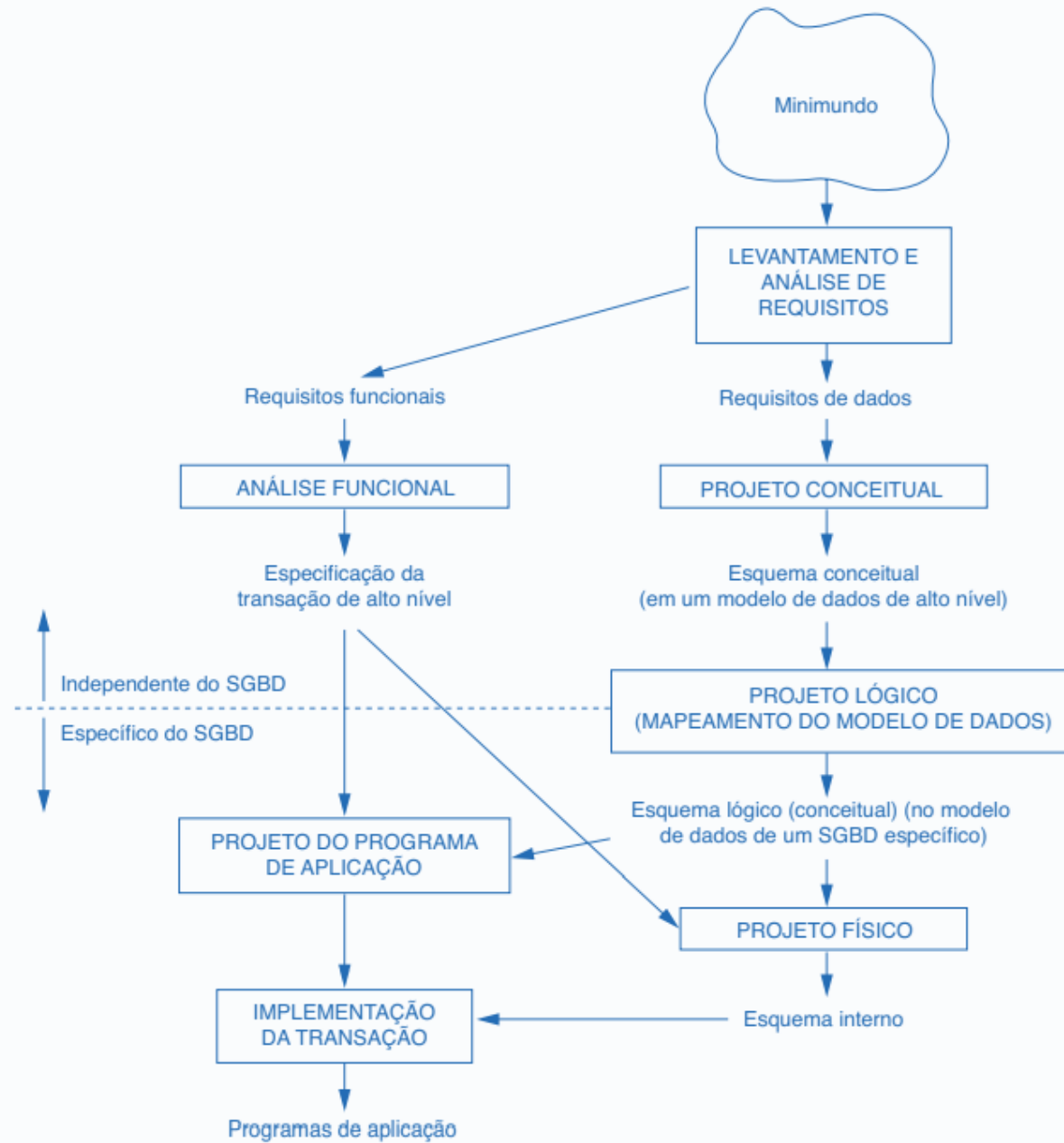


```
select instrutor.nome  
from instrutor  
where instrutor.nome_dept = 'History';
```

Projeto de Banco de Dados



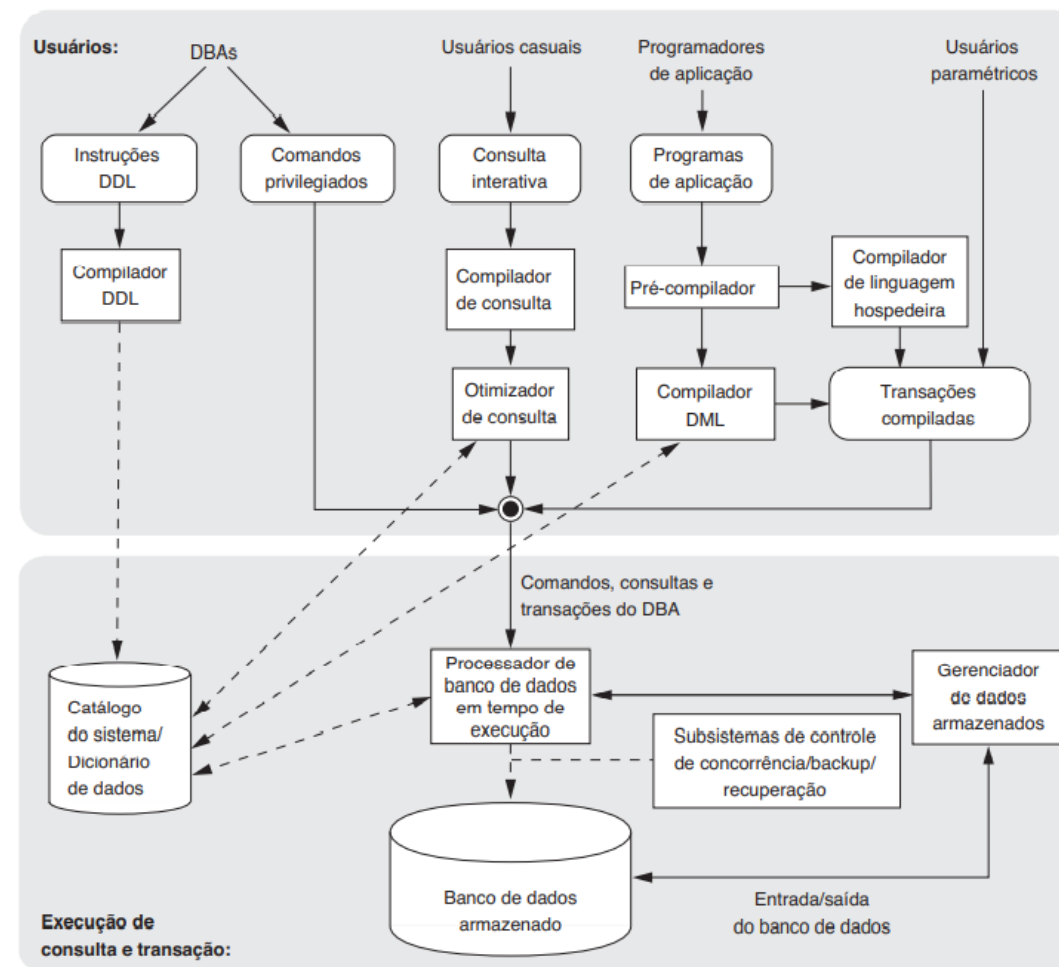
Principais Fases do Projeto de Banco de Dados



Arquitetura de banco de dados

Jogo de Palavras

- Instruções.
- Compilador.
- Catálogo.
- Consulta.
- Transação.
- DBA – *Data Base administrator*.
- DDL – *Data Definition Language*.
- DML – *Data Manipulation Language*.



Modelo Relacional

Conceitos

Modelo de dados Relacional

Relações, Tuplas, Atributos, Domínios

Nome da relação

ALUNO

Atributos

Nome	Cpf	Telefone_residencial	Endereco	Telefone_comercial	Idade	Media
Bruno Braga	305.610.243-51	(17)3783-1616	Rua das Paineiras, 2918	NULL	19	3,21
Carlos Kim	381.620.124-45	(17)3785-4409	Rua das Goiabeiras, 125	NULL	18	2,89
Daniel Davidson	422.111.232-70	NULL	Avenida da Paz, 3452	(17)4749-1253	25	3,53
Roberta Passos	489.220.110-08	(17)3476-9821	Rua da Consolação, 265	(17)3749-6492	28	3,93
Barbara Benson	533.690.123-80	(17)3239-8461	Rua Jardim, 7384	NULL	19	3,25

Tuplas

PLANO DE ENSINO

Resultados Esperados

EMENTA/OBJETIVOS

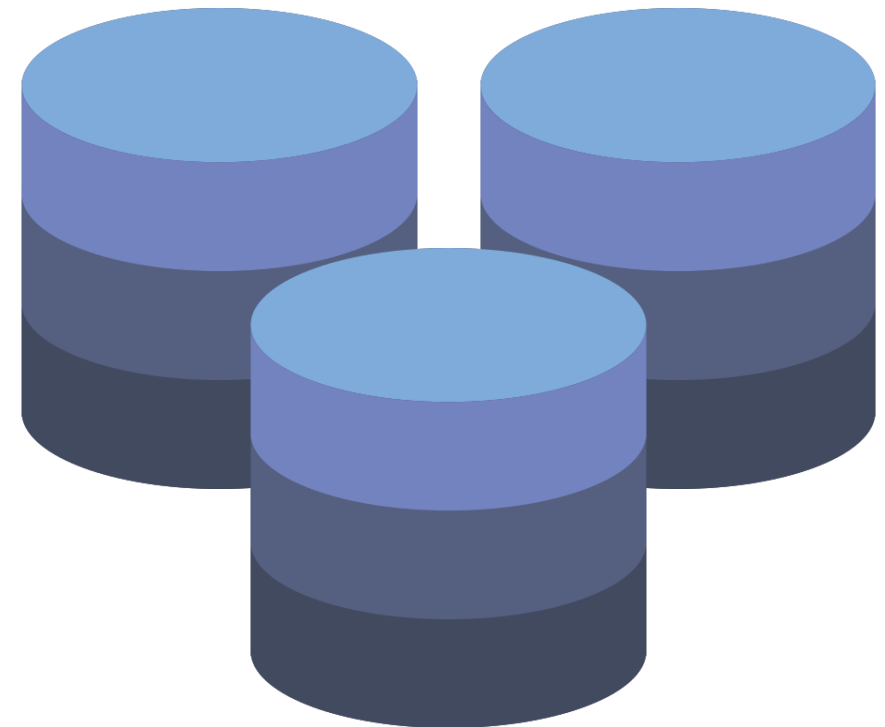
A disciplina de **Arquitetura de Banco de Dados** é destinada a estudantes dos cursos de computação



- **Analisar** um problema de armazenamento de dados (**Requisitos**).
- **Representar** a solução em modelos conceitual, lógico e físico (**Modelagem**).
- **Criar** e/ou **atualizar** o banco de dados (Inserir, Alterar, Recuperar, Excluir) (**SQL**).
- **Criar** modelos a partir de coleções de dados (**Engenharia Reversa**).

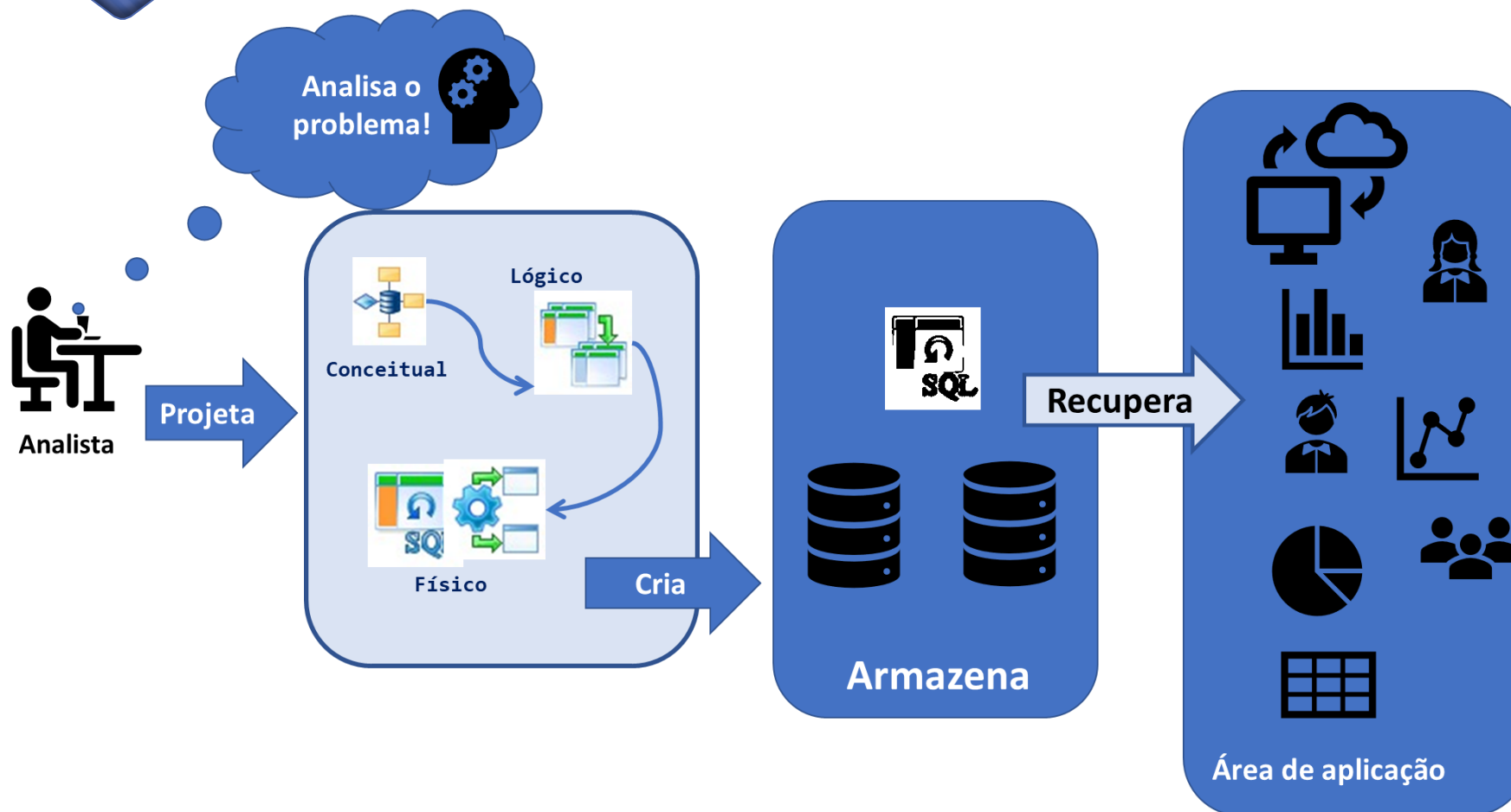
TEMAS DE ESTUDO

- **TE01** – Componentes da arquitetura de banco de dados.
- **TE02** – Processo de Modelagem de Dados Relacional.
- **TE03** – Esquemas em banco de dados relacional.
- **TE04** – Manipulação de dados utilizando Linguagem SQL (*Structured Query Language*) em Sistemas Gerenciadores de Banco de Dados (SGBD).
- **TE05** – Operadores de Álgebra Relacional em consultas sobre dados.
- **TE06** – Controle de acesso aos dados.
- **TE07** – Concorrência e controle de transação.



Mapa Mental

Evolução das Competências



Processo de Avaliação

Como?



Formativa:

- Lista de exercícios.
- Resolução de problemas.
- Atividades em Equipe.
- Feedbacks

Avaliações Somativas: Trabalhos

- Avaliação Somativa 1 Modelagem.
- Avaliação Somativa 2 SQL Avançado.
- Feedbacks

Avaliações Somativas: Provas

- Prova Individual 1 (Teórica e Prática).
- Prova Individual 2 (Teórica e Prática).
- Feedbacks.

Agora imagine uma Universidade

Quais dados existem?

- Quais dados?
- Como são organizados?
- Como são armazenados?
- Como se relacionam?
- Como são inseridos?
- Atualizados? Consultados?
- Por quem e quando?



Pessoas

Docentes
Estudantes
Funcionários



Departamentos



Cursos



Disciplinas



Turmas



Currículos

PROJETO

Discussão sobre o projeto de Banco de Dados da disciplina, englobando todos os tópicos principais apresentados durante o semestre.