

# Aula 2: Conceitos e Arquiteturas de SGBD's

## Modelos de Dados, Esquemas e Instâncias

Uma das características fundamentais da abordagem de base de dados é que ela fornece algum nível de abstração de dados, pela omissão de detalhes de armazenamento de dados que não são necessários para a maioria dos usuários. O **modelo de dados** é a principal ferramenta que fornece esta abstração. Ele é um conjunto de conceitos que podem ser usados para descrever a estrutura de uma base de dados. Por estrutura de uma base de dados entende-se os tipos de dados, relacionamentos e restrições pertinentes aos dados. Muitos modelos de dados também definem um conjunto de operações para especificar como recuperar e modificar a base de dados.

### Categorias de Modelos de Dados

Muitos modelos de dados têm sido propostos. Pode-se classificar os modelos de dados baseando-se nos tipos de conceitos que fornecem para descrever a estrutura da base de dados. **Modelos de Dados Conceituais** ou de Alto-Nível fornecem conceitos próximos à percepção dos usuários. Já os **Modelos de Dados Físicos** ou de Baixo-Nível fornecem conceitos que descrevem os detalhes de como os dados são armazenados no computador.

Modelos de alto-nível utilizam conceitos tais como Entidades, Atributos e Relacionamentos. Uma entidade é um objeto que é representado na base de dados. Um atributo é uma propriedade que descreve algum aspecto de um objeto. Relacionamentos entre objetos são facilmente representados em modelos de dados de alto-nível, que são algumas vezes chamados de Modelos Baseados em Objeto devido, principalmente, a sua característica de descreverem objetos e seus relacionamentos.

Modelos de Dados de Baixo-Nível descrevem como os dados são armazenados no computador, representando informações em formato de registros, ordem dos registros e caminho de acesso. Um Caminho de Acesso é uma estrutura de que facilita a busca de um registro particular na base de dados.

### Esquemas e Instâncias

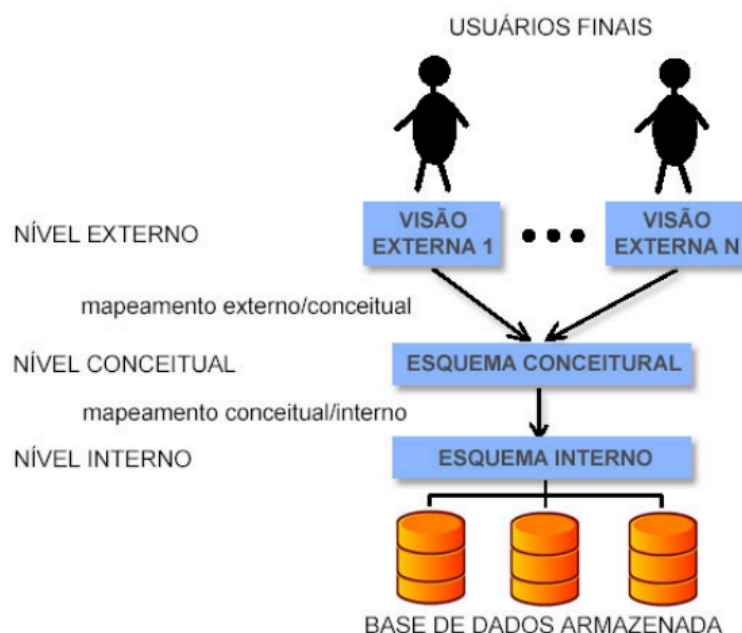
Em qualquer modelo de dados é importante distinguir entre a descrição da base de dados e a base de dados propriamente dita. A descrição de uma base de dados é chamada **Esquema da Base**

**de Dados.** Um esquema de base de dados é especificado durante o projeto da base de dados, sendo que a expectativa de mudanças não é grande. A forma de visualização de um esquema é chamada Diagrama do Esquema. Muitos modelos de dados têm certas convenções para, diagramaticamente, mostrar esquemas especificados no modelo.

Os dados atualmente existentes em uma base de dados podem mudar com relativa frequência. Os dados da base de dados em um particular momento do tempo são chamados **Instâncias da Base de Dados** (ou Ocorrências ou Estados). A base-esquema é algumas vezes chamada de Base-Intencional e uma instância é chamada de Base-Extensional do esquema.

## Arquitetura e Independência de Dados de SGBD's

A arquitetura mais difundida na literatura é a Arquitetura **Três-Esquemas** (também conhecida como arquitetura ANSI/SPARC), proposta por Tsichritzis & Klug em 1978. A meta desta arquitetura, exibida na Figura 1, é separar as aplicações de usuários da base de dados física.



**Figura 1:** Arquitetura Três-Esquemas.

Nesta arquitetura, esquemas podem ser definidos em três níveis:

1. O nível interno tem um esquema interno que descreve a estrutura de armazenamento físico da base de dados. O esquema interno usa um modelo de dados físico e descreve todos os detalhes de armazenamento de dados e caminhos de acesso à base de dados;
2. O nível conceitual tem um esquema conceitual que descreve a estrutura de toda a base de dados. O esquema conceitual é uma descrição global da base de dados, que omite detalhes da estrutura de armazenamento físico e se concentra na descrição de entidades, tipos de

dados, relacionamentos e restrições. Um modelo de dados de alto-nível ou um modelo de dados de implementação podem ser utilizados neste nível.

3. O nível externo ou visão possui esquemas externos ou visões de usuários. Cada esquema externo descreve a visão da base de dados de um grupo de usuários da base de dados. Cada visão descreve, tipicamente, a parte da base de dados que um particular grupo de usuários está interessado e esconde deste o restante da base de dados. Um modelo de dados de alto-nível ou um modelo de dados de implementação podem ser usados neste nível.

Muitos SGBD's não separam os três níveis completamente. Pode acontecer que alguns SGBD's incluam detalhes do nível interno no esquema conceitual. Em muitos SGBD's que permitem visões, os esquemas externos são especificados com o mesmo modelo de dados usado no nível conceitual.

Note que os três esquemas são apenas descrições dos dados. A arquitetura Três-Esquemas pode ser utilizada para explicar conceitos de independência de dados, que podem ser definidos como a capacidade de alterar o esquema de um nível sem ter que alterar o esquema no próximo nível superior. Dois tipos de independência de dados podem ser definidos:

- **Independência Lógica de Dados:** É a capacidade de alterar o esquema conceitual sem ter que mudar os esquemas externos ou programas de aplicação. Pode-se mudar o esquema conceitual para expandir a base de dados, com a adição de novos tipos de registros (ou itens de dados), ou reduzir a base de dados removendo um tipo de registro. Neste último caso, esquemas externos que se referem apenas aos dados remanescentes não devem ser afetados;
- **Independência Física de Dados:** É a capacidade de alterar o esquema interno sem ter que alterar o esquema conceitual externo. Mudanças no esquema interno podem ser necessárias devido a alguma reorganização de arquivos físicos para melhorar o desempenho nas recuperações e/ou modificações. Após a reorganização, se nenhum dado foi adicionado ou perdido, não haverá necessidade de modificar o esquema conceitual.

## Linguagens de Base de Dados

**Linguagem de Definição de Dados** (“Data Definition Language” - DDL): é utilizada pelo DBA e projetistas de base de dados para definir seus esquemas. O SGBD tem um compilador para processar descrições em DDL e construir a descrição do esquema armazenado no catálogo;

**Linguagem de Manipulação de Dados** (“Data Manipulation Language” - DML): uma vez que o esquema é compilado e a base de dados preenchida com dados, os usuários têm que ter algum modo de manipular os dados. Manipulações comuns como recuperação, inserção, remoção e modificação de dados são realizadas pela DML.

# Exercícios

1. O que é modelo de dados?
2. O que é um esquema? Cite dois exemplos.
3. O que é uma instância? Cite dois exemplos.
4. Cite e explique os três níveis da Arquitetura Três-Esquemas.
5. Qual a diferença entre independência lógica de dados e independência física?
6. Qual a diferença entre as linguagens DDL e DML?