

Banco de Dados

Arquiteturas de Gerenciamento de Bancos de Dados

André Santanchè

Laboratory of Information Systems - LIS

Instituto de Computação - UNICAMP

Museu Exploratório de Ciências da Unicamp

Outubro 2018

Objetivo

- Apresentar um panorama sobre a **Arquitetura** de um Sistema Gerenciador de Banco de Dados, ou seja, a forma como ele se organiza e interage com as aplicações.

Sumário

- Recapitulando
- Motivação: Para que um Banco de Dados?
- Sistema Gerenciador de Bancos de Dados
- Arquitetura
 - Componentes do SGBD
 - SGBD & rede

Motivação

- Aplicações computacionais de todos os portes trabalham com grandes volumes de dados
 - Dados produzidos por uma pessoa (uso de dispositivos móveis, redes sociais etc.)
 - Gerenciamento de uma farmácia
 - Dados geográficos e pontos de referência (desde o bairro até o planeta)
 - Sistema bancário
 - Dados de um hospital (vários tipos)
 - Dados da Web e redes sociais

Motivação

- Grandes volumes de dados e suas relações complexas justificam a criação de estratégias específicas para gerenciá-los

Motivação

- Grandes volumes de dados e suas relações complexas justificam a criação de estratégias específicas para gerenciá-los



Bancos de Dados

Aplicações Tradicionais

- Bancos de dados numéricos e tradicionais
- Exemplos:
 - Gerenciamento de uma farmácia
 - Sistema bibliotecário
 - Sistema bancário

Aplicações

- Aplicações mais recentes
 - Bancos de Dados Multimídia - vídeos no Youtube; exames de um hospital; etc.
 - Sistemas de Informação Geográfica (GIS)
 - Dados em rede: e.g., redes sociais
 - etc.

Banco de Dados

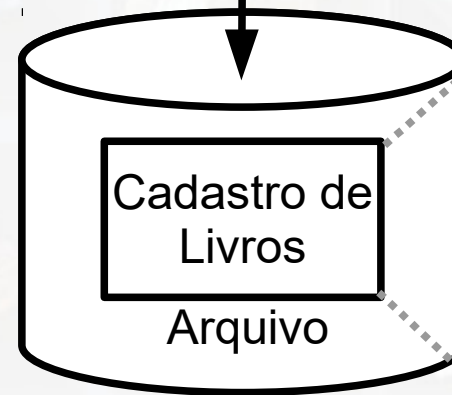
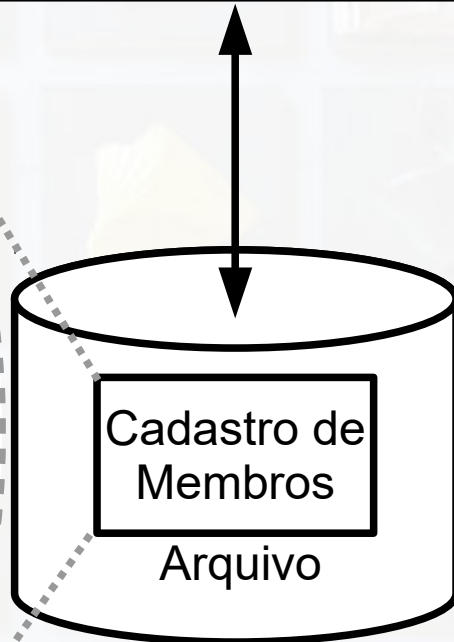
Aplicação Exemplo

- Gerenciamento de uma biblioteca
- Serviços:
 - Cadastro de membros associados
 - Registro do acervo (ex.: livros, revistas etc.)
 - Controle de empréstimos

Sistemas Isolados

Programa de
Controle de Membros

Programa de
Registro do Acervo



Sistemas/Arquivos Isolados

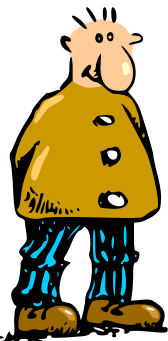
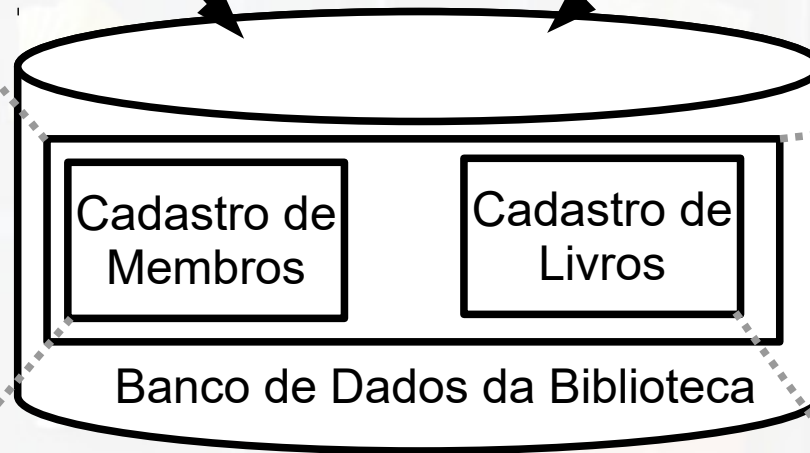
- Redundância não controlada
 - Repetição
 - Inconsistência
- Barreiras para relacionamento entre arquivos
- Dificuldades com:
 - acesso concorrente
 - integridade e recuperação em caso de crash
 - segurança e controle de acesso

Banco de Dados

Compartilhamento de Dados

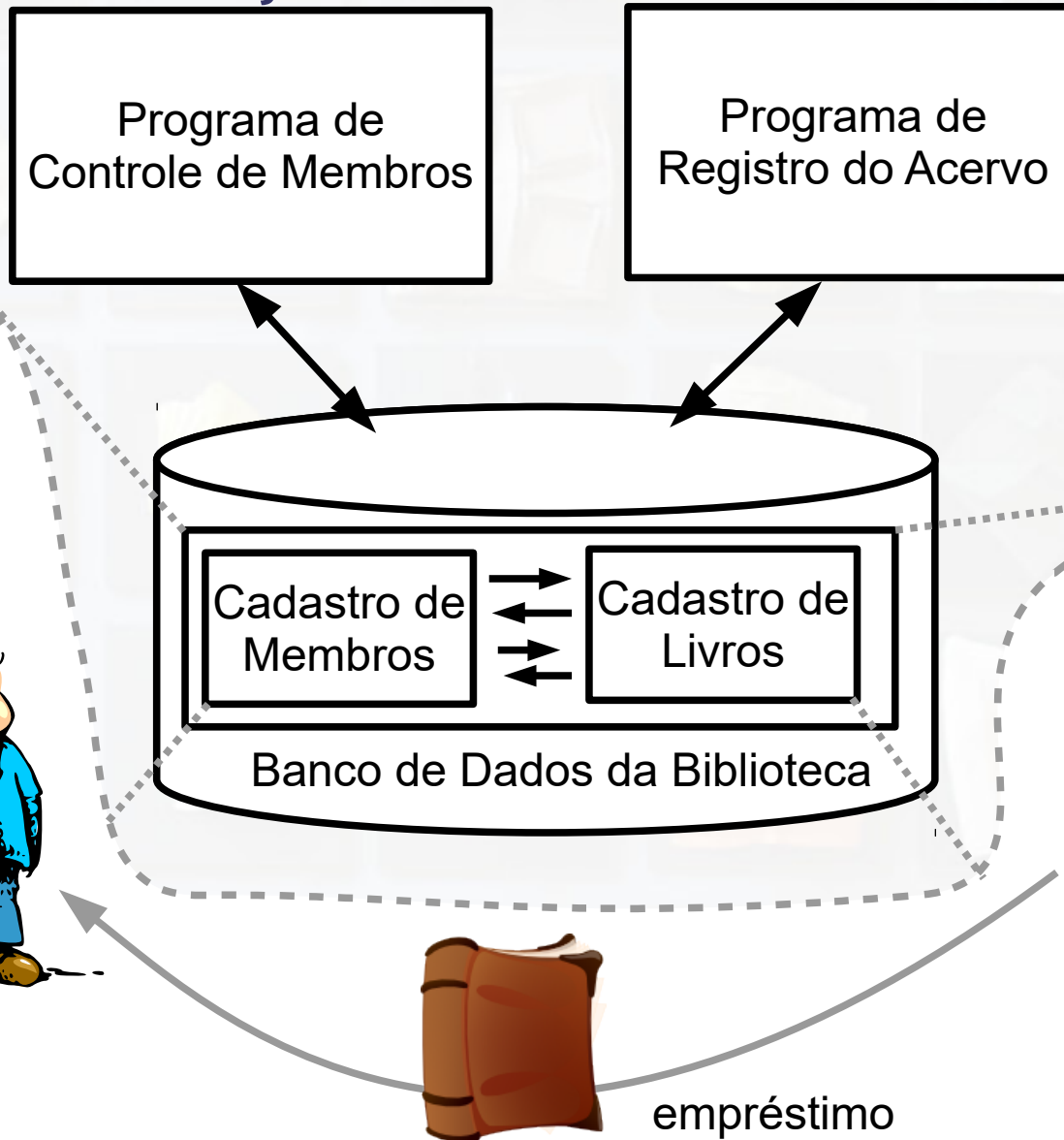
Programa de
Controle de Membros

Programa de
Registro do Acervo



Banco de Dados

Relações entre os dados



Sistema Gerenciador de Banco de Dados (SGBD)

- Sistema de software com finalidade genérica
- Projetado para a definição, construção e manipulação de bancos de dados
- Pode atender várias aplicações

SGBD

Programa de
Controle de Membros

Programa de
Registro do Acervo

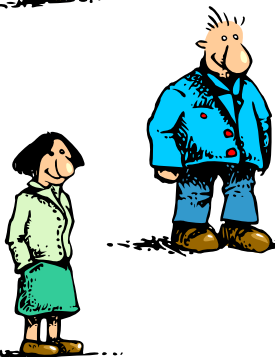
SGBD
Sistema Gerenciador
de Banco de Dados

Cadastro de
Membros

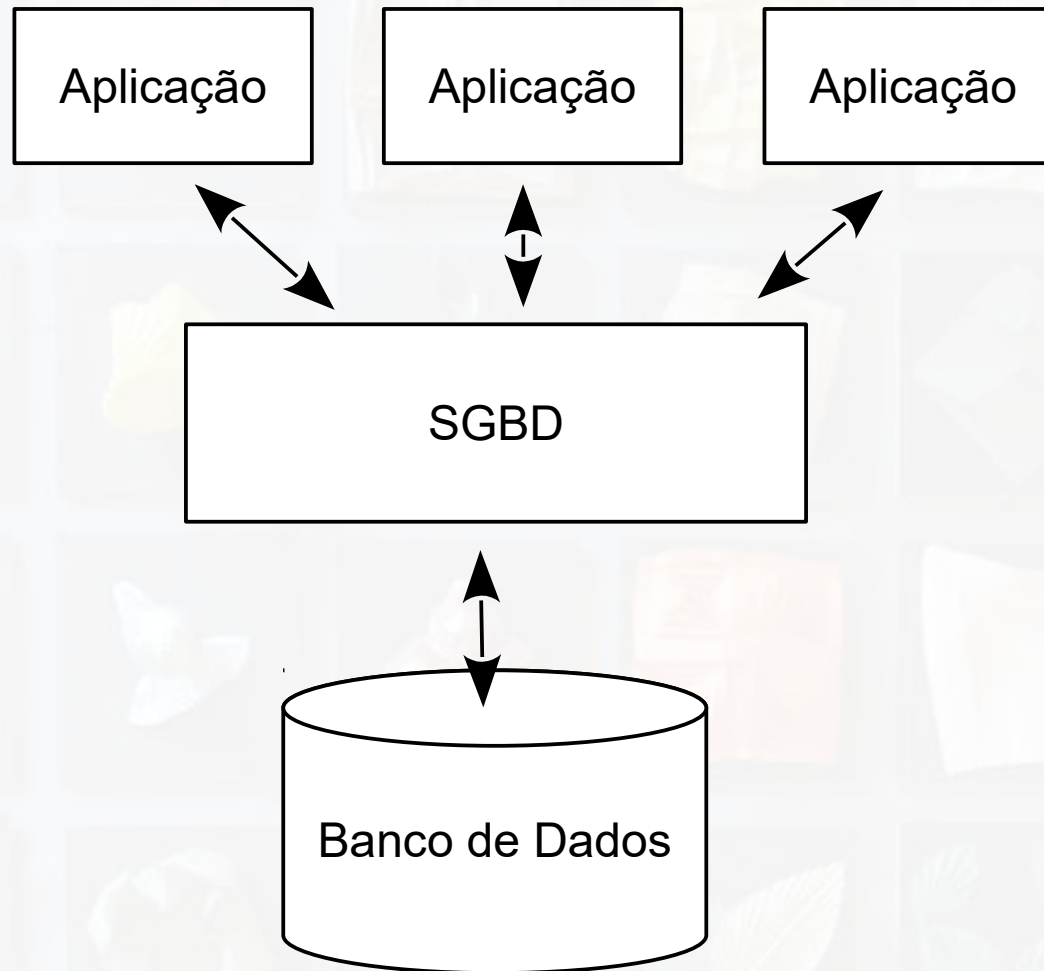
Cadastro de
Livros

Banco de Dados da Biblioteca

empréstimo



Arquitetura



Arquitetura Perspectivas

- (A) Componentes do SGBD
- (B) SGBD & rede

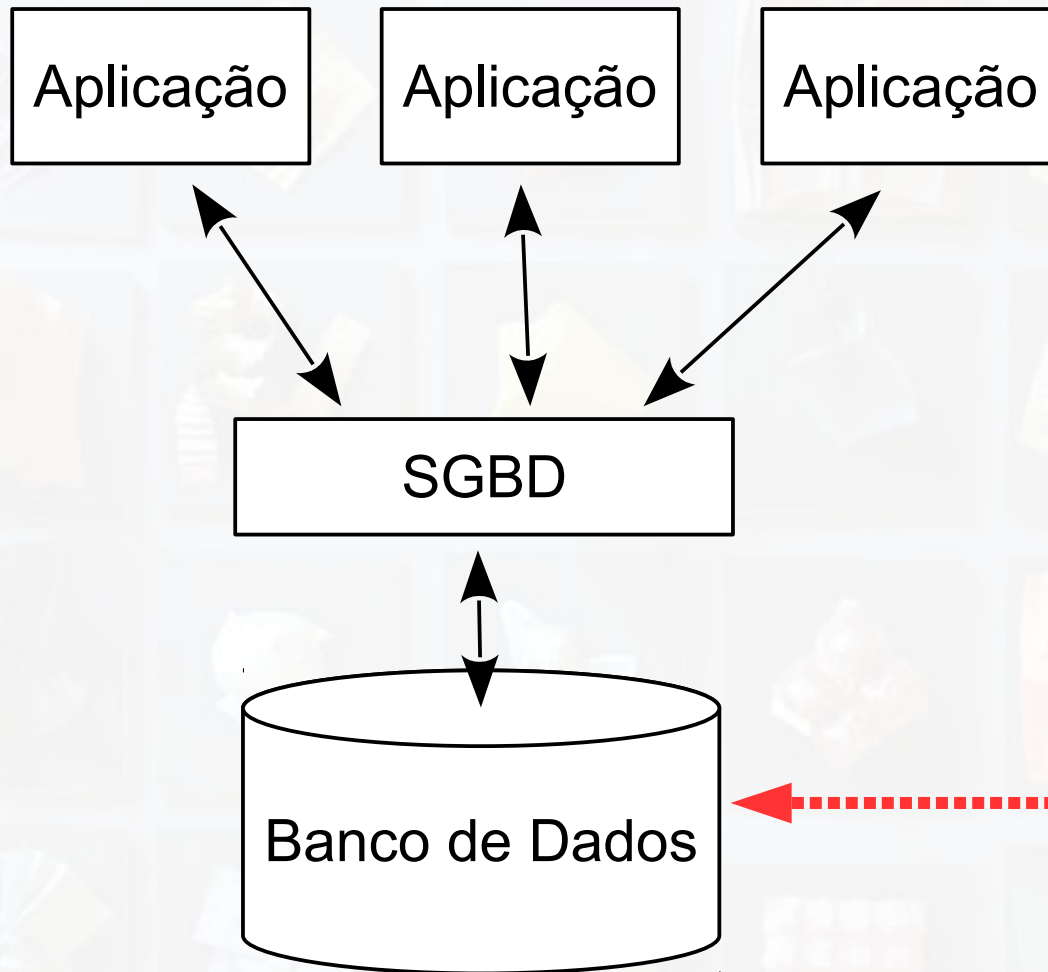


Arquitetura - Perspectivas

(A) Componentes do SGBD

Arquitetura

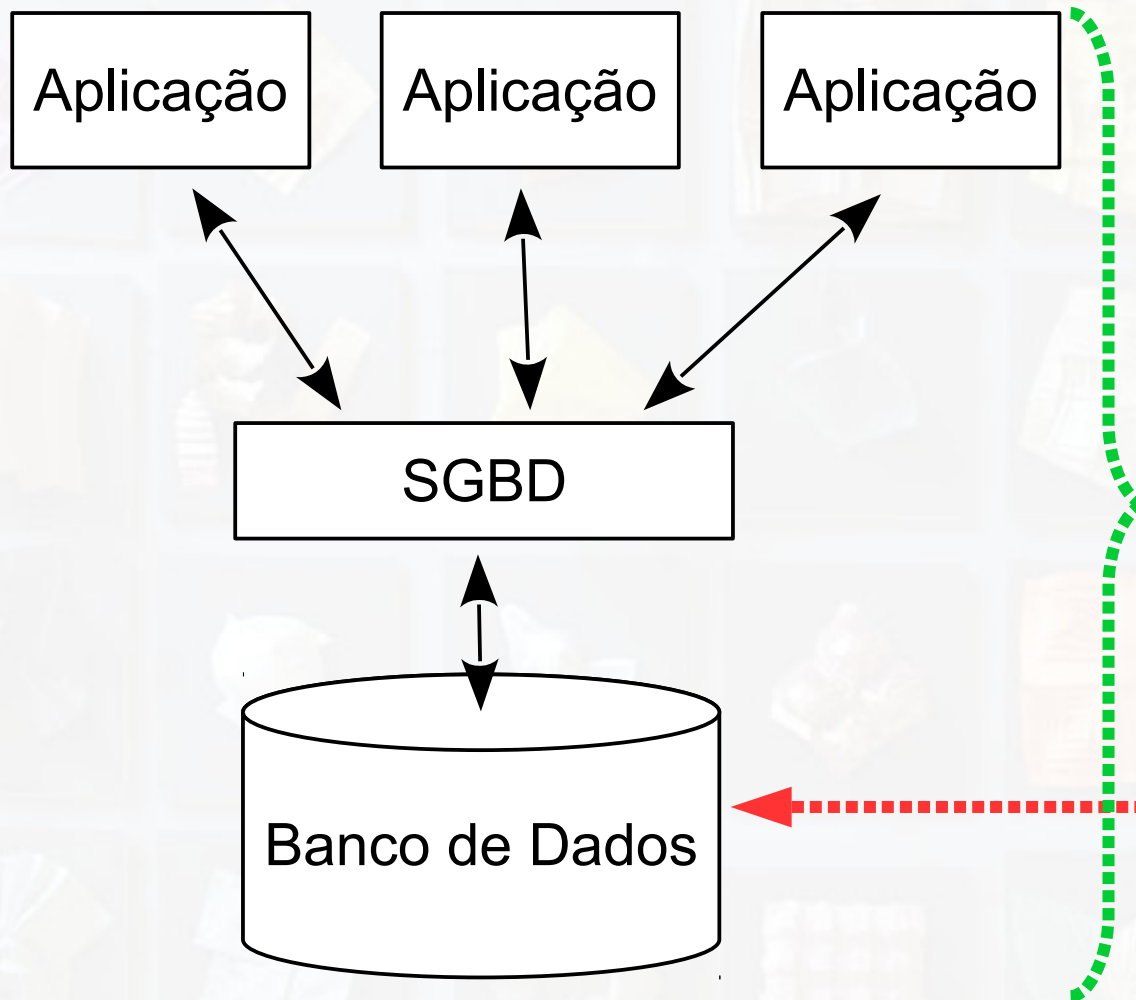
Questões



(1) Como os dados são armazenados aqui?
[esquema/instâncias]

Arquitetura

Questões

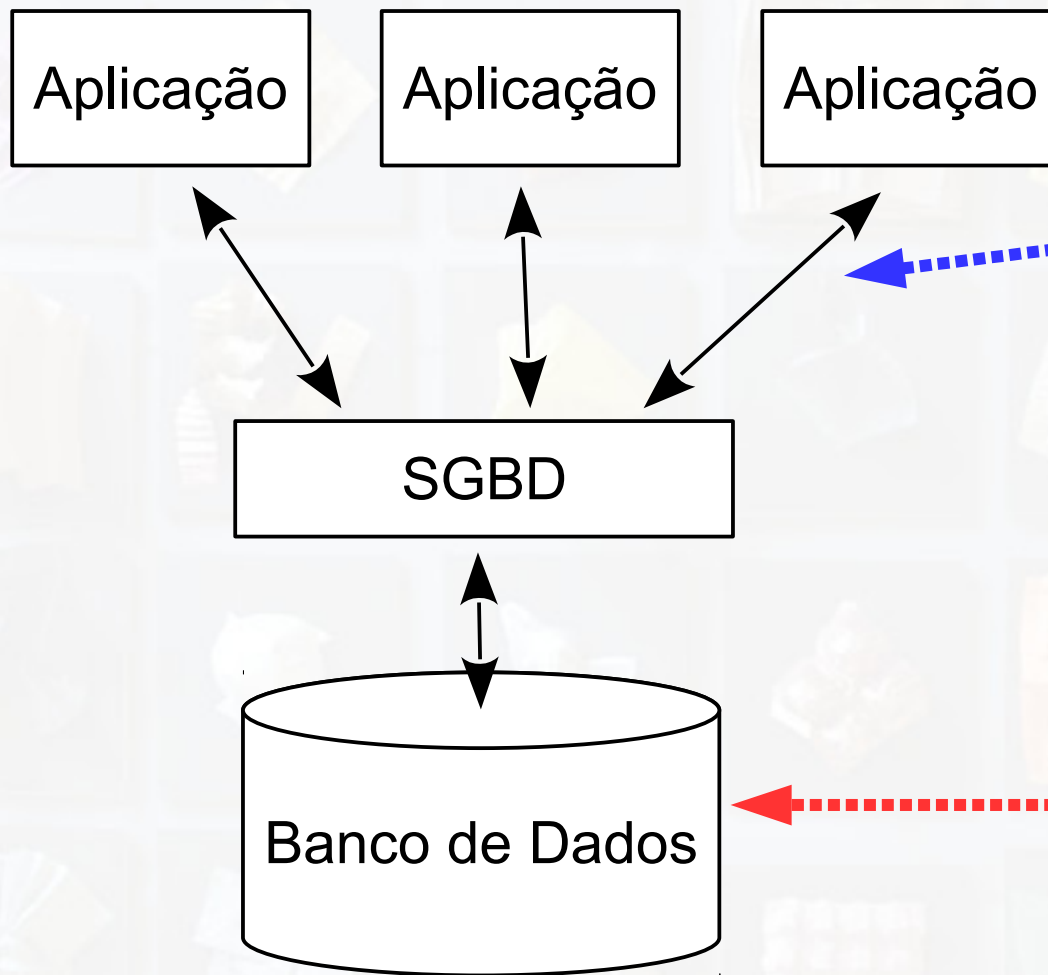


(2) Como estes dados são definidos?
[modelagem]

(1) Como os dados são armazenados aqui?
[esquema/instâncias]

Arquitetura

Questões

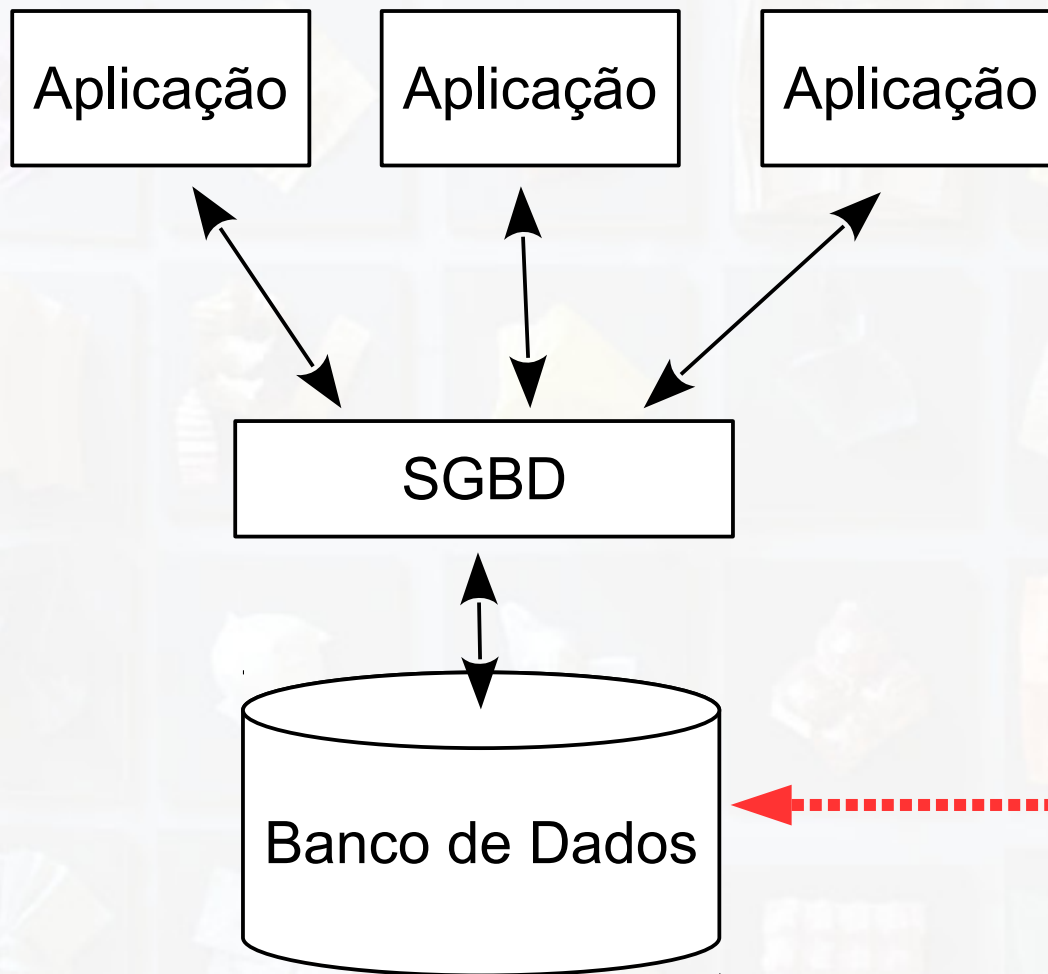


(3) Como acontece esta comunicação?
[linguagem de manipulação de dados]

(2) Como estes dados são definidos?
[modelagem]

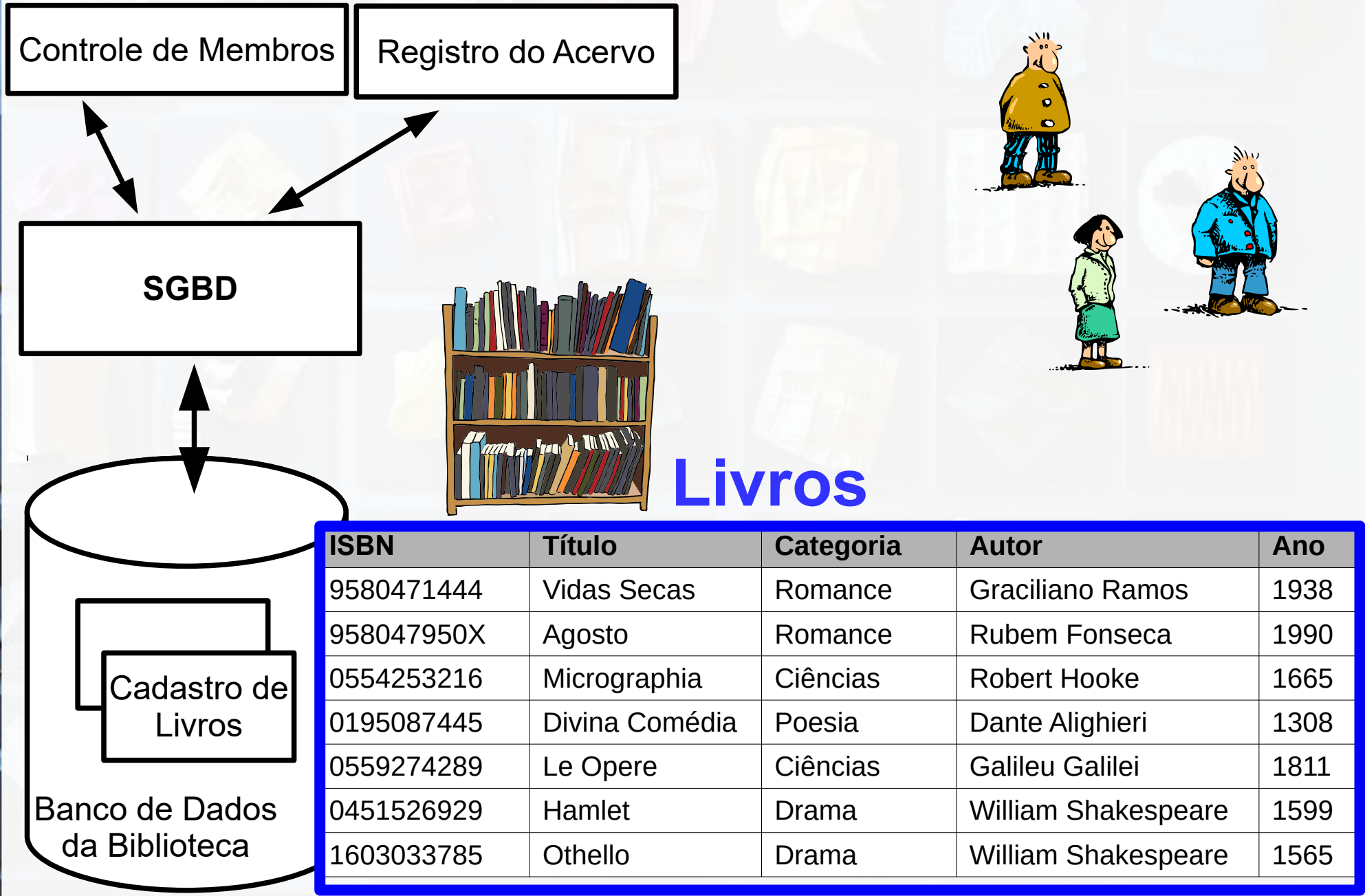
(1) Como os dados são armazenados aqui?
[esquema/instâncias]

(1) Esquema/Instâncias



Como os dados são armazenados aqui?

Cadastro de Livros - Biblioteca



ISBN	Título	Categoria	Autor	Ano
9580471444	Vidas Secas	Romance	Graciliano Ramos	1938
958047950X	Agosto	Romance	Rubem Fonseca	1990
0554253216	Micrographia	Ciências	Robert Hooke	1665
0195087445	Divina Comédia	Poesia	Dante Alighieri	1308
0559274289	Le Opere	Ciências	Galileu Galilei	1811
0451526929	Hamlet	Drama	William Shakespeare	1599
1603033785	Othello	Drama	William Shakespeare	1565

Cadastro de Livros - Biblioteca



Livros

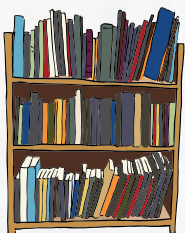
ISBN	Título	Categoria	Autor	Ano
9580471444	Vidas Secas	Romance	Graciliano Ramos	1938
958047950X	Agosto	Romance	Rubem Fonseca	1990
0554253216	Micrographia	Ciências	Robert Hooke	1665
0195087445	Divina Comédia	Poesia	Dante Alighieri	1308
0559274289	Le Opere	Ciências	Galileu Galilei	1811
0451526929	Hamlet	Drama	William Shakespeare	1599
1603033785	Othello	Drama	William Shakespeare	1565

Esquema

Instâncias

Cadastro de Livros - Biblioteca

- Esquema - descrição da tabela
- Instâncias - dados da tabela seguindo um esquema (estado do banco)



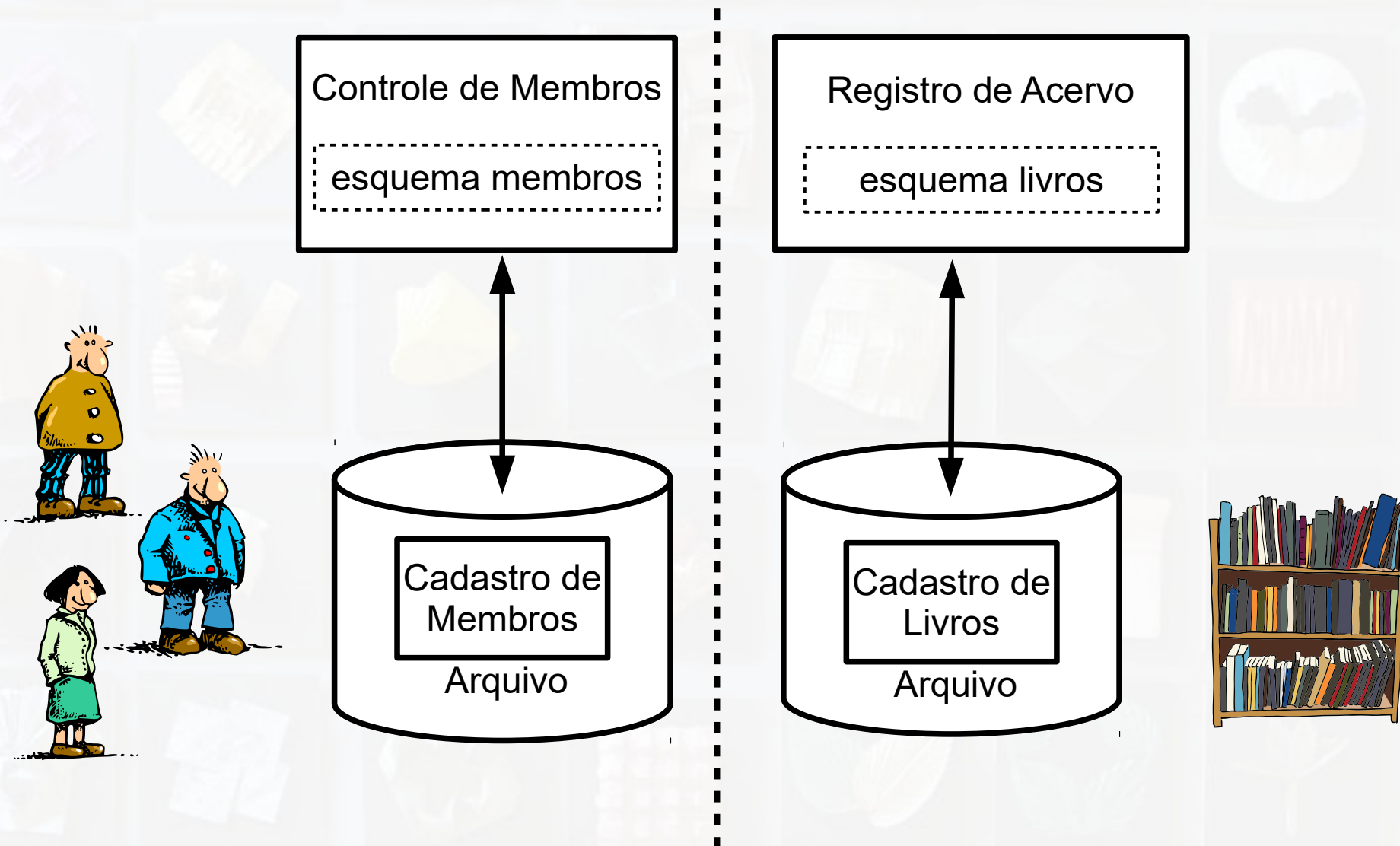
Livros

ISBN	Título	Categoria	Autor	Ano
9580471444	Vidas Secas	Romance	Graciliano Ramos	1938
958047950X	Agosto	Romance	Rubem Fonseca	1990
0554253216	Micrographia	Ciências	Robert Hooke	1665
0195087445	Divina Comédia	Poesia	Dante Alighieri	1308
0559274289	Le Opere	Ciências	Galileu Galilei	1811
0451526929	Hamlet	Drama	William Shakespeare	1599
1603033785	Othello	Drama	William Shakespeare	1565

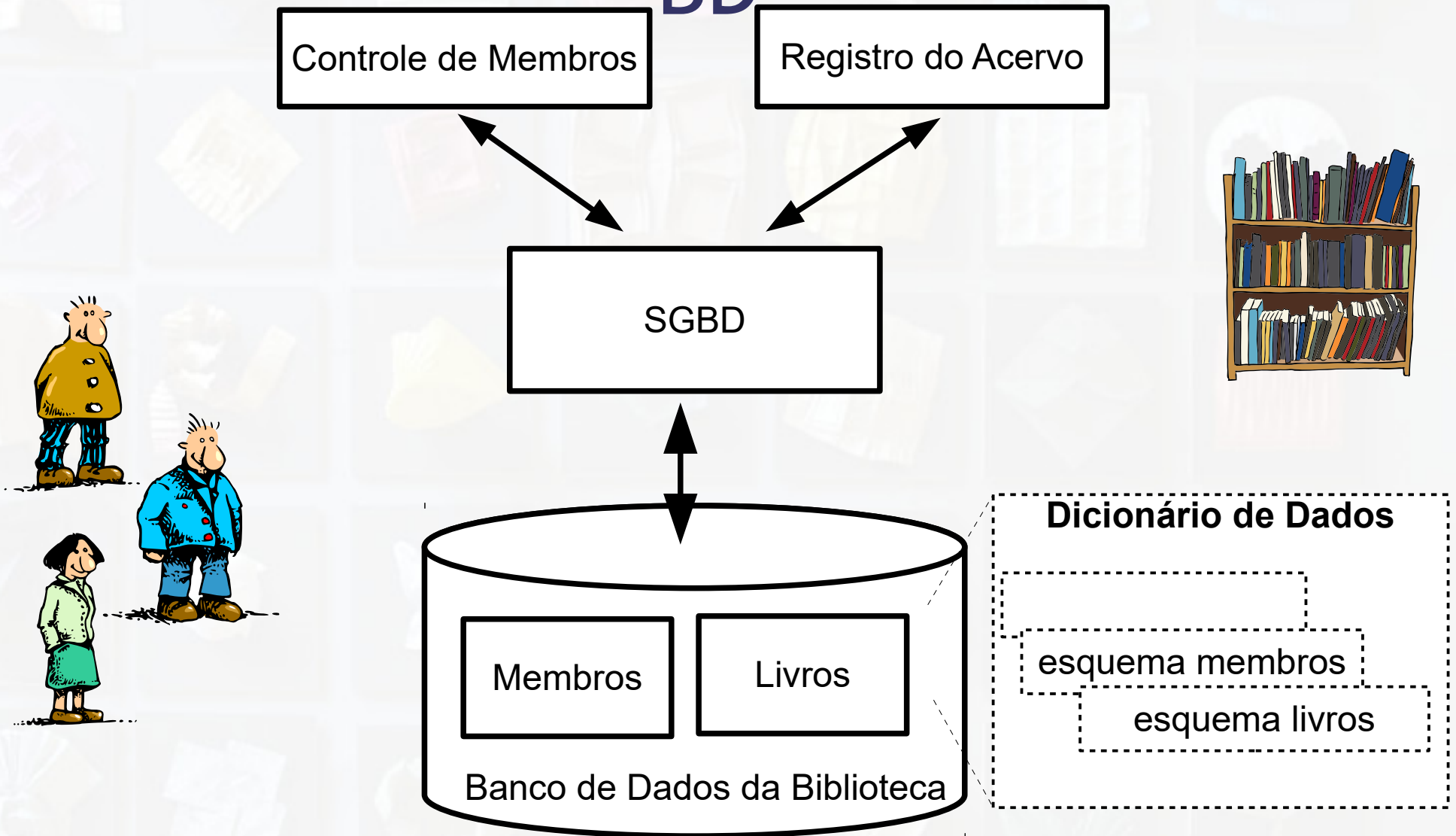
Esquema

Instâncias

Passado: Esquema na Aplicação



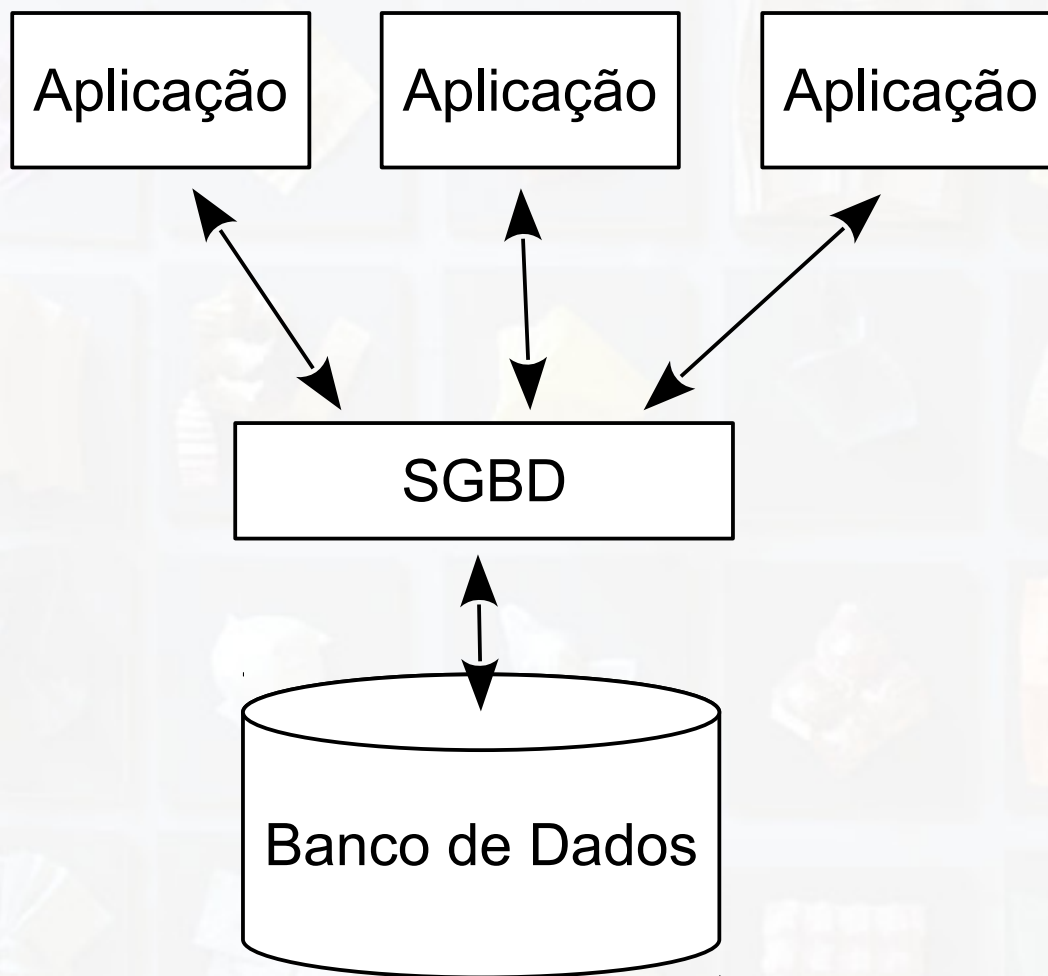
Presente: Dicionário de Dados no BD



Esquema e Modelo

- Esquema de um Banco de Dados é produzido a partir de um modelo.

(2) Modelagem



Como estes dados são definidos?

Universo de Discurso ou Mini-mundo

- “Um banco de dados representa algum aspecto do mundo real, às vezes chamado de **mini-mundo** ou de **universo de discurso** (UoD - Universe of Discourse).”

(Elmasri & Navathe, 2011)

Abstração



fez
empréstimo



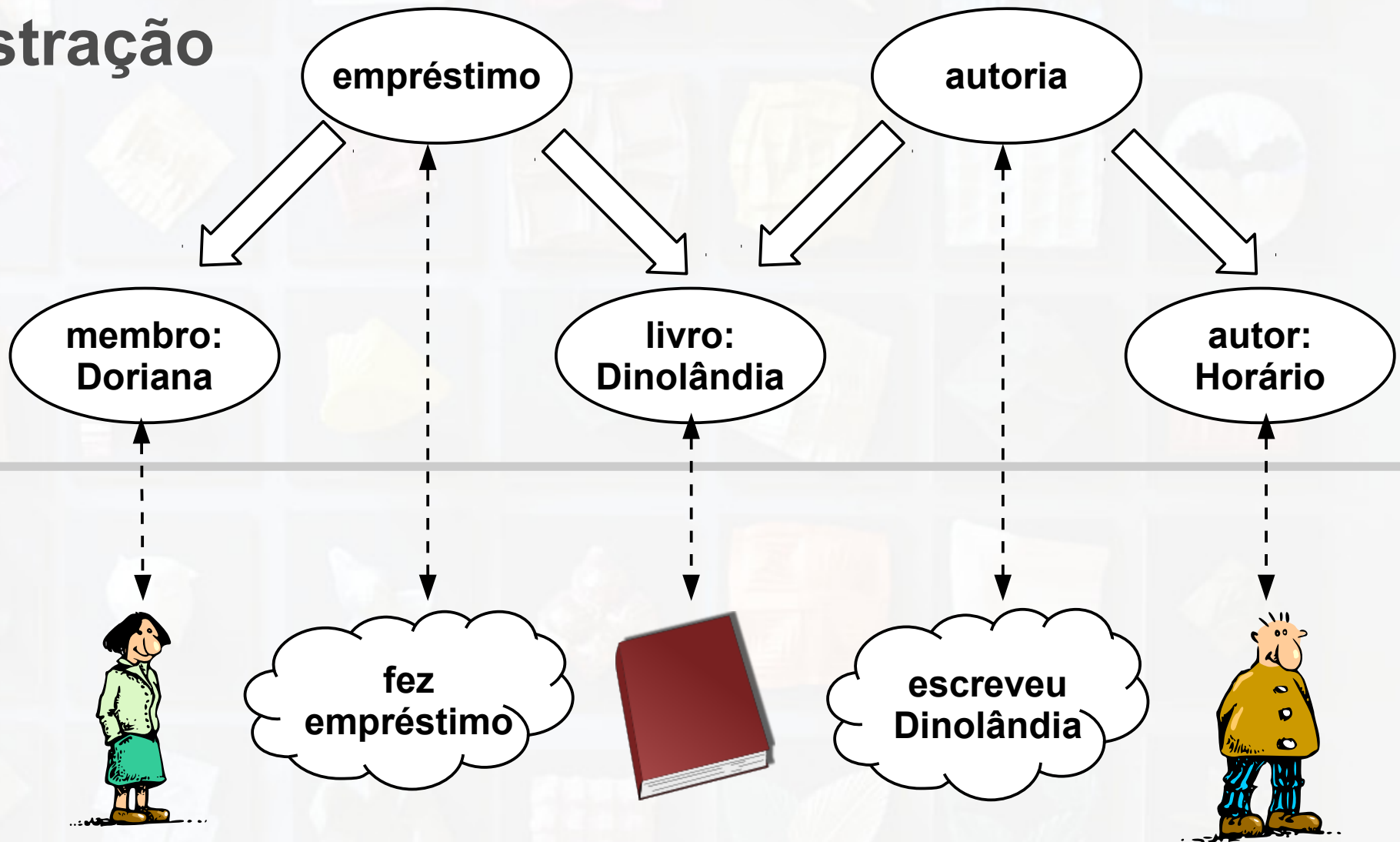
escreveu
Dinolândia



universo de discurso

Abstração

abstração



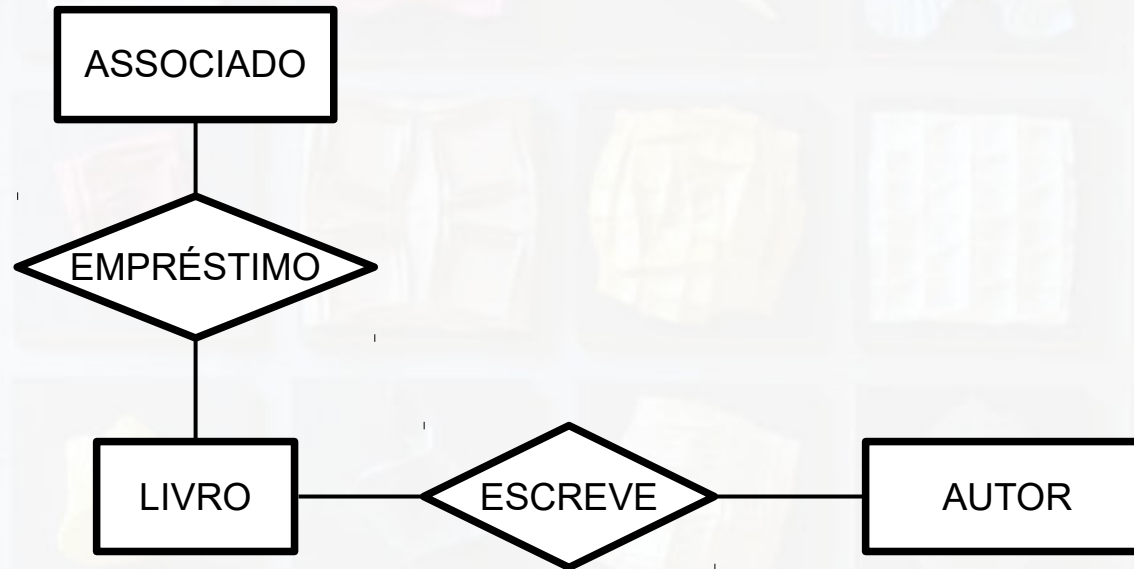
universo de discurso

Modelo

- Coleção de ferramentas conceituais
- Descrevem: dados, relações, semântica e restrições

Abstração

modelo



universo de discurso

Projeto de um BD

Usuários
Finais



Visão Externa 1

Visão Externa 2

Visão Externa n

requisitos

requisitos

requisitos

Modelo Conceitual

esquema conceitual

Independente de SGBD

Dependente de SGBD

Modelo Lógico

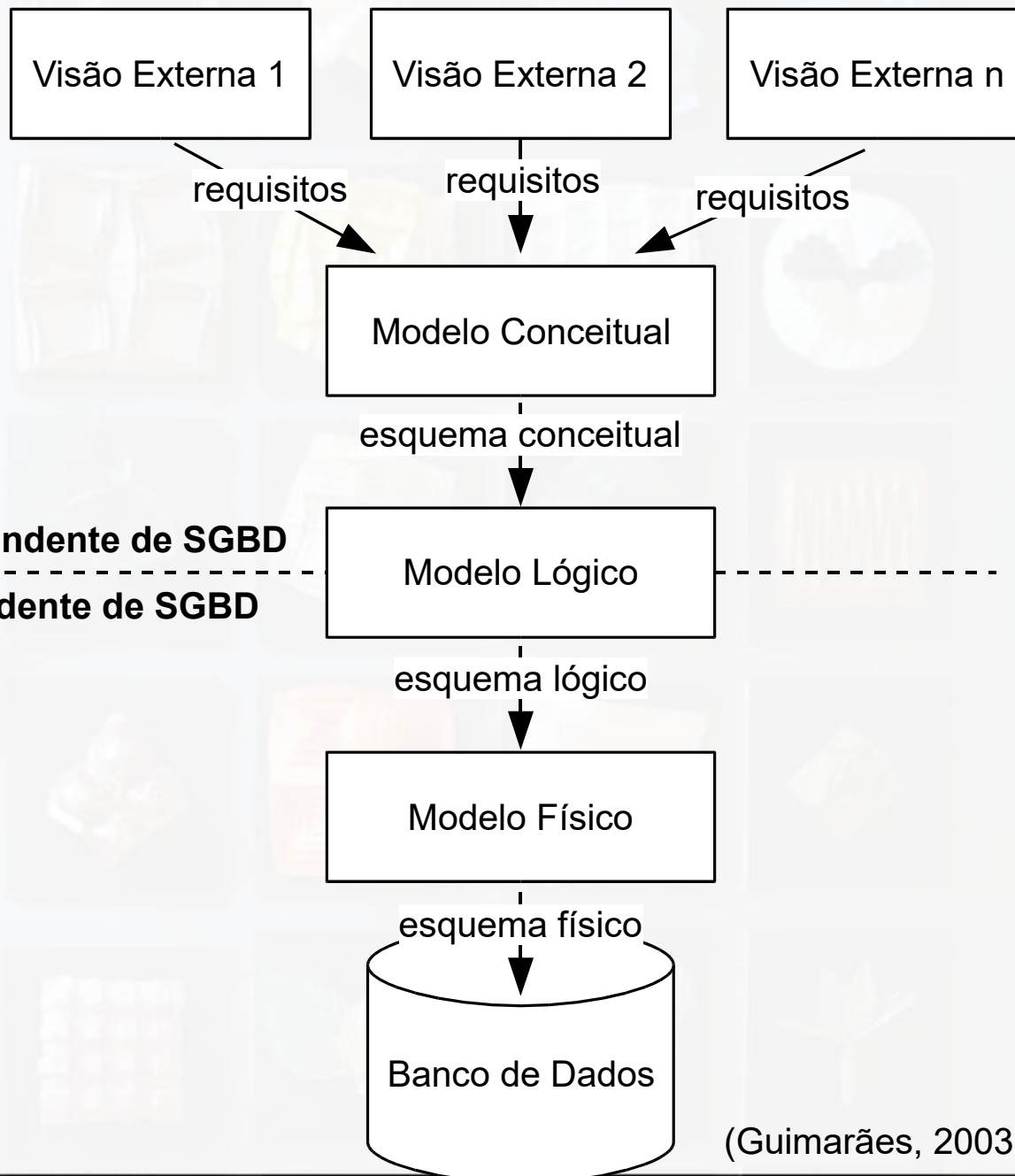
esquema lógico

Modelo Físico

esquema físico

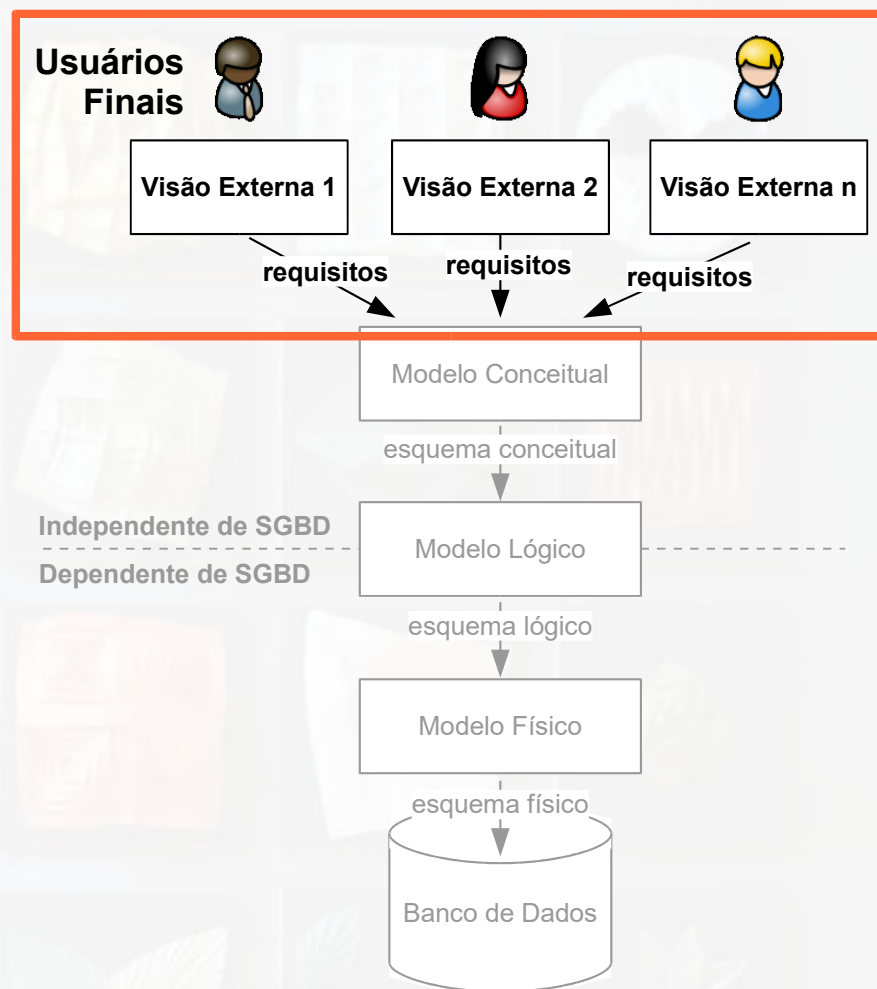
Banco de Dados

(Guimarães, 2003)



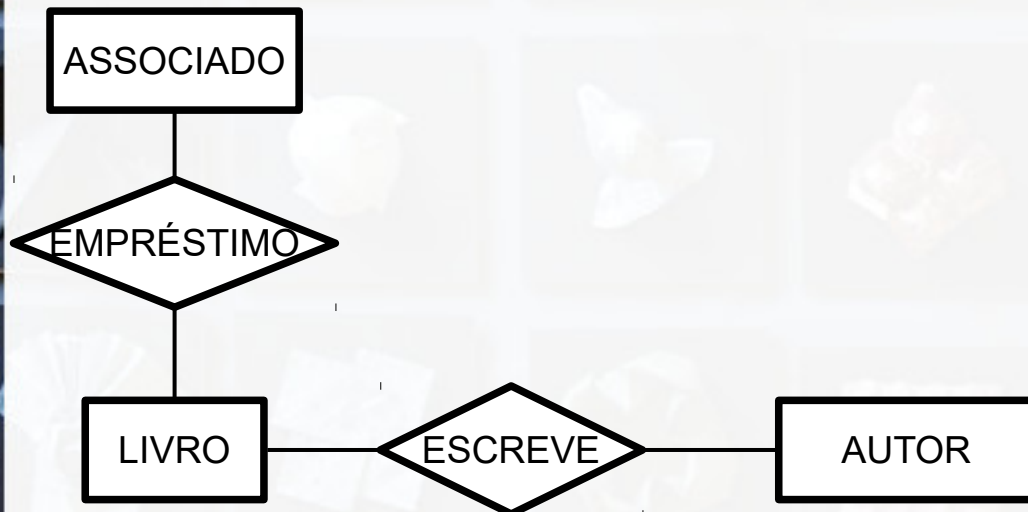
Visão Externa

- Guiada pelos requisitos dos usuários
- Usualmente representada em documentos textuais
- Visão - recorte do esquema

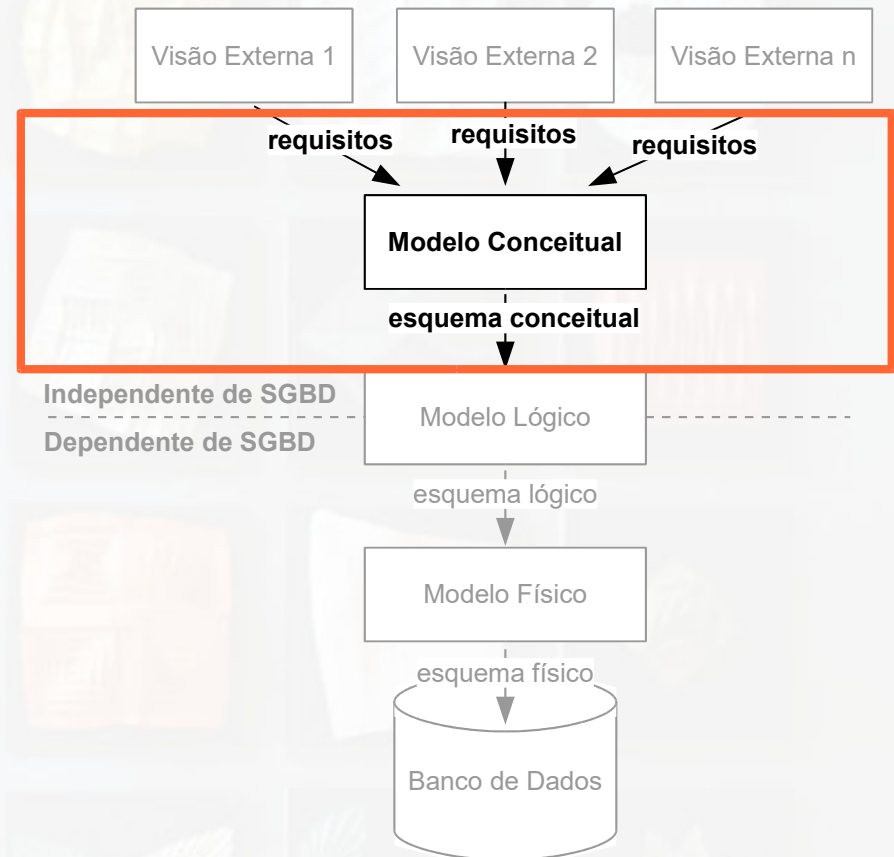


Modelo Conceitual

- Próximo ao modo como os usuários percebem os dados
- Oculta detalhes de armazenamento físico

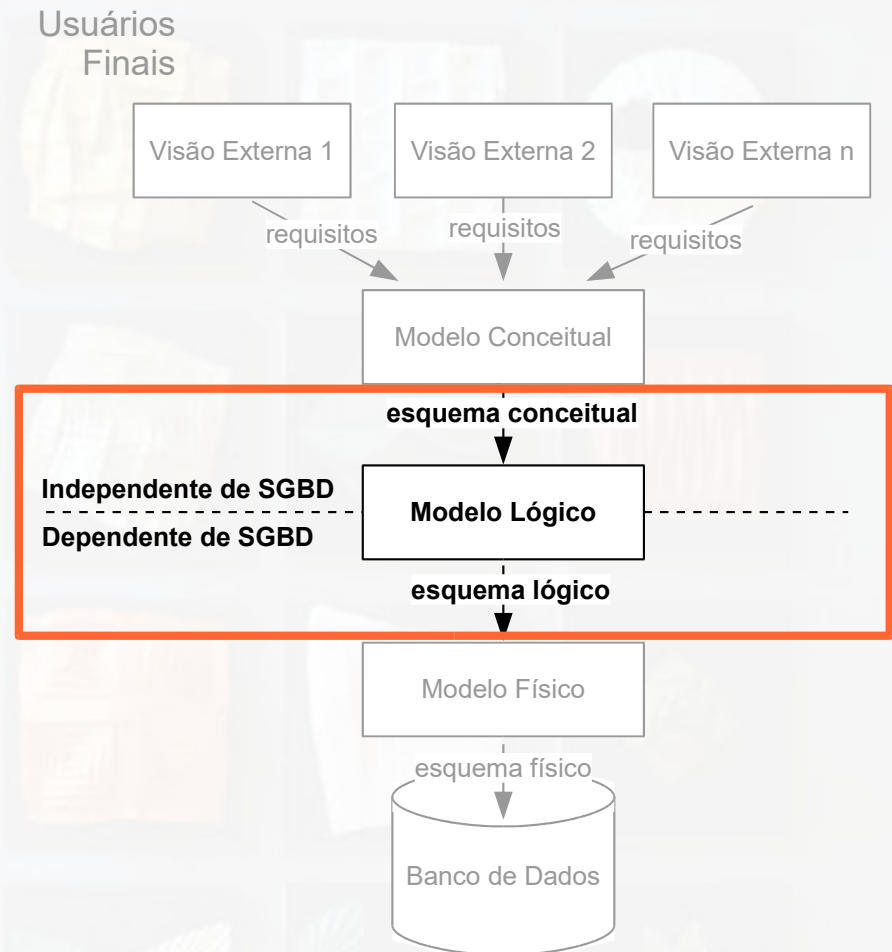


Usuários
Finais



Modelo Lógico

- Associado a um “modelo de dados de implementação” (Elmasri, 2005)



Modelo Lógico Relacional



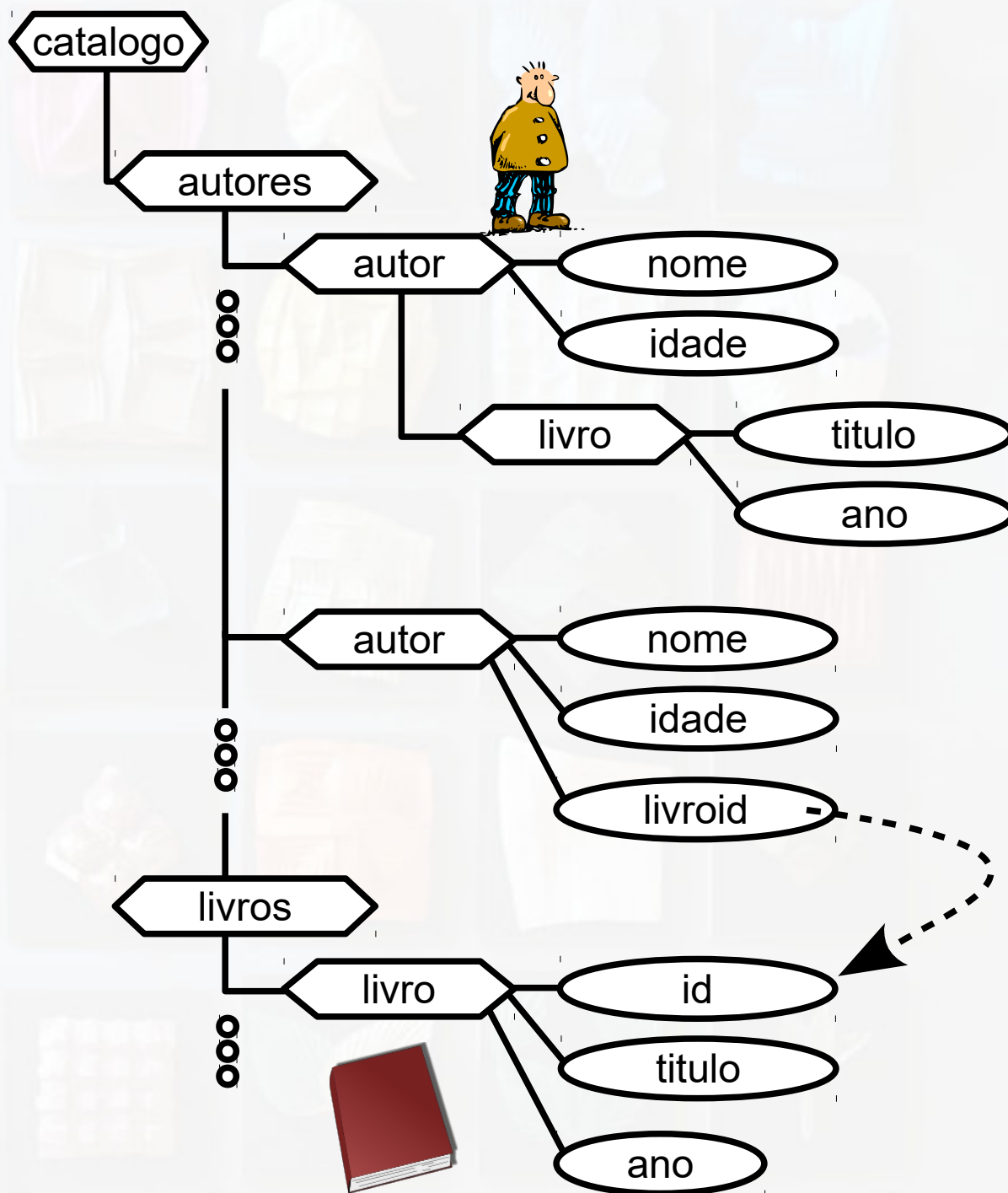
Livros

ISBN	Título	Categoria	Autor	Ano
9580471444	Vidas Secas	Romance	Graciliano Ramos	1938
958047950X	Agosto	Romance	Rubem Fonseca	1990
0554253216	Micrographia	Ciências	Robert Hooke	1665
0195087445	Divina Comédia	Poesia	Dante Alighieri	1308
0559274289	Le Opere	Ciências	Galileu Galilei	1811
0451526929	Hamlet	Drama	William Shakespeare	1599
1603033785	Othello	Drama	William Shakespeare	1565

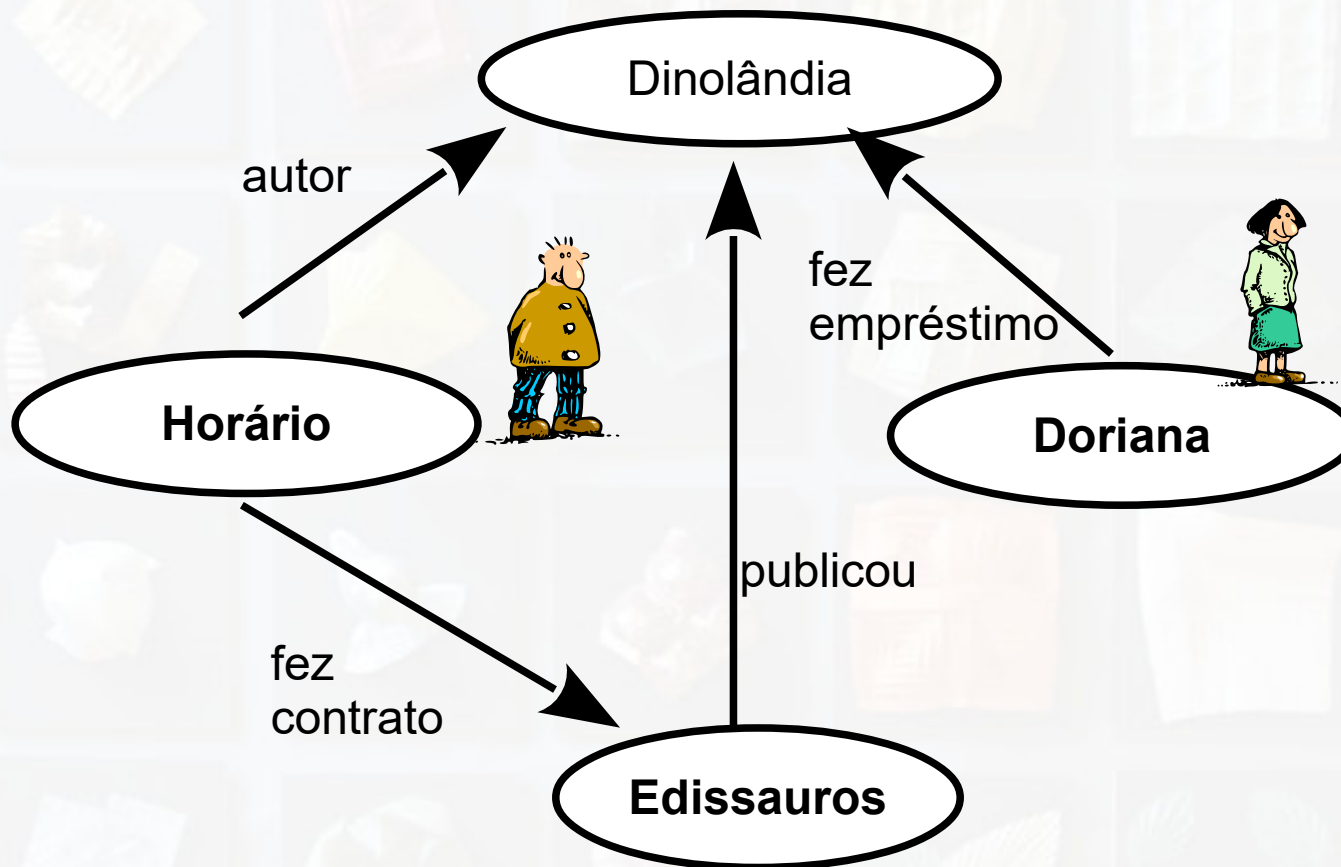
Modelo Lógico Hierárquico

■ Exemplo

- XML
- JSON

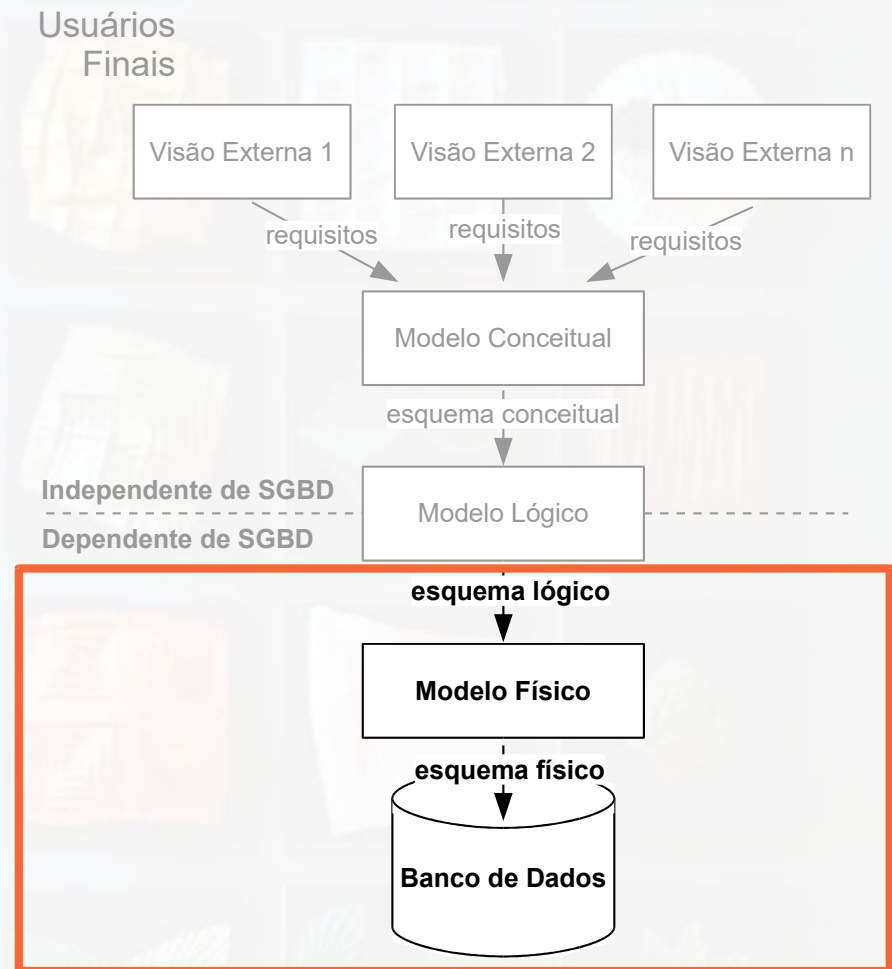


Modelo Lógico de Grafos

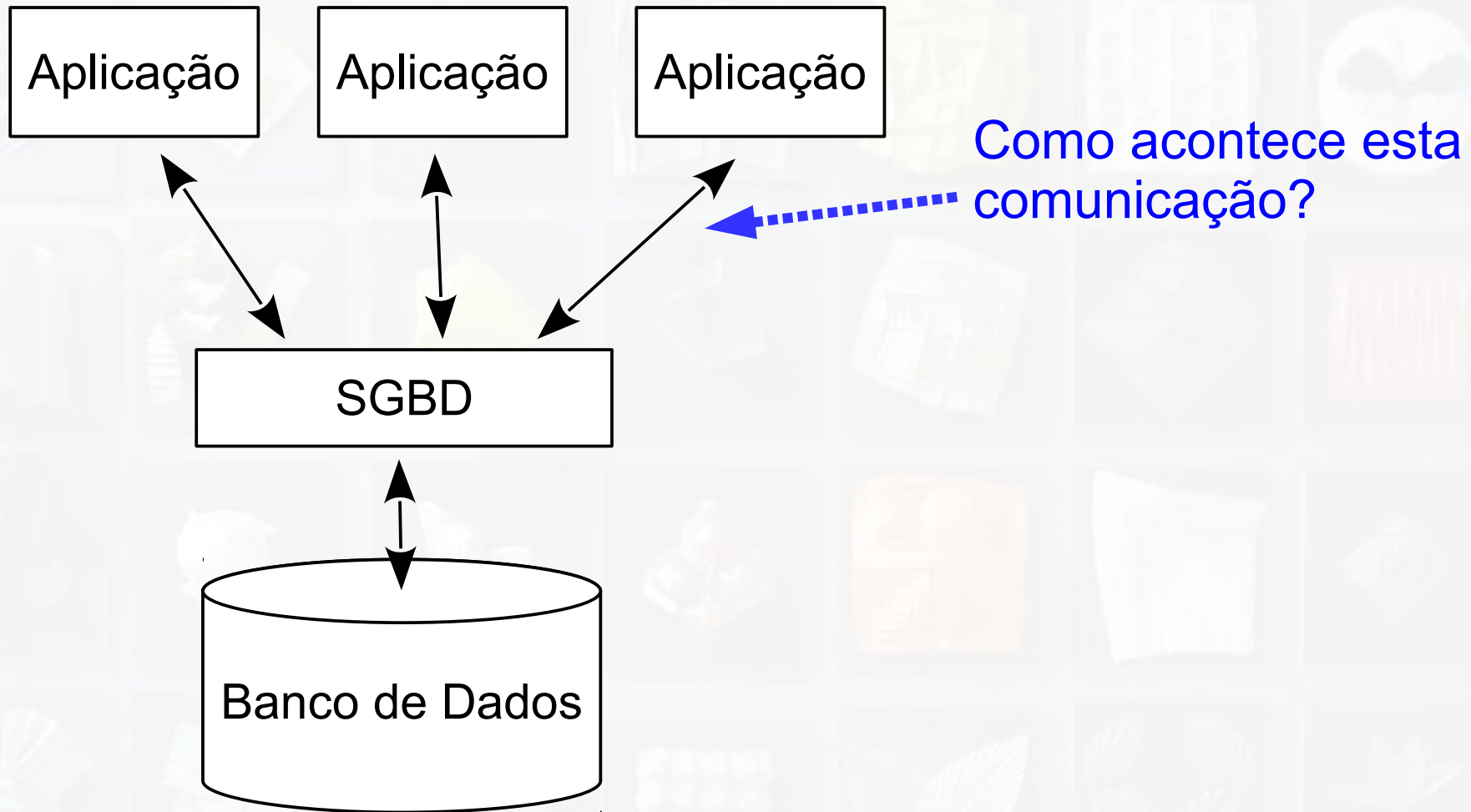


Modelo Físico

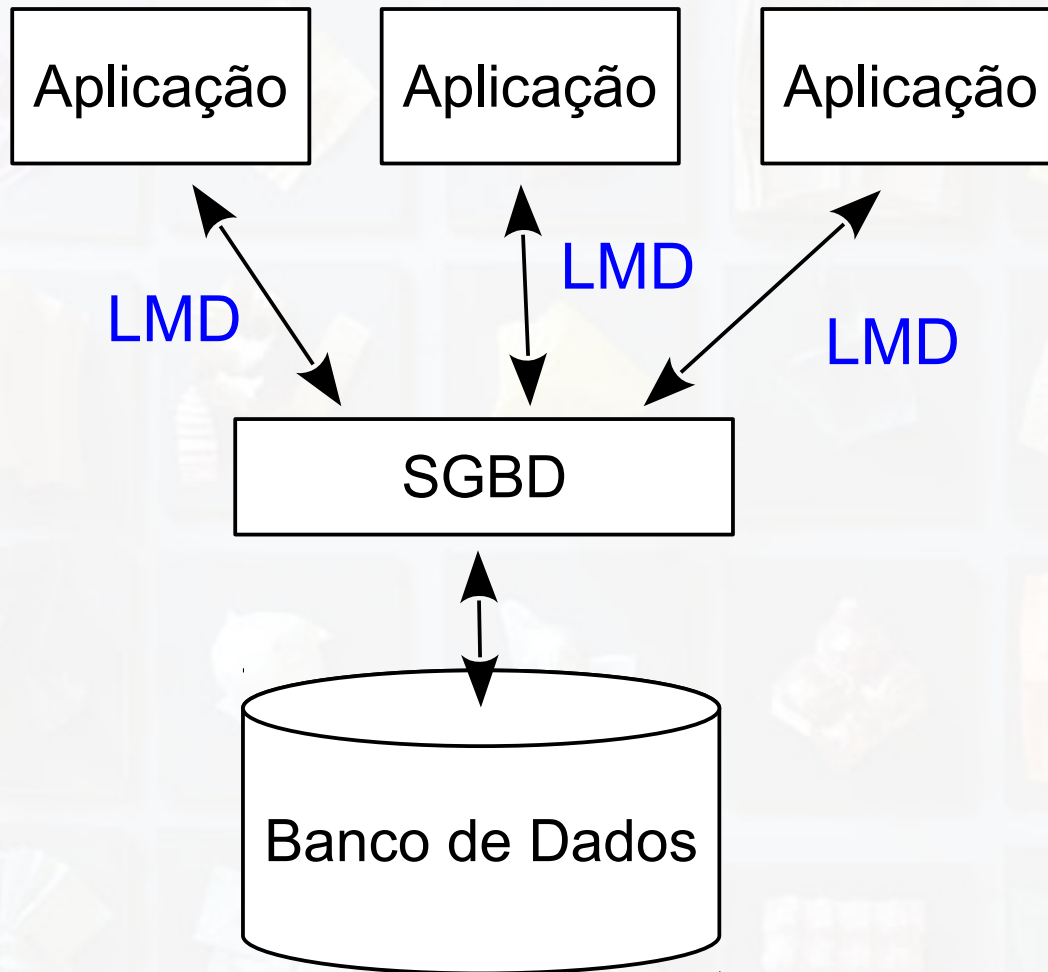
- Descreve a estrutura de armazenamento físico



(3) Linguagem de Manipulação de Dados



Linguagem de Manipulação de Dados (LMD)



Execução de Consulta

Passos Típicos (Elmasri, 2010)

Consulta em linguagem de alto nível

Análise Léxica,
Análise Sintática
e Validação

Forma intermediária de consulta

Otimizador
de Consulta

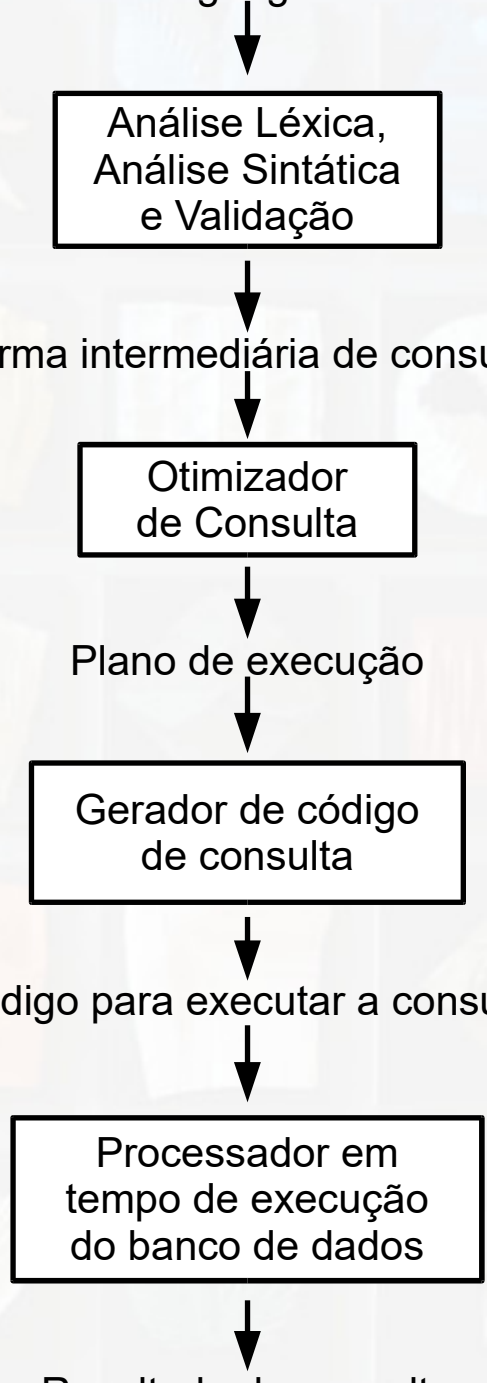
Plano de execução

Gerador de código
de consulta

Código para executar a consulta

Processador em
tempo de execução
do banco de dados

Resultado da consulta



Análise e Validação

- **Análise e Validação**
 - Análise léxica
 - Análise sintática
 - Validação
- **Representações internas:**
 - árvore de consulta
 - grafo de consulta

Consulta em linguagem de alto nível

Análise Léxica,
Análise Sintática
e Validação

Forma intermediária de consulta

Otimizador
de Consulta

Plano de execução

Gerador de código
de consulta

Código para executar a consulta

Processador em
tempo de execução
do banco de dados

Resultado da consulta

Estratégia de Execução

- Consulta possui muitas estratégias de execução possíveis
- Planejamento da Estratégia de Execução
 - Otimização → processo de escolha da estratégia adequada (razoavelmente eficiente)

Consulta em linguagem de alto nível

Análise Léxica,
Análise Sintática
e Validação

Forma intermediária de consulta

Otimizador
de Consulta

Plano de execução

Gerador de código
de consulta

Código para executar a consulta

Processador em
tempo de execução
do banco de dados

Resultado da consulta

Código da Consulta

- Pode ser:
 - Executado diretamente
 - modo interpretado
 - Armazenado e executado quando necessário
 - modo compilado

Consulta em linguagem de alto nível

Análise Léxica,
Análise Sintática
e Validação

Forma intermediária de consulta

Otimizador
de Consulta

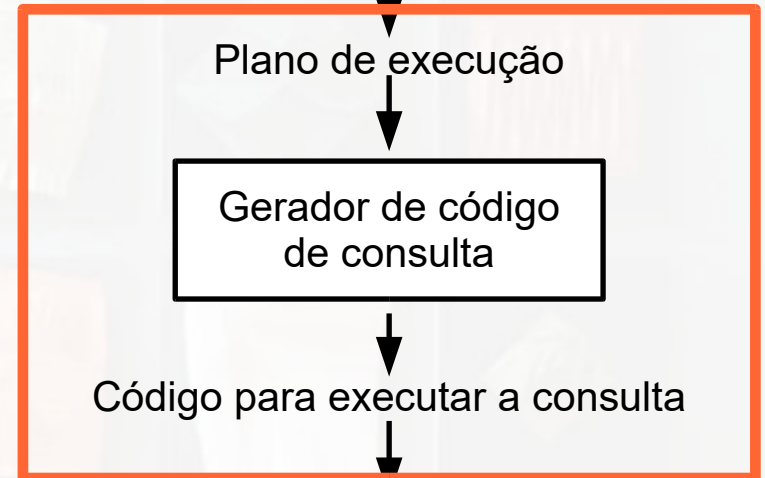
Plano de execução

Gerador de código
de consulta

Código para executar a consulta

Processador em
tempo de execução
do banco de dados

Resultado da consulta



Execução do Código

- Processador executa código da consulta
- Produz resultado da execução

Consulta em linguagem de alto nível

Análise Léxica,
Análise Sintática
e Validação

Forma intermediária de consulta

Otimizador
de Consulta

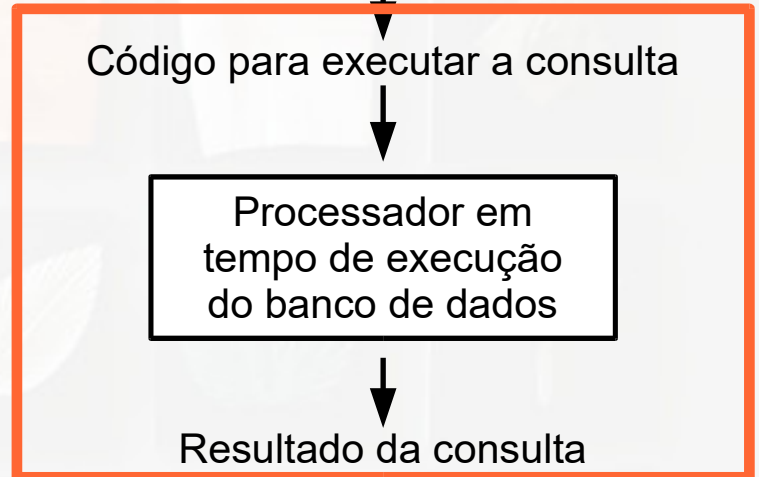
Plano de execução

Gerador de código
de consulta

Código para executar a consulta

Processador em
tempo de execução
do banco de dados

Resultado da consulta



Consultas Declarativas

- “O quê” ao invés de “Como”
- Otimização de consulta
 - Solução razoavelmente eficiente (Elmasri, 2011)
 - Solução ótima pode ser muito custosa

Linguagem de Manipulação de Dados (LMD)

- Para manipulação e recuperação de dados
- LMDs em BD:
 - Fundamentação formal
 - Subsidiária otimização
- LMDs <> linguagens de programação
 - não se espera que sejam “Turing completas”.
 - não pensadas para cálculos complexos.
 - suportam acessos simples e eficientes a extensos conjuntos de dados


(Ramakrishnan, 2003)

LMDs - Operações Exemplo

- Incluir registros sobre um paciente em um BD
- Recuperar quais as transferências feitas por um cliente no mês de agosto/2018
- Aumentar em 15% o salário de todos os funcionários de uma empresa
- Recuperar os cinco filmes mais assistidos no Netflix

Exemplo de Linguagem: SQL

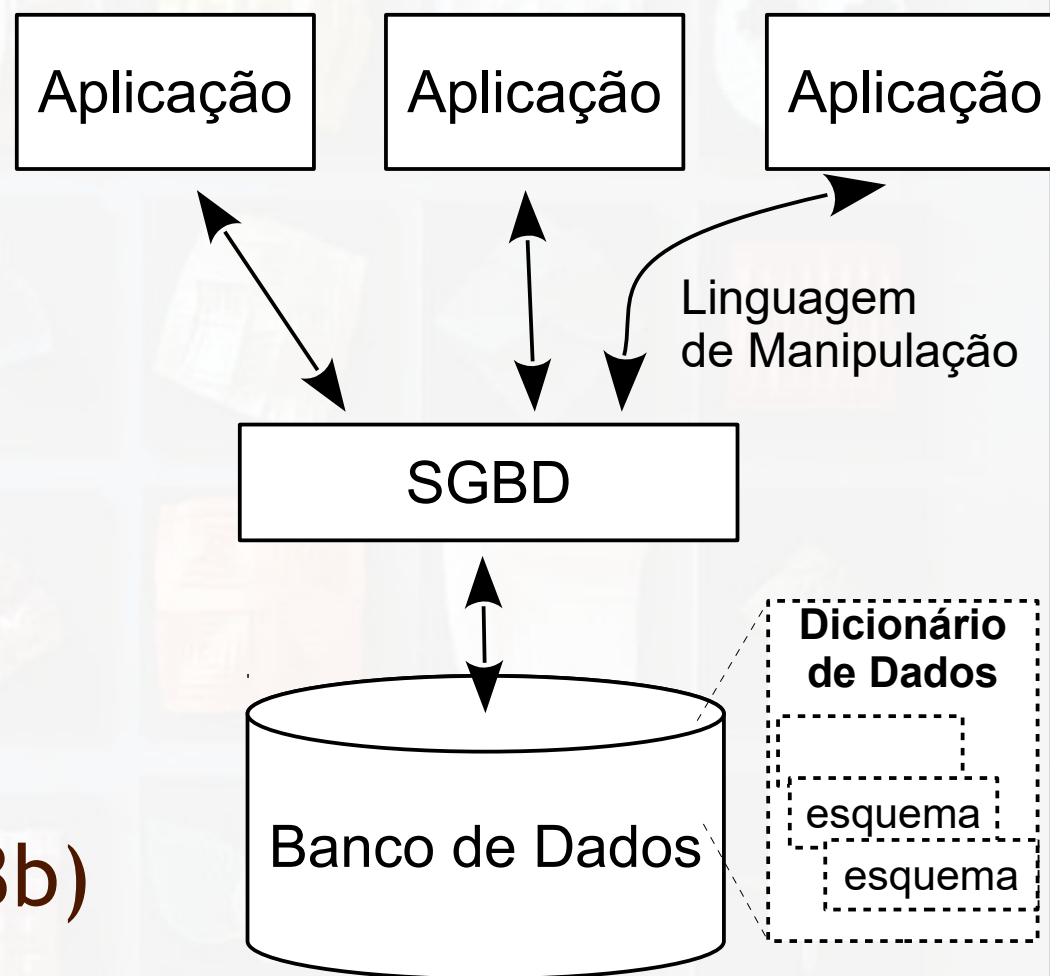
- SQL - Structured Query Language
- Criada pela IBM Research
- Exemplo: `SELECT Marca, Modelo FROM Taxi`



<u>Placa</u>	Marca	Modelo	AnoFab
DAE6534	Ford	Fiesta	1999
DKL4598	Wolkswagen	Gol	2001
DKL7878	Ford	Fiesta	2001
JDM8776	Wolkswagen	Santana	2002
JJM3692	Chevrolet	Corsa	1999

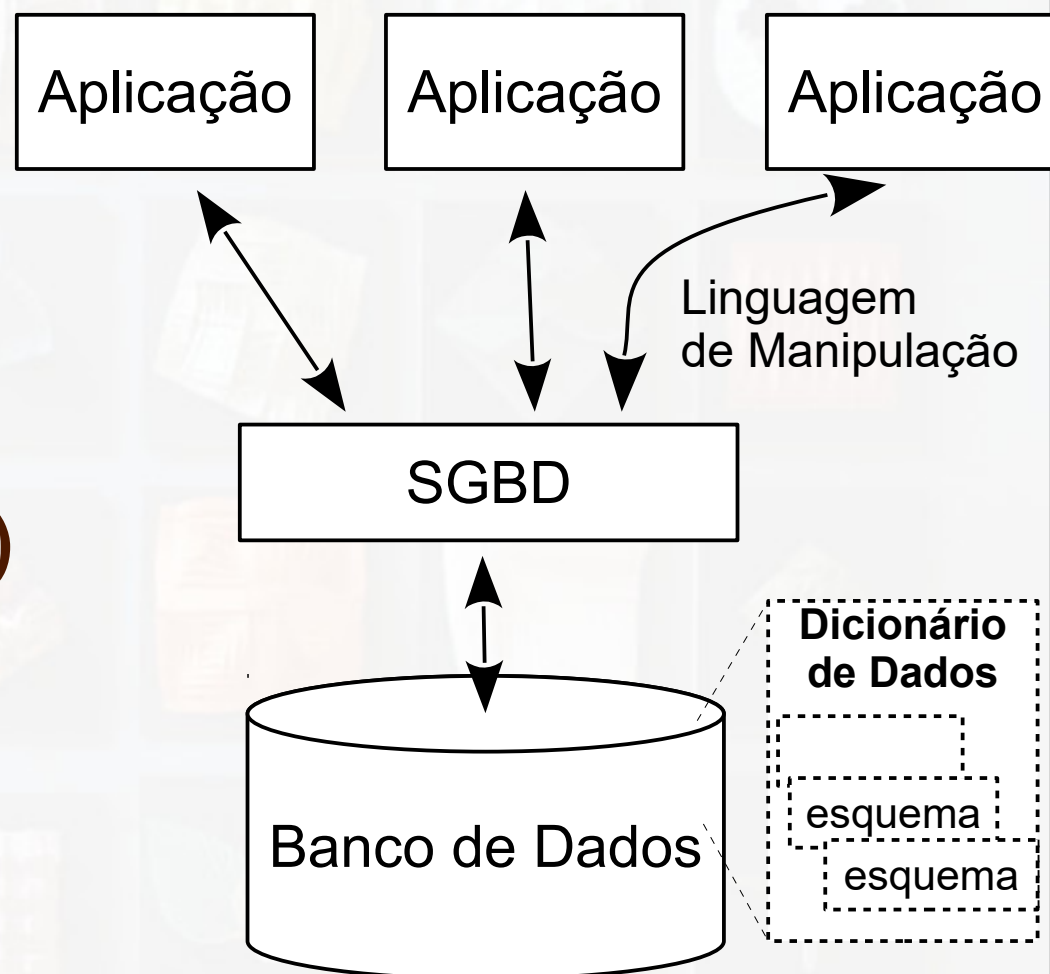
Vantagens de um SGBD

- Independência de dados
 - Acesso eficiente
 - Tempo reduzido no desenvolvimento de aplicações
 - Segurança e integridade de dados
- (Ramakrishnan, 2003b)



Vantagens de um SGBD

- Administração de dados uniforme
 - Acesso concorrente
 - Recuperação contra *crashes*
- (Ramakrishnan, 2003b)

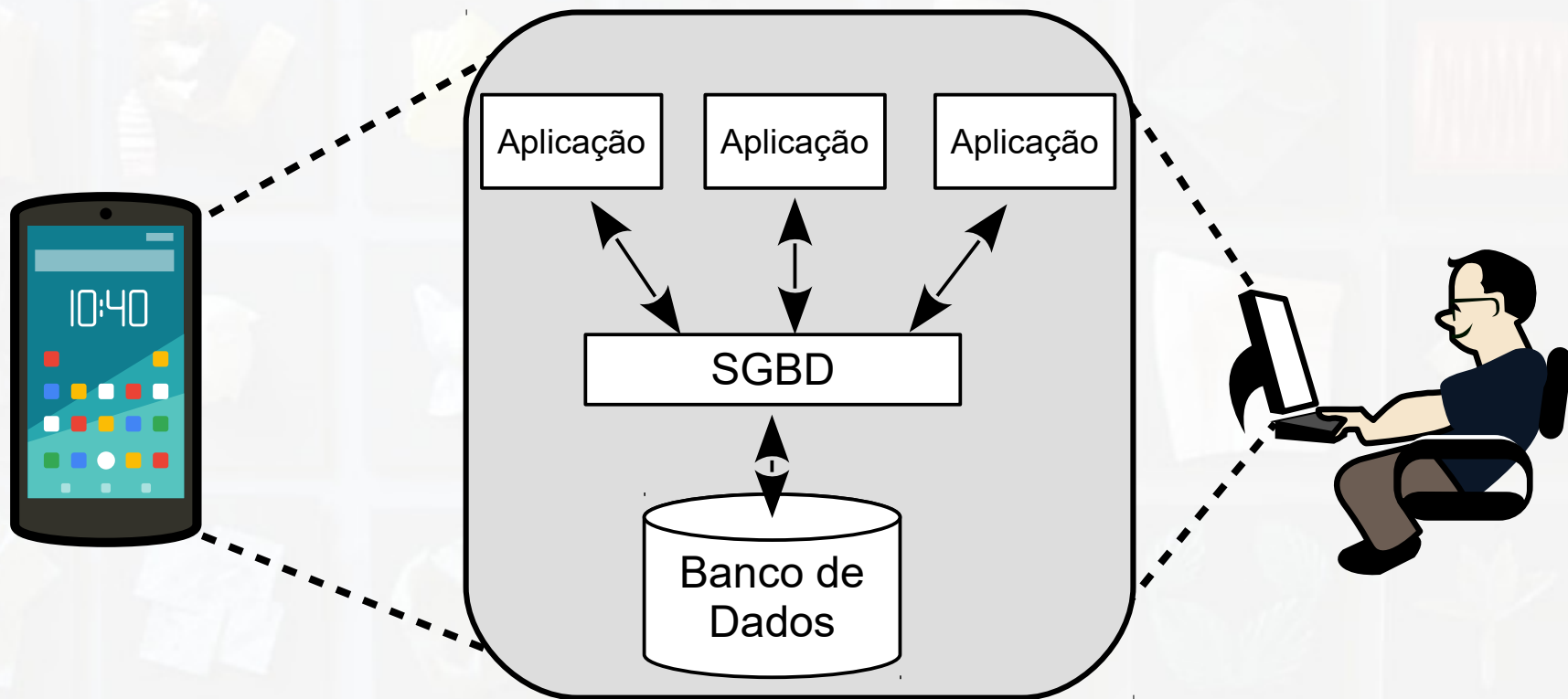




Arquitetura - Perspectivas (B) SGBD & Rede

Arquitetura Local

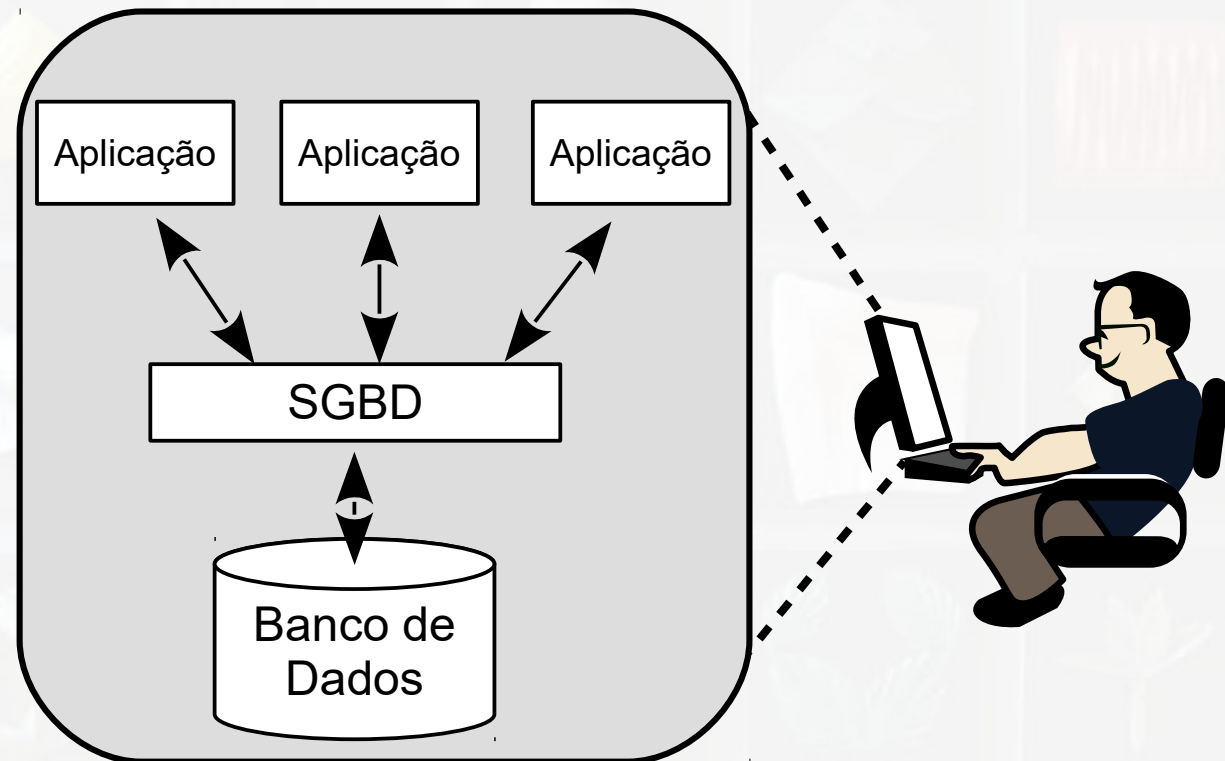
- SGBD executado no mesmo dispositivo da aplicação



Arquitetura Local

Exemplos Desktop

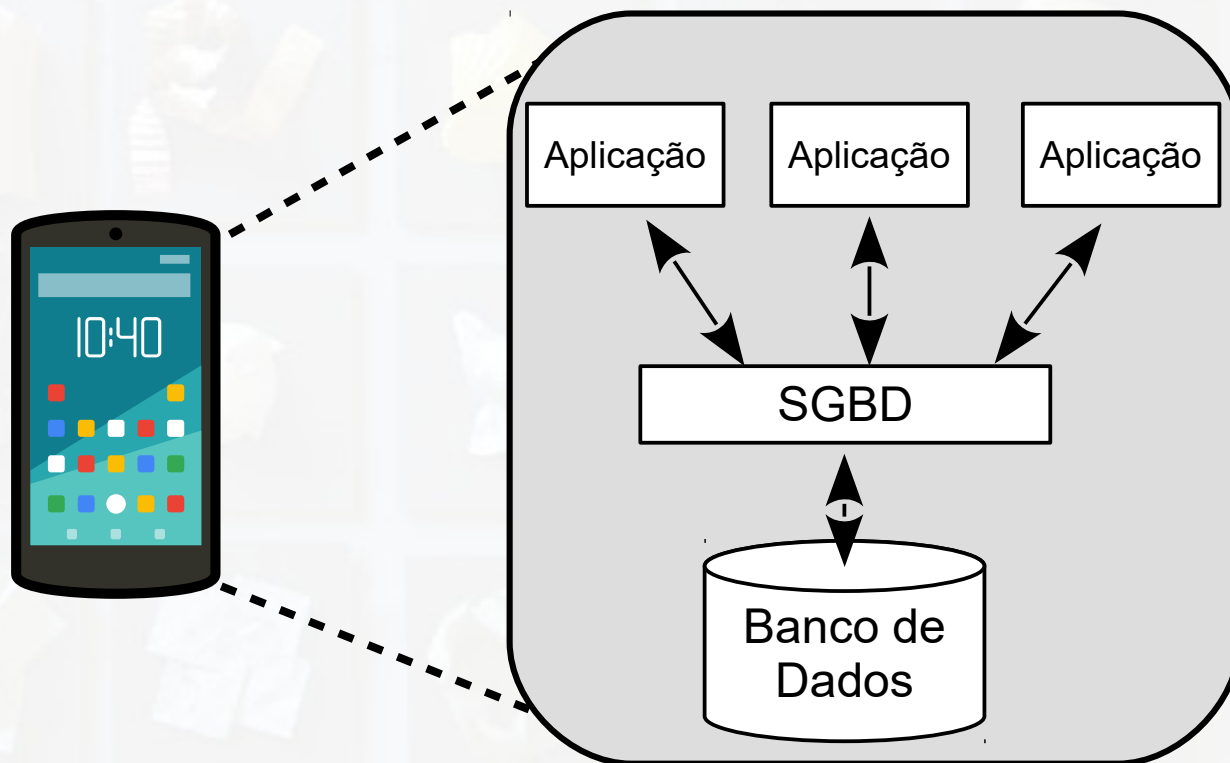
- Sistema de loja simples
- Dados para análise de um pesquisador



Arquitetura Local

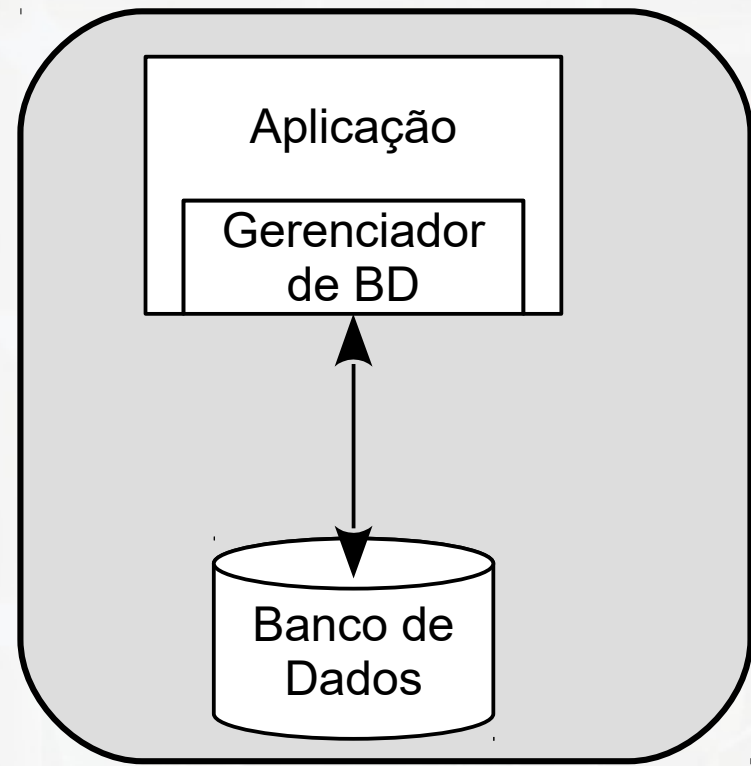
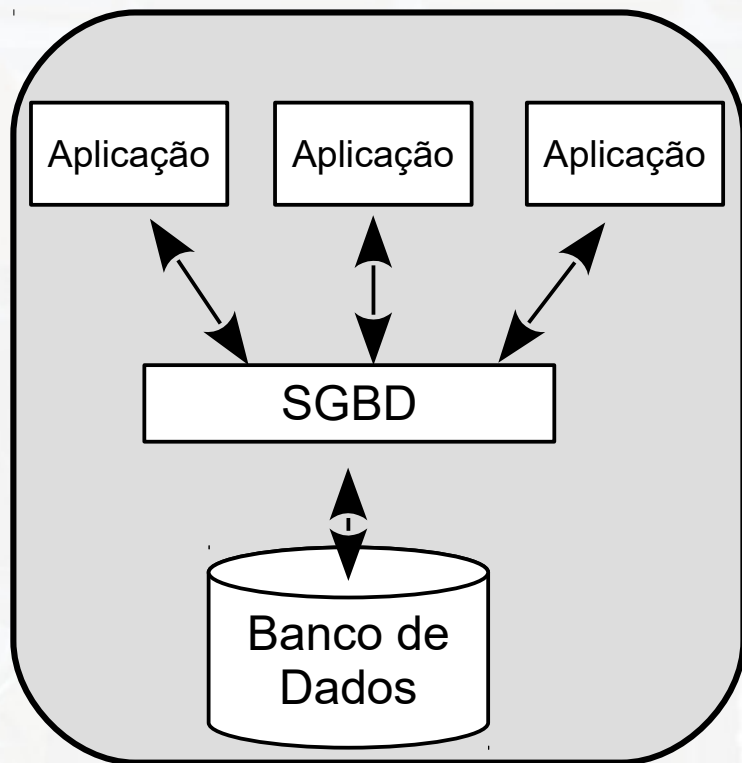
Exemplos - Celular

- Agenda local
- Dados de perfil e registro de operações

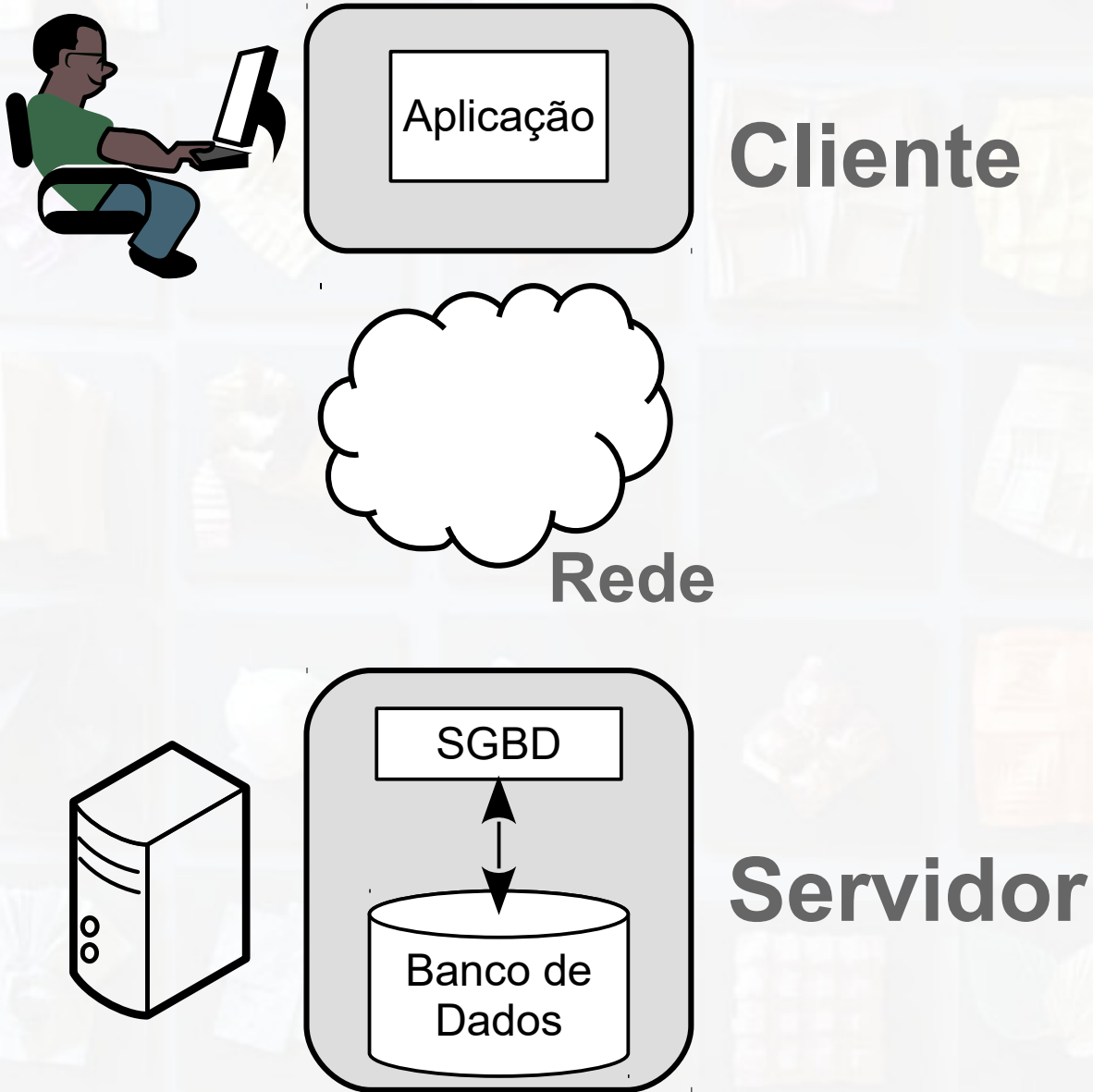


Arquitetura Local

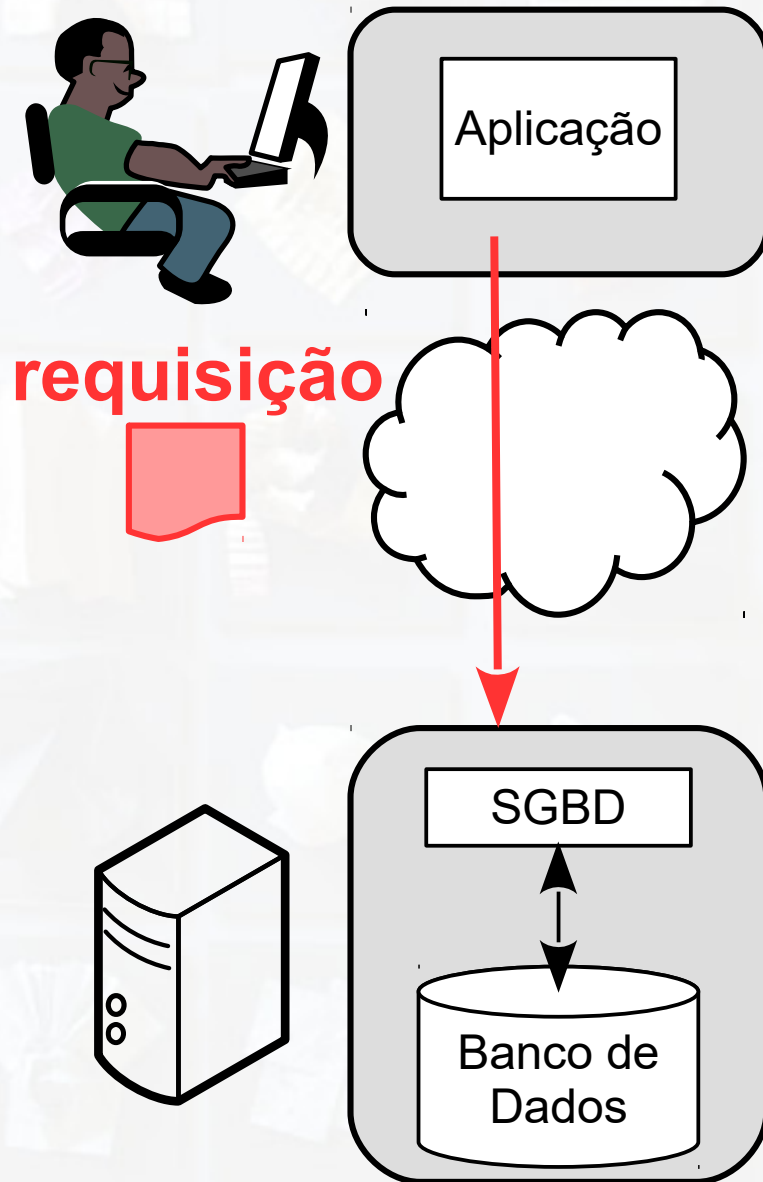
- O SGBD pode ser um sistema independente ou uma biblioteca da aplicação (embutido).



Arquitetura Cliente-Servidor

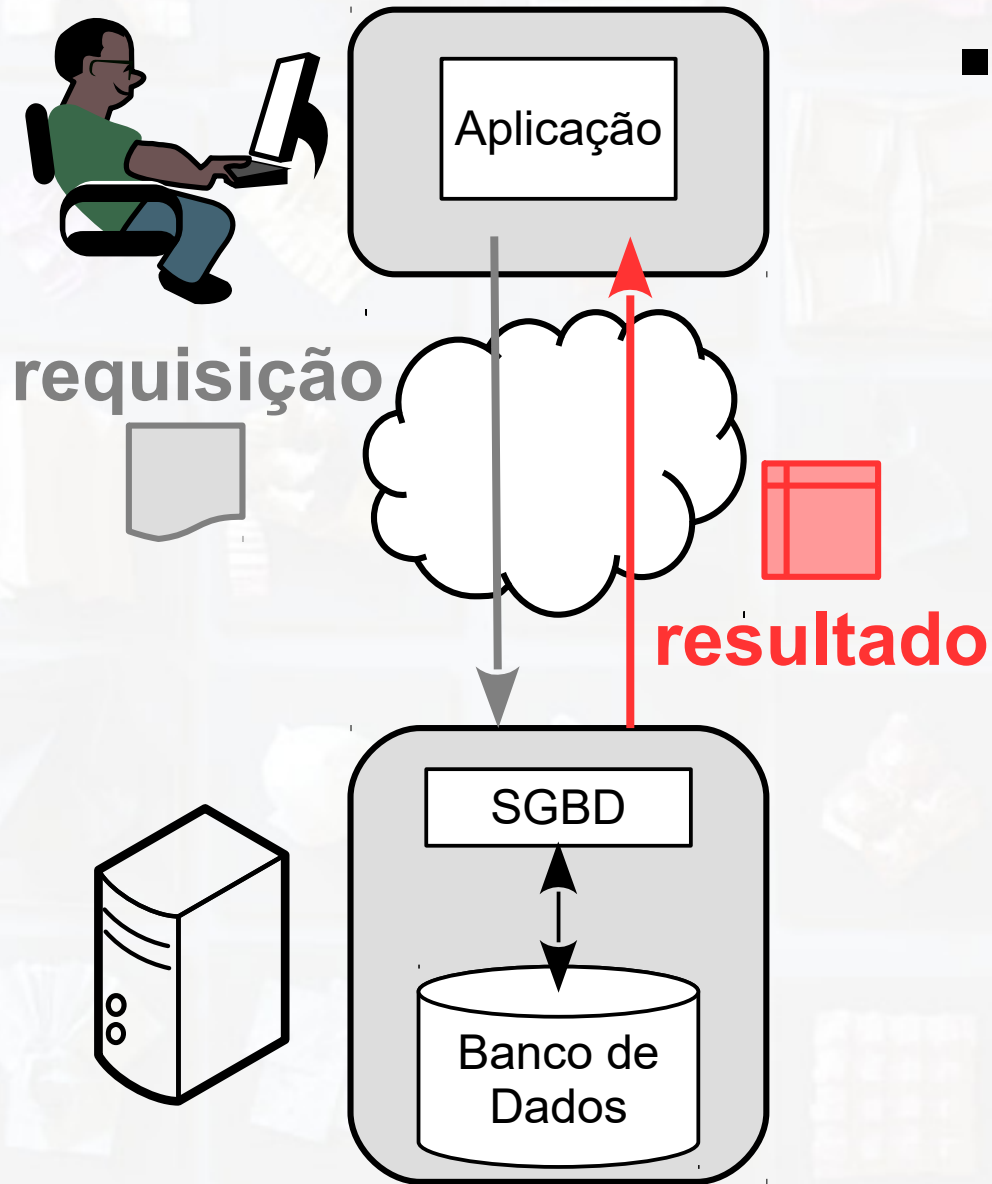


Arquitetura Cliente-Servidor



- Aplicação (cliente) envia consulta para obter dados

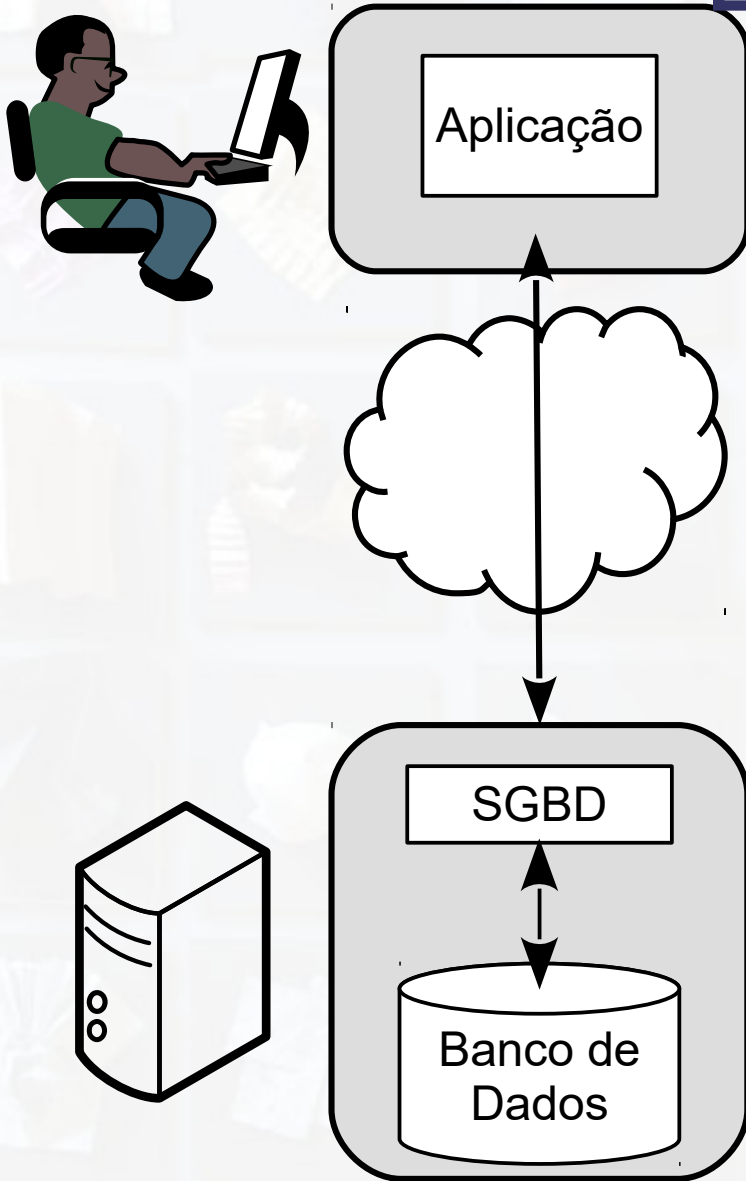
Arquitetura Cliente-Servidor



- SGBD (servidor) processa consulta e retorna dados para cliente

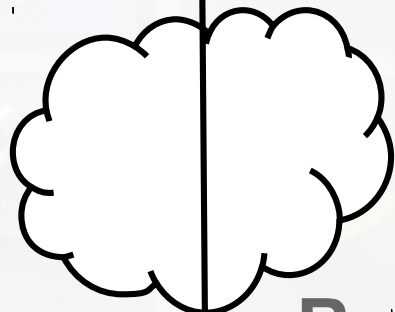
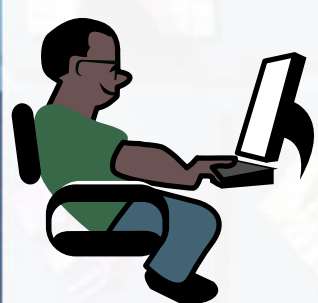
Arquitetura Cliente-Servidor

Exemplos

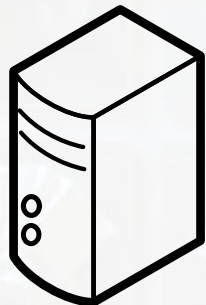
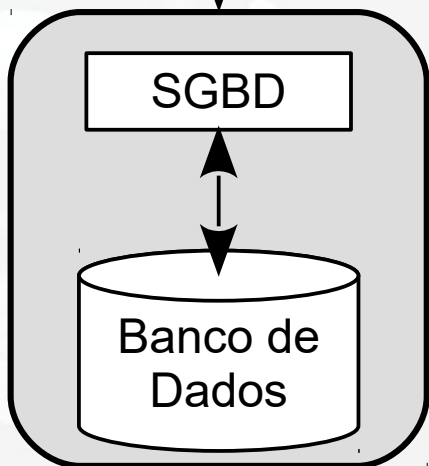


- Controle de estoque de conjunto de farmácias
- Inscritos em uma conferência
- Reclamações para a prefeitura georeferenciadas

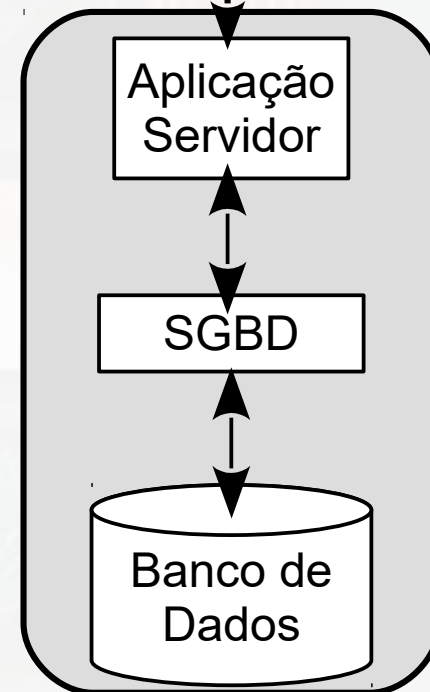
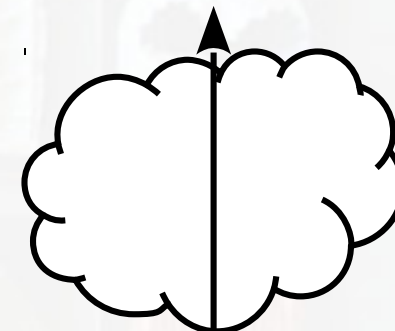
Arquitetura de Duas Camadas



Rede



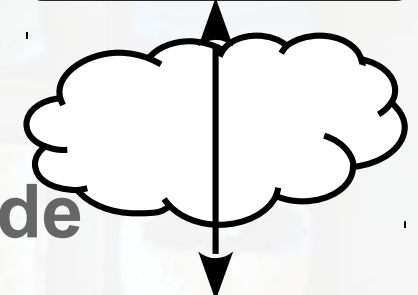
**Camada
Cliente**



**Camada
Servidor**

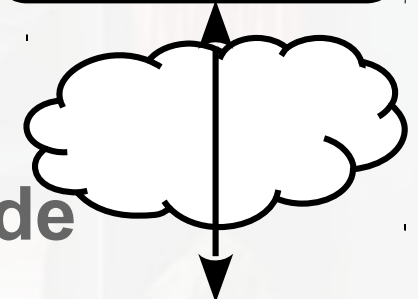
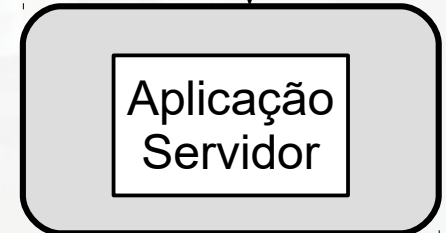
Arquitetura de Três Camadas

Camada de Apresentação



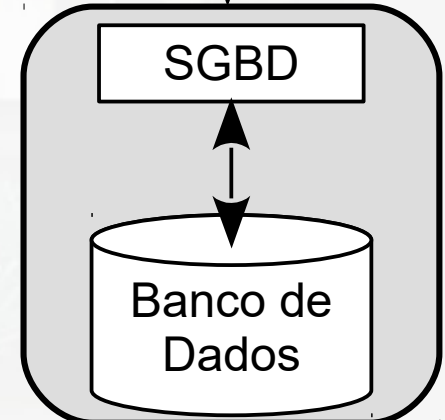
Rede

Camada Lógica de Negócios



Rede

Camada de Serviços de BD

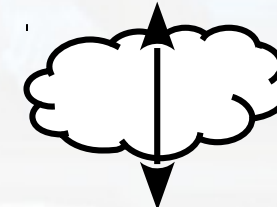


Arquitetura de n Camadas

Camada de Apresentação



Aplicação Ciente



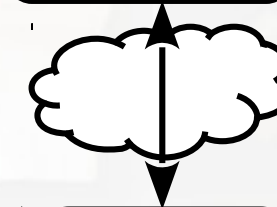
Camada Web



Camada Lógica de Negócios



Aplicação Servidor

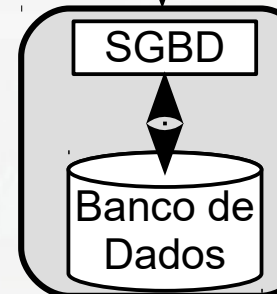


Camada de Serviços de BD



SGBD

Banco de Dados



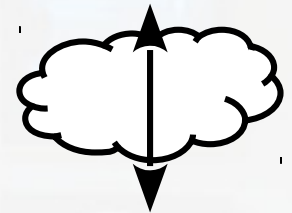
n Camadas Exemplos

- Lojas Virtuais
- Sistemas da administração pública
- Sistemas integrados de grandes corporações

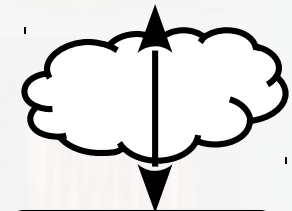
Camada de Apresentação



Aplicação Ciente



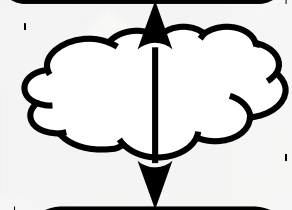
Camada Web



Camada Lógica de Negócios



Aplicação Servidor

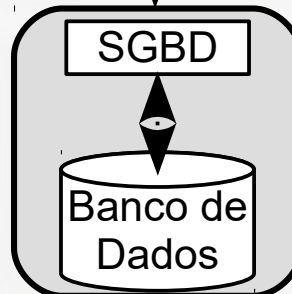


Camada de Serviços de BD



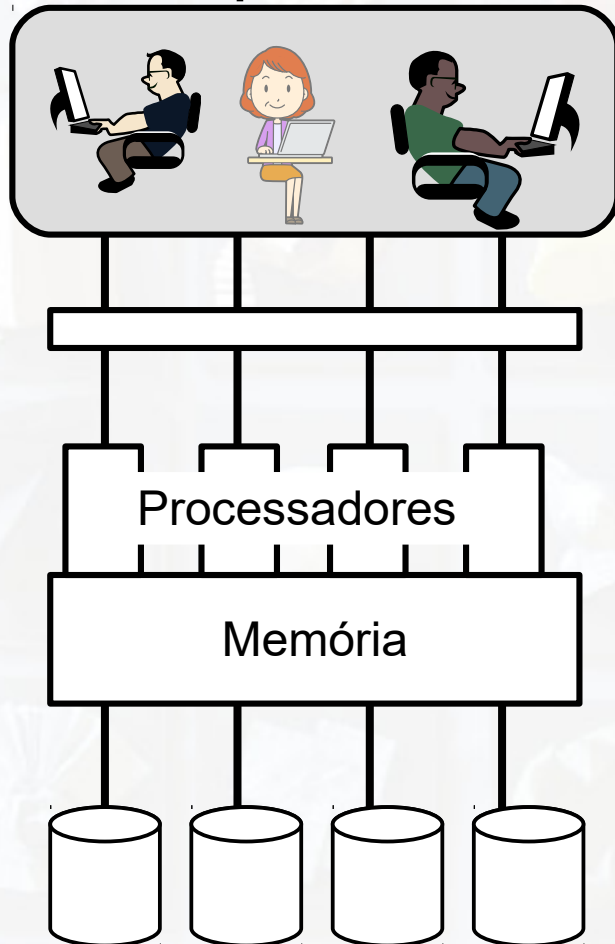
SGBD

Banco de Dados

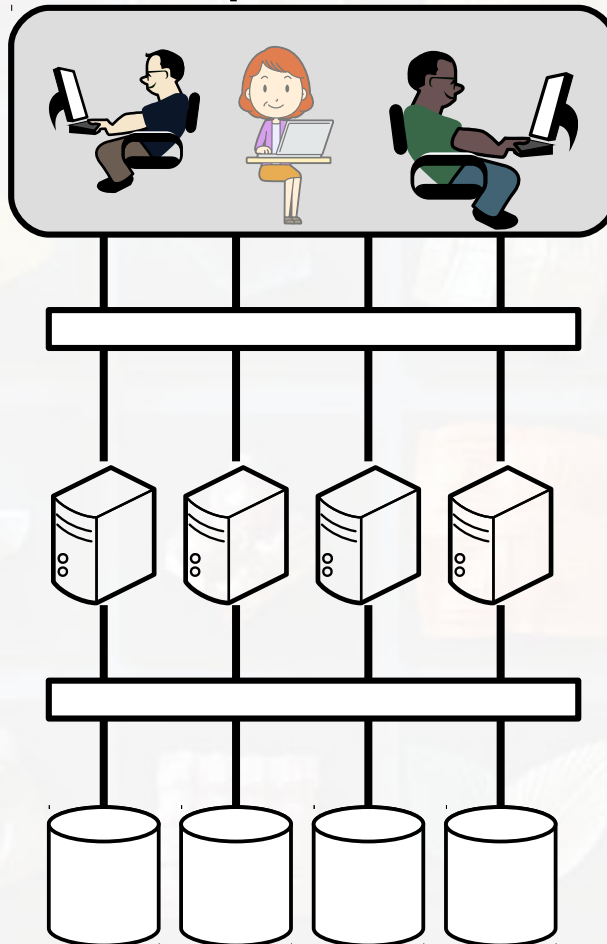


Bancos de Dados Paralelos e Distribuídos

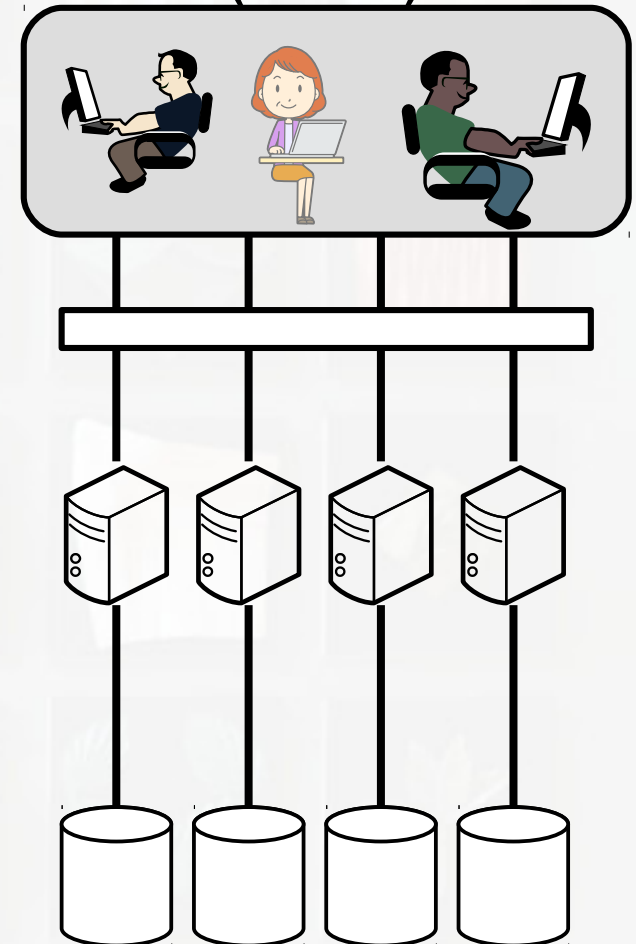
Memória
Compartilhada



Disco
Compartilhado

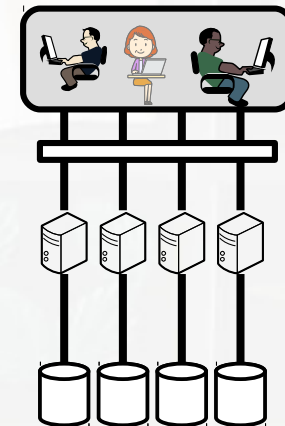
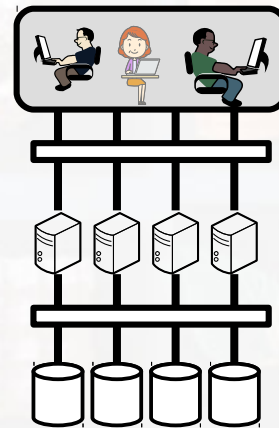
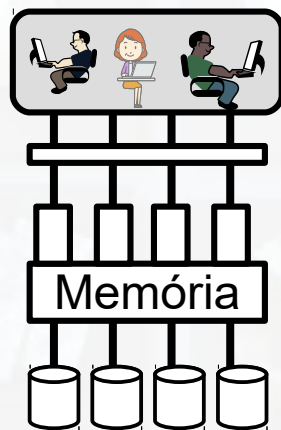


Nada
Compartilhado
(rede)



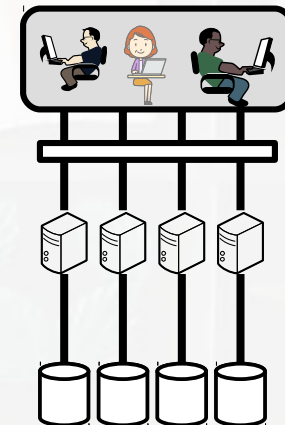
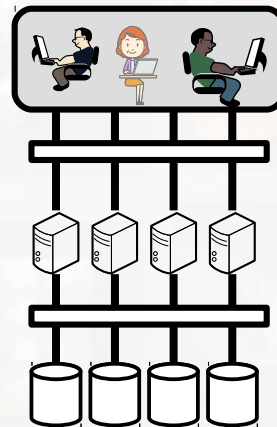
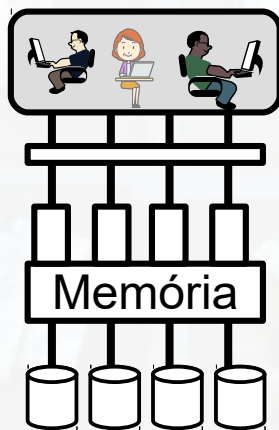
Bancos de Dados Paralelos e Distribuídos - Desafios

- Como distribuir os dados entre os dispositivos de armazenamento
- Como sincronizar dados distribuídos
- Como lidar com sistemas distribuídos no planeta



Bancos de Dados Paralelos e Distribuídos - Exemplos

- Redes sociais
- Busca de páginas na Web
- Sistemas de mapas globais
- Dados de pesquisa compartilhados - e.g., biodiversidade

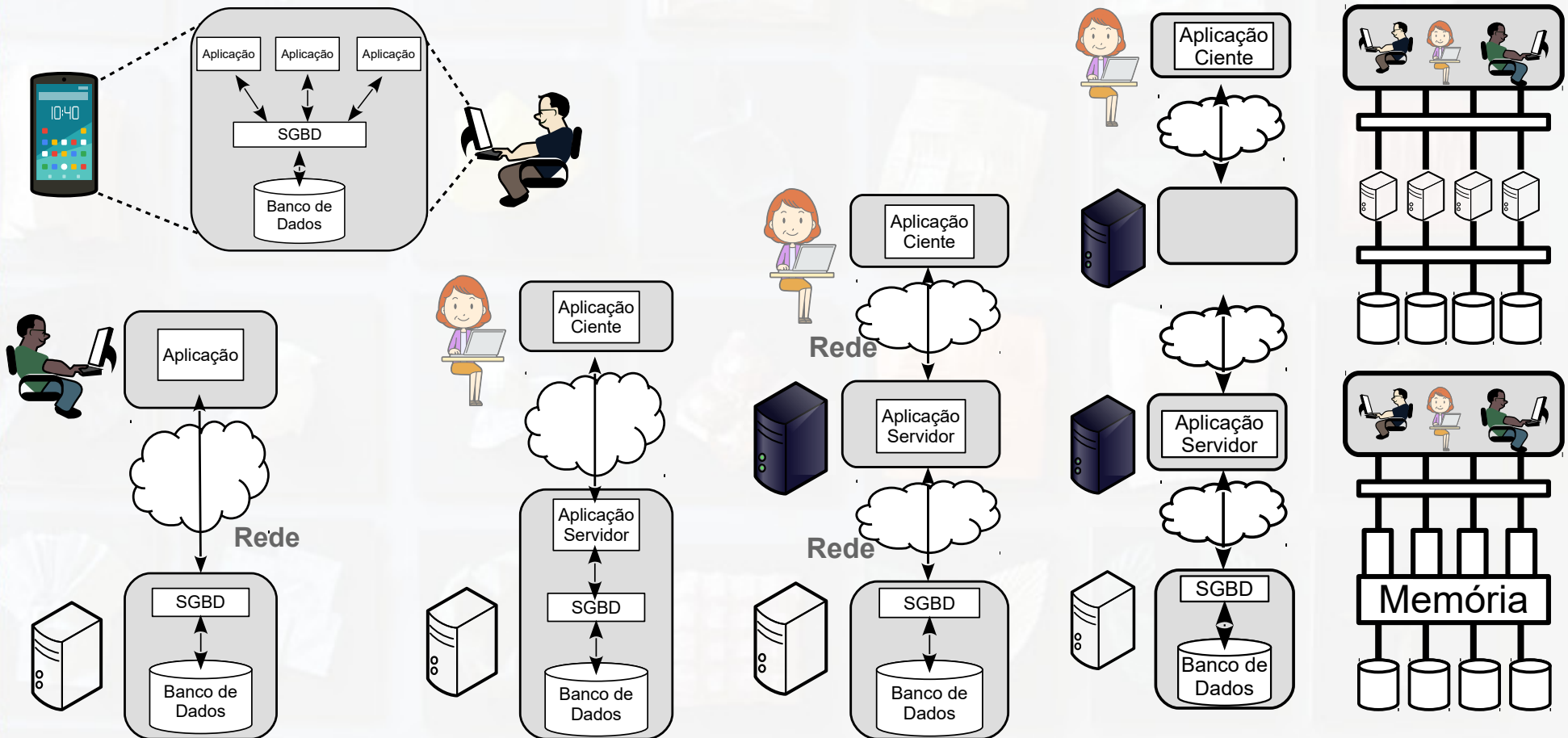


Sumarizando

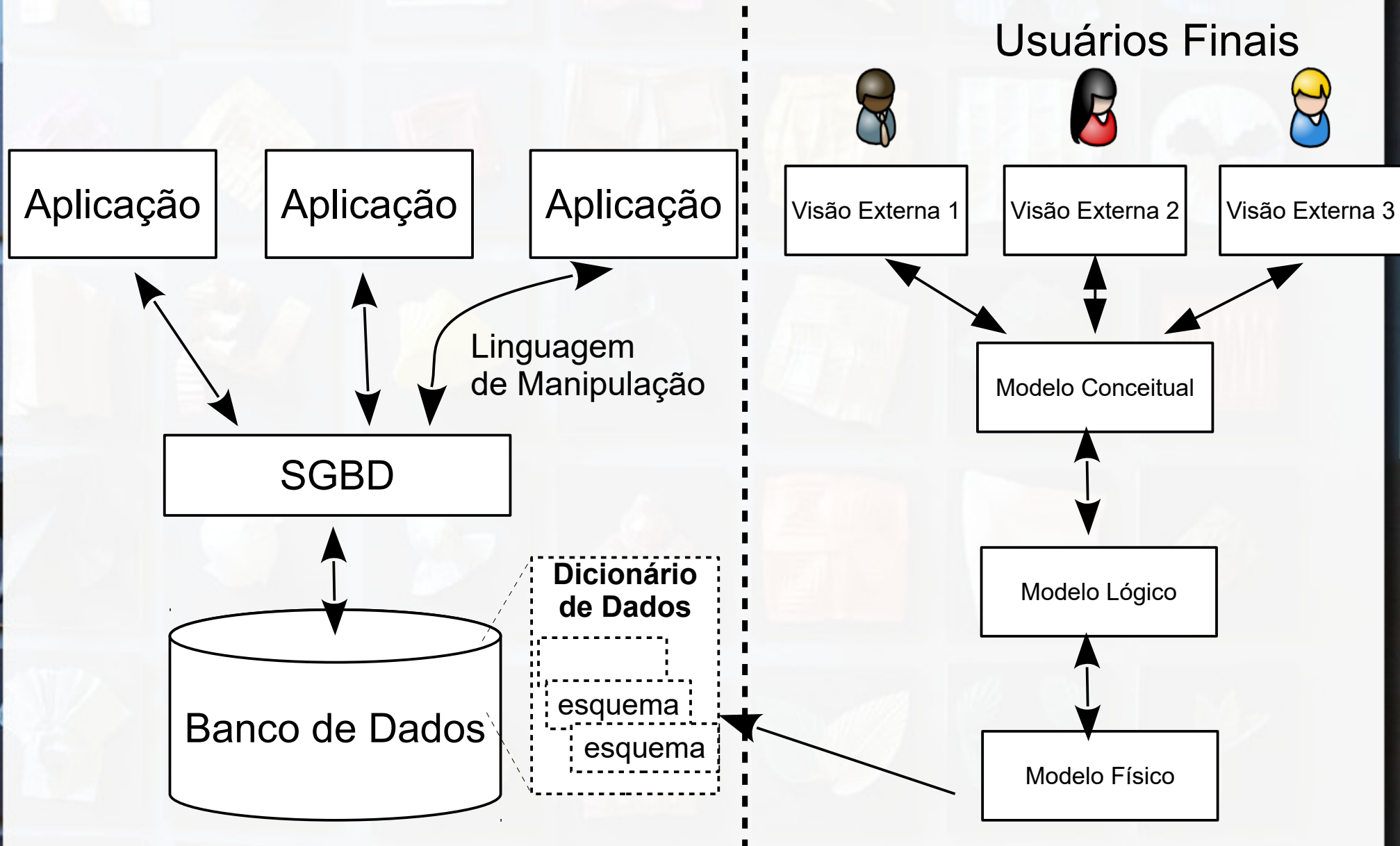
- Grandes volumes de dados justificam a criação de SGBDs

Sumarizando

- Há diferentes arquiteturas de SGBD conforme o cenário



Sumarizando



O que está mudando?

- **Dados estão por toda a parte**
 - não somente centralizados em um banco
 - produzidos de forma distribuída e interligados
- **Modelagem e semântica ganham importância**
 - Web Semântica e ontologias
- **Data deluge e Big Data**
 - novas abordagens (NoSQL)
 - processamento e armazenamento descentralizados
 - bancos de dados em memória

Referências

- Chen, Peter Pin-Shan (1976) **The entity-relationship model - toward a unified view of data**. ACM Trans. Database Systems, ACM, 1, 9-36.
- Elmasri, Ramez; Navathe, Shamkant B. (2005) **Sistemas de Bancos de Dados**. Addison-Wesley, 4ª edição em português.
- Guimarães, Célio (2003) **Fundamentos de Bancos de Dados: Modelagem, Projeto e Linguagem SQL**. Editora UNICAMP, 1ª edição.
- Heuser, Carlos Alberto (2004) **Projeto de Banco de Dados**. Editora Sagra Luzzato, 5ª edição.

Referências

- Ramakrishnan, Raghuram; Gehrke, Johannes (2003) **Database Management Systems**. McGraw-Hill, 3rd edition.

André Santanchè

<http://www.ic.unicamp.br/~santanche>

Licença

- Estes slides são concedidos sob uma Licença Creative Commons. Sob as seguintes condições: Atribuição, Uso Não-Comercial e Compartilhamento pela mesma Licença.
- Mais detalhes sobre a referida licença Creative Commons veja no link:
https://creativecommons.org/licenses/by-nc-sa/4.0/deed.pt_BR
- Agradecimento a Goran Konjevod [<https://www.flickr.com/photos/23913057@N05/>] por sua fotografia “50-50 Show III” usada na capa e nos fundos, disponível em [<https://flic.kr/p/advD33>] vide licença específica da fotografia.