

Introdução

Bancos de Dados

Luiz Celso Gomes-Jr
gomesjr@dainf.ct.utfpr.edu.br

Mundo de Dados



Mundo de Dados

- Curtidas em Redes Sociais
- Mensagens Instantâneas
- Páginas na Web
- Notas dos alunos
- Fotos
- Localização de Pokémons
- Sinais de televisão
- Saldo de contas correntes
- Produtos à venda

Dados X Aplicações

Curtidas



Notas dos alunos



Localização Pokémons



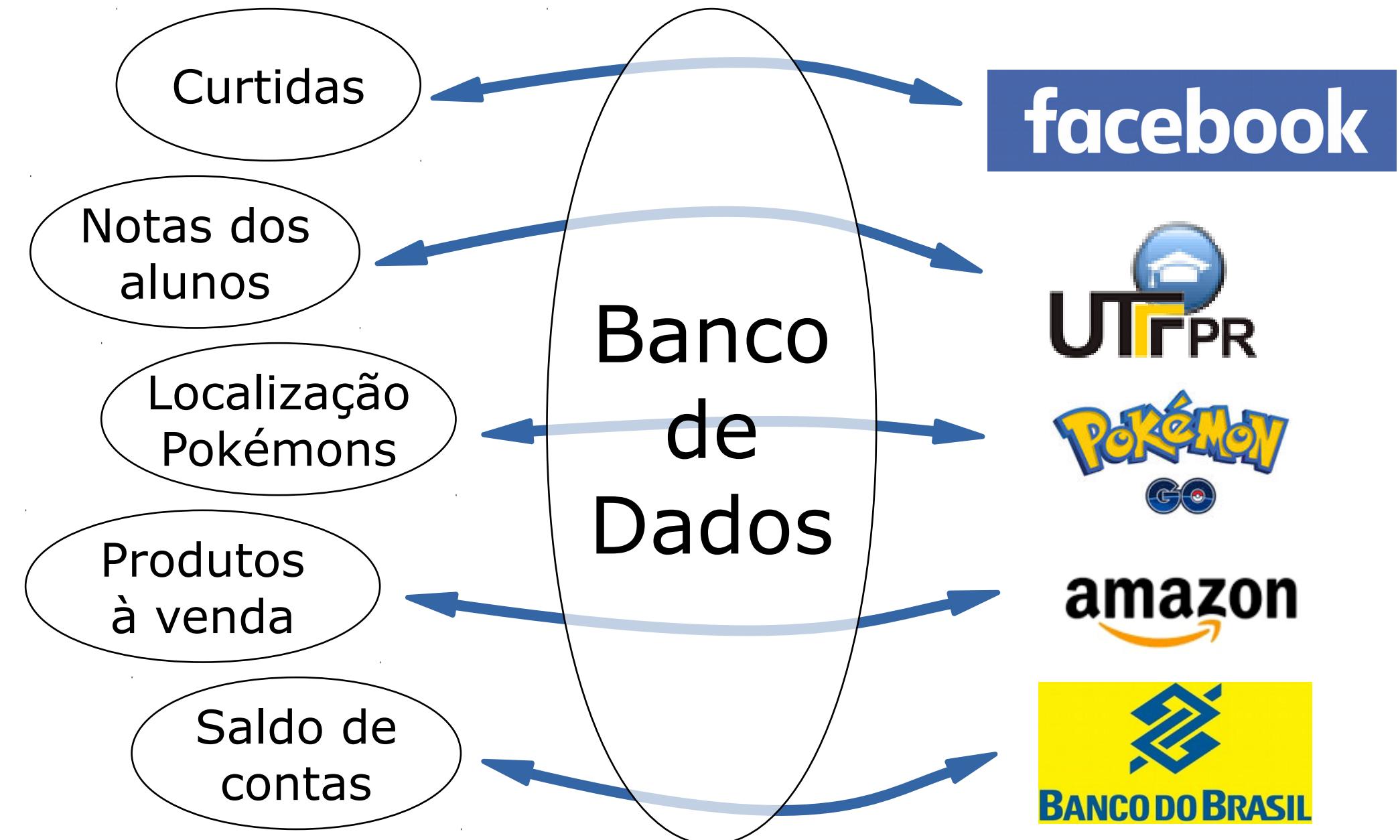
Produtos à venda



Saldo de contas



Dados X Aplicações



Dados e aplicações

- Diversas aplicações geram e lidam com uma **grande quantidade de dados**
- É importante garantir a **integridade** dos dados e disponibilizá-los da forma mais **eficiente** possível
- **Sistemas de Gerenciamento de Bancos de Dados (SGBDs)** são responsáveis pelo armazenamento, gerenciamento e disponibilização dos dados

Um Sistema de Gerenciamento de Bancos de Dados provê:

Armazenamento e acesso a uma
grande quantidade de dados
persistentes de forma **eficiente,**
confiável, conveniente, segura e
multi-usuário.

Super SGBD



Preciso implementar um sistema! Quem cuidará dos meus dados??

Eu,
o Super SGBD

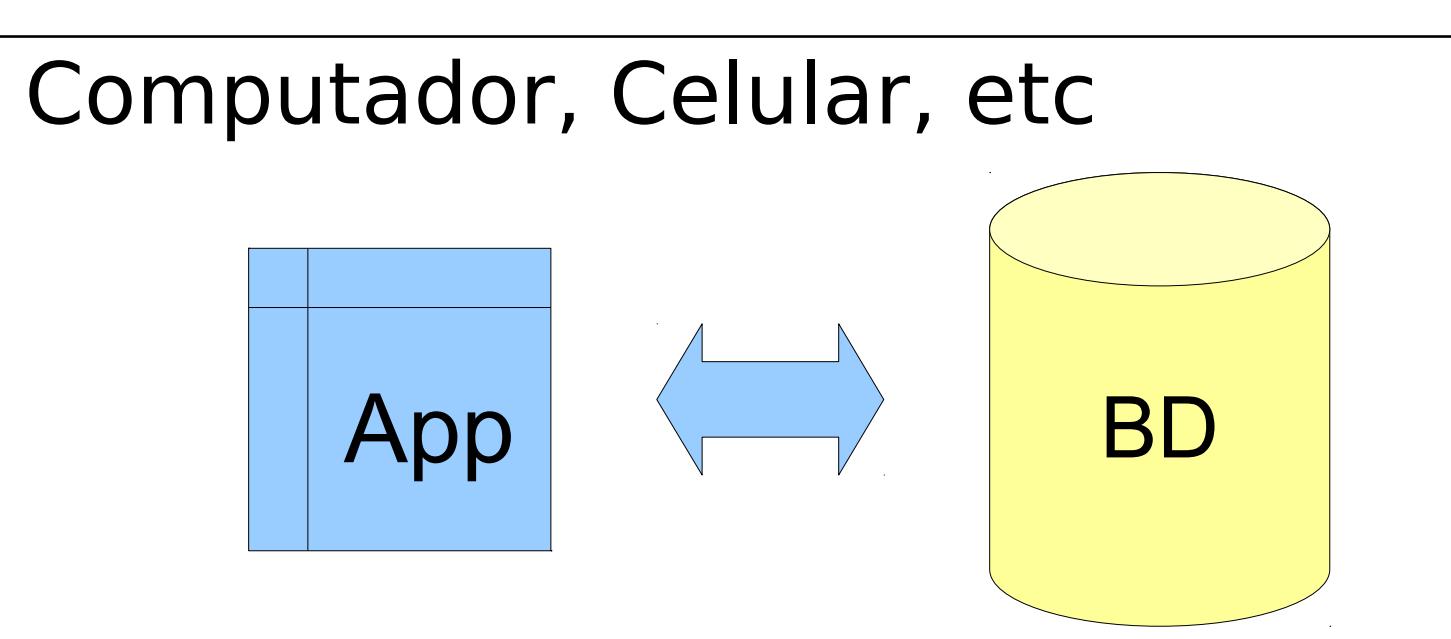
SGBDs desempenham muitas tarefas importantes que sobrecarregariam os programadores. Veremos ao longo do curso o quanto eles facilitam nosso trabalho.

Arquiteturas típicas para SGBDs

- O SGBD intermedia a manipulação dos dados para as aplicações
- Como esta intermediação é feita e quais elementos estão envolvidos depende do contexto da aplicação
- Aplicações podem demandar diversas configurações de arquiteturas locais ou distribuídas (ou híbridas)

Arquitetura local

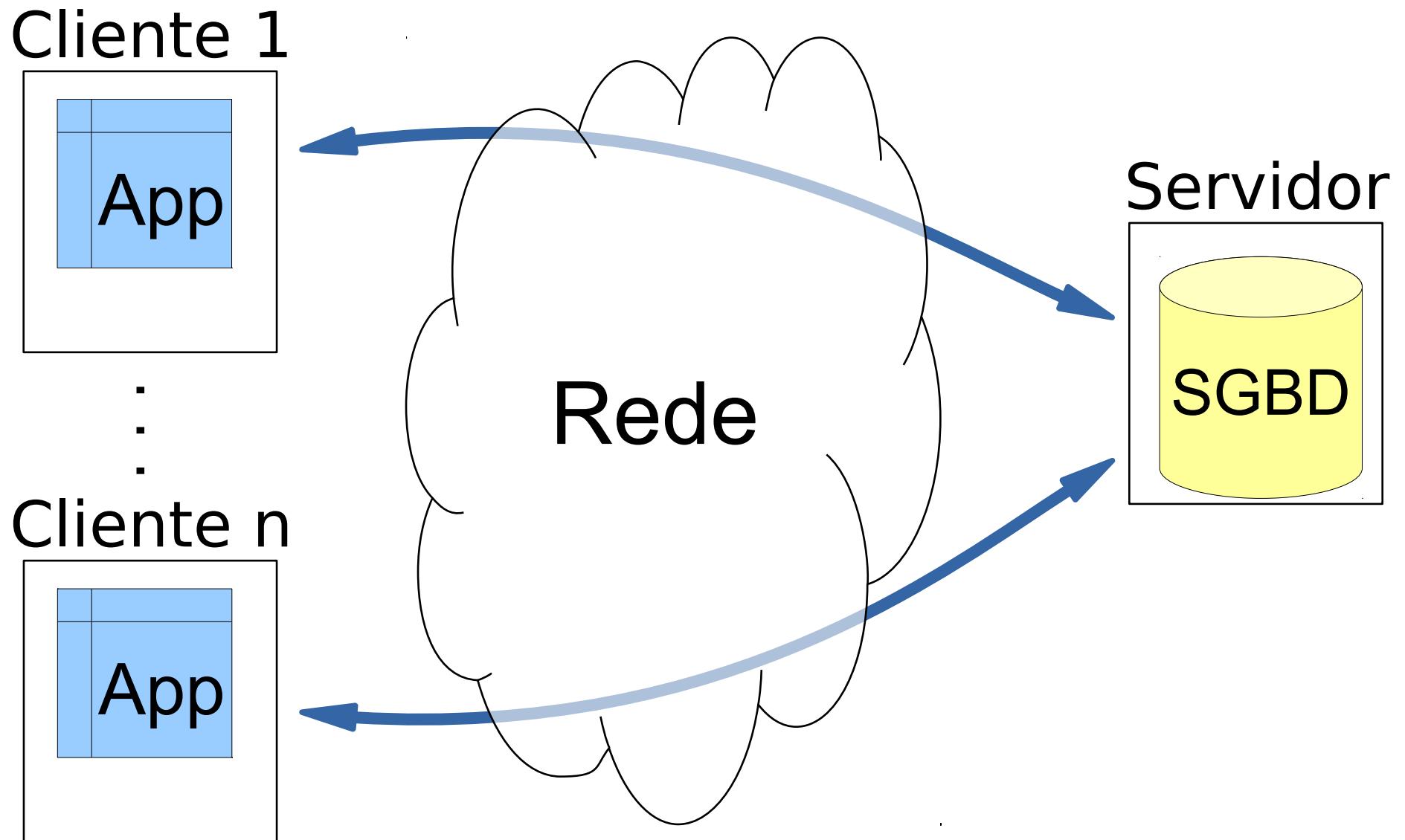
- Não costumam ser classificados como SGBDs
- O banco é um software ou biblioteca executado no mesmo dispositivo da aplicação



Arquitetura Cliente-Servidor

- Usada na maioria das instituições
- Usuário acessa a aplicação por um dispositivo Cliente (desktop, laptop, celular...)
- Aplicação envia consultas para obter dados do SGBD (Servidor)
- SGBD processa consulta e retorna dados para serem exibidos no Cliente
- Exemplos: Folha de pagamentos, iTunes

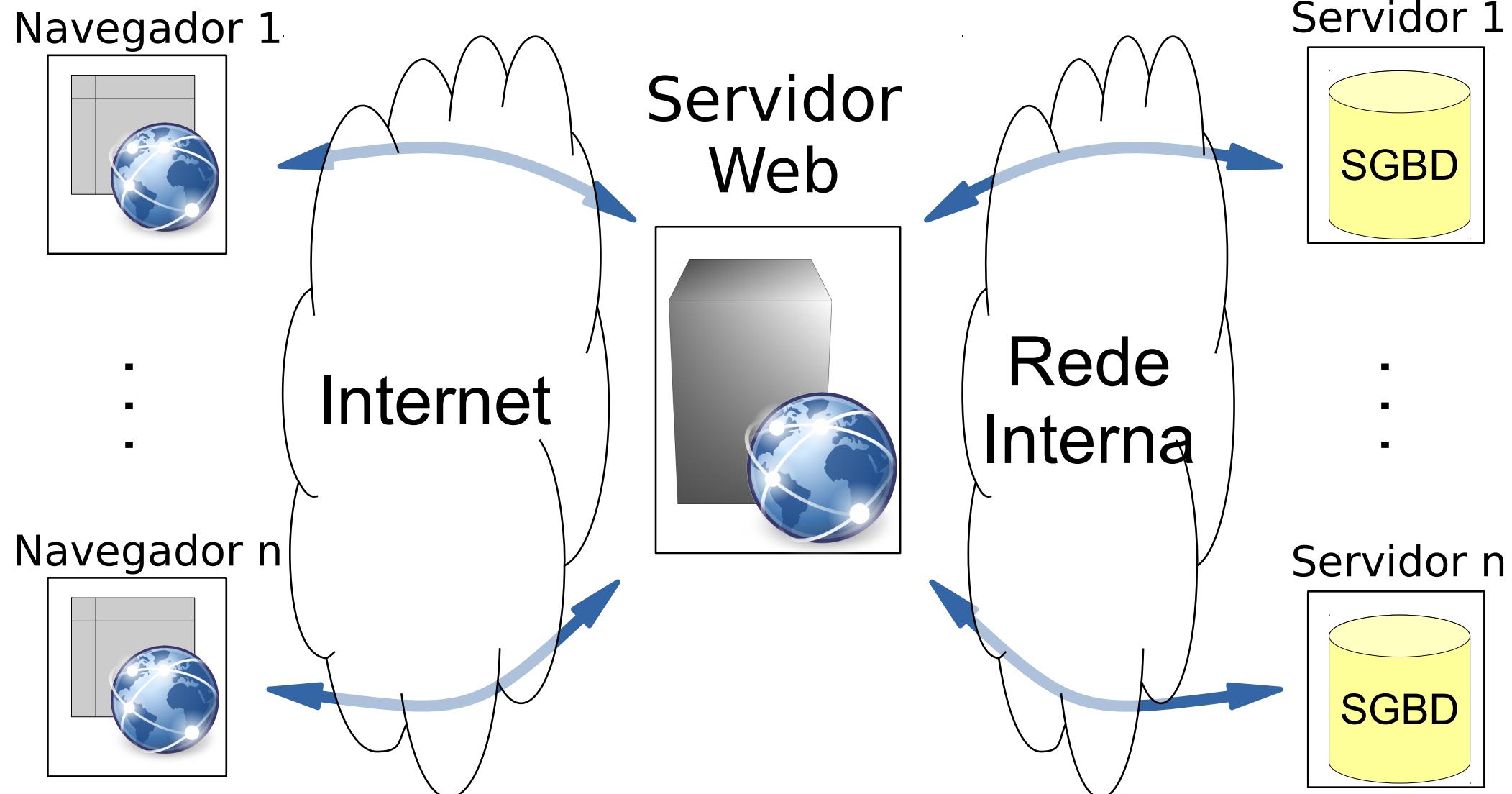
Arquitetura Cliente-Servidor



Arquitetura Web 1.0

- Usada na maioria dos sites “normais”
- Usuário usa o navegador para requisitar páginas para um Servidor Web
- Servidor Web envia consultas a um ou mais SGBDs para obter dados e montar a página
- Exemplos: bancos online, sites de empresas
- Muitas apps e sites como Facebook, Google precisam de arquiteturas mais complexas. Veremos estes casos no fim do curso.

Arquitetura Web 1.0



Exemplo: Facebook

1. Usuário abre o navegador e entra em facebook.com
2. Servidor Web do facebook recebe a requisição do usuário
3. Servidor Web do facebook obtém dados do mural de um SGBD interno
4. Servidor Web do facebook obtém dados de propaganda de um outro SGBD interno
5. Servidor Web do facebook monta a página e envia para o navegador exibir

Conceitos importantes

- **Modelo de Dados** – Como representar os dados (listas, tabelas, árvores, grafos...)
- **Esquema X Dados** – Separação entre a definição (tipo) e o conteúdo (instâncias, variáveis)
- **Linguagem de Consulta** – Como requisitar dados para o banco de forma conveniente (exemplo: como dizer para o SGBD “obtenha o salário médio de todos os funcionários que têm mais de 30 anos”)

Domadores de SGBDs

- **Implementador:** quem constrói o sistema (em geral grandes empresas como MS, IBM, Oracle)
- **Projetista:** Faz análise e cria o esquema (define os tipos ou classes)
- **Desenvolvedor** de aplicações: programador que usa o banco de dados (faz consultas, modifica dados)
- **Administrador** (DBA): Gerencia o carregamento de dados, segurança, backups, eficiência, etc.

Uma breve história dos SGBDs

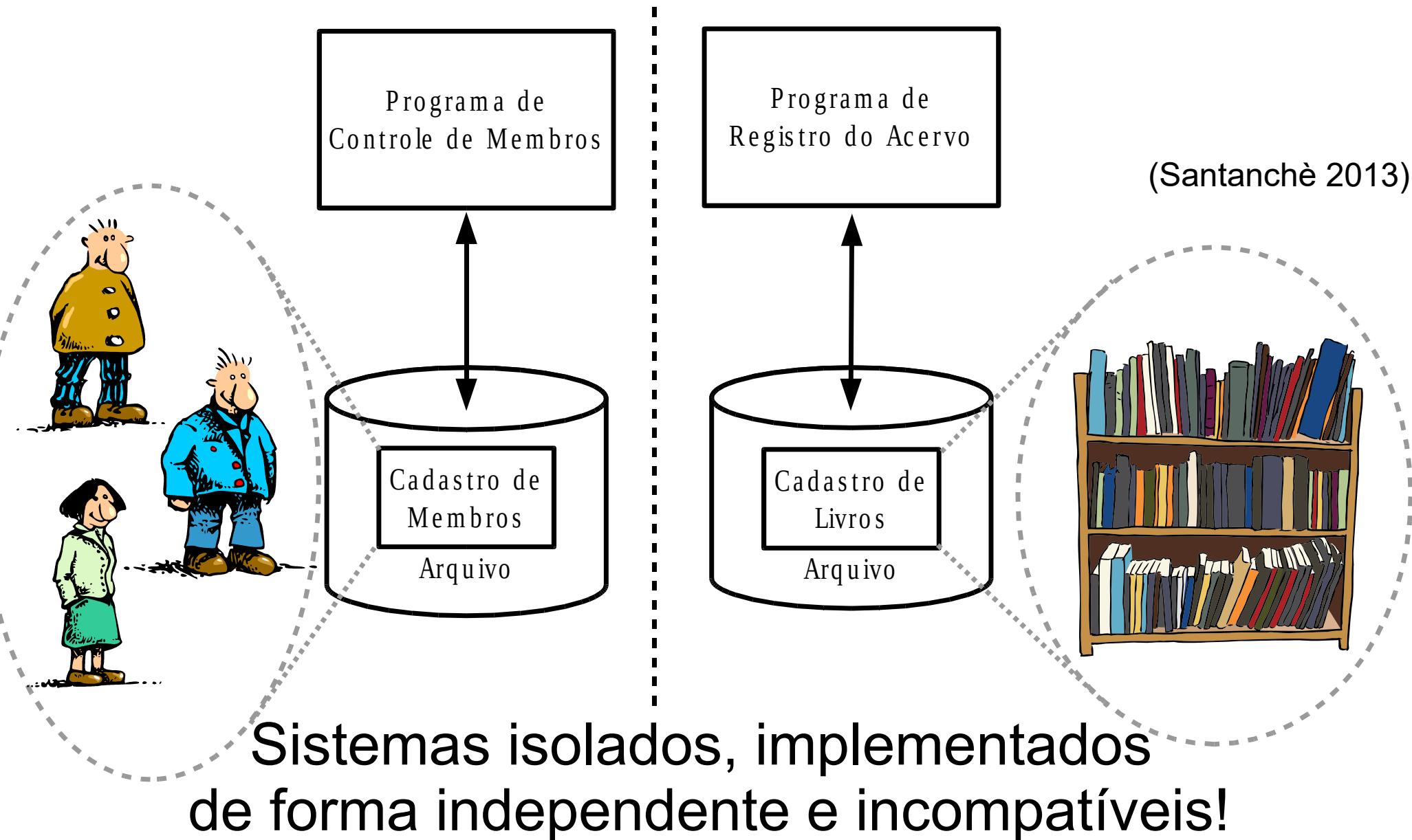
- Caos (sem SGBDs)
- Bancos de dados relacionais surgem, dão ordem ao caos, dominam indústria, fazem bilionários
- Web explode e inicia um novo caos!!!

Caos (até ~1980)

- Avanços em discos e fitas magnéticas permitem armazenamento de grandes quantidades de dados
- Ninguém sabe como armazenar os dados
- Cada um faz do seu jeito
- Métodos populares: hierárquico, rede
- Manutenção cara, difícil integração entre sistemas, bugs, inconsistência...



Exemplo: Biblioteca



Sistemas/Arquivos Isolados

- Redundância não controlada
 - Repetição
 - Inconsistência
- Barreiras para referenciar dados em outros sistemas
- Dificuldades com:
 - acesso concorrente
 - integridade e recuperação em caso de falha
 - segurança e controle de acesso

SGBDs (80's)

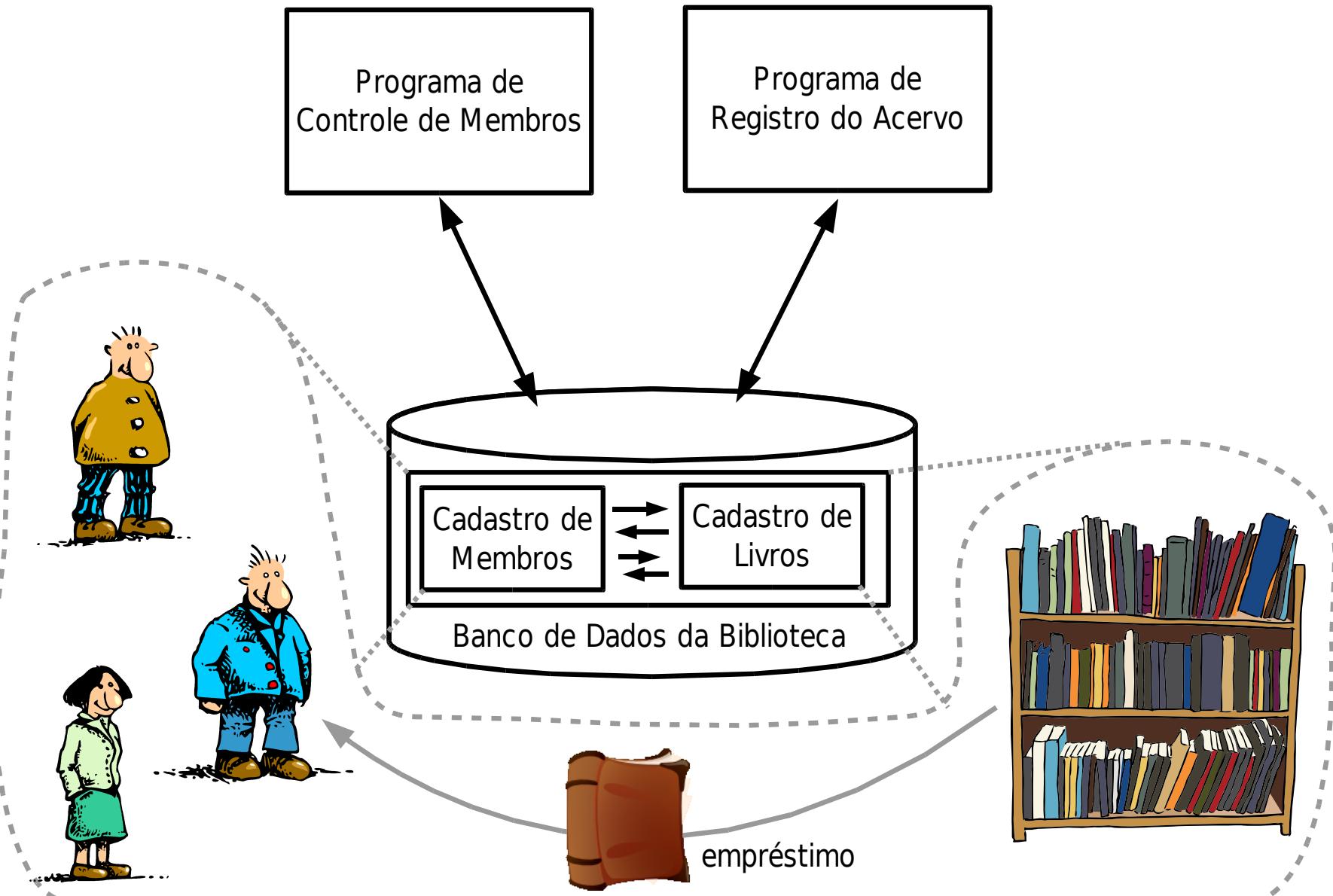
- Modelos conceituais sólidos (Modelo Relacional)
- Separação entre visão lógica e implementação física
- Linguagens de consulta poderosas (SQL)
- Implementações comerciais eficientes



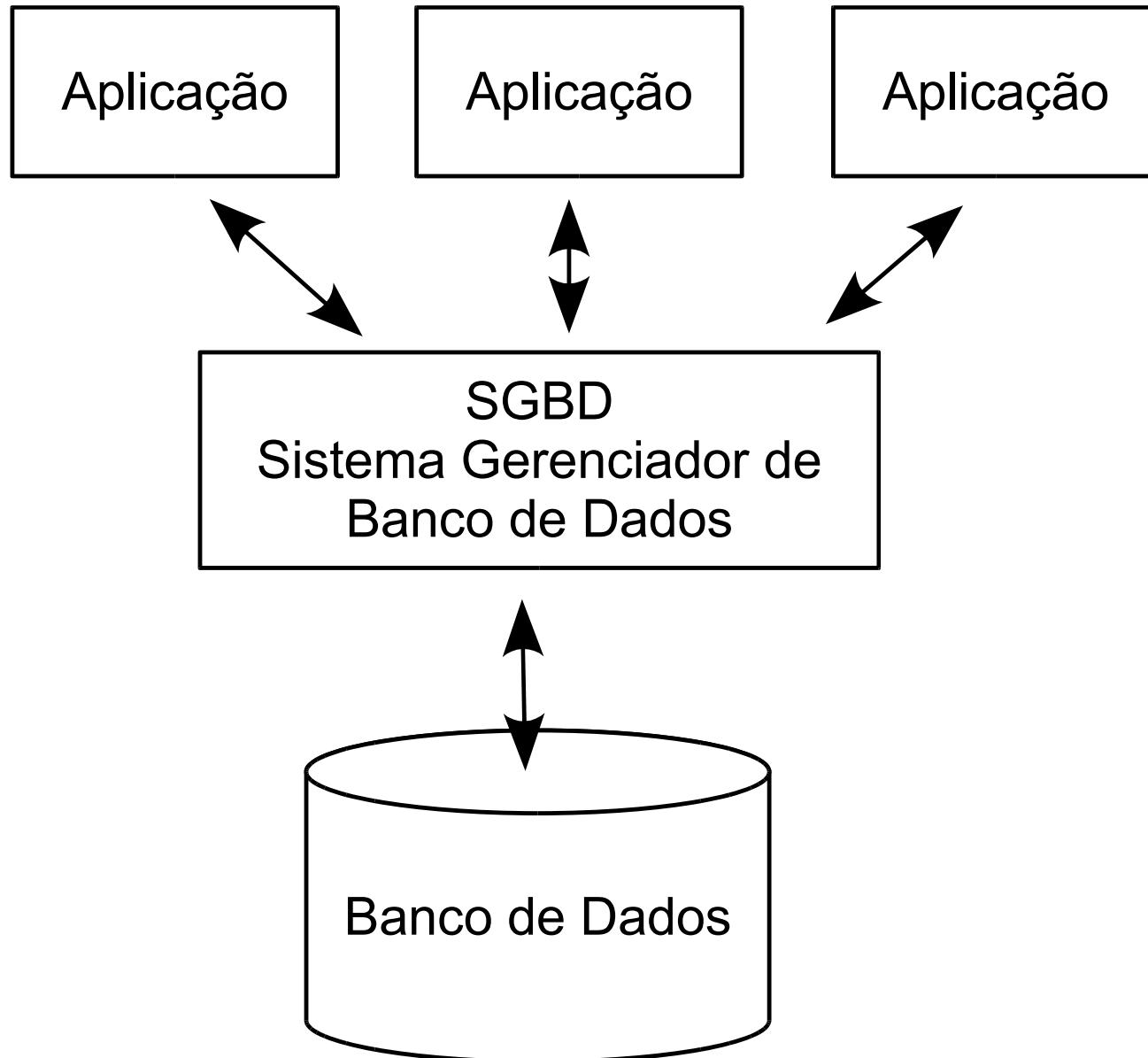
Benefícios dos SGBDs

- Independência de dados
- Acesso eficiente
- Tempo reduzido no desenvolvimento de aplicações
- Segurança e integridade de dados
- Administração de dados uniforme
- Acesso concorrente
- Recuperação de falhas

SGBD



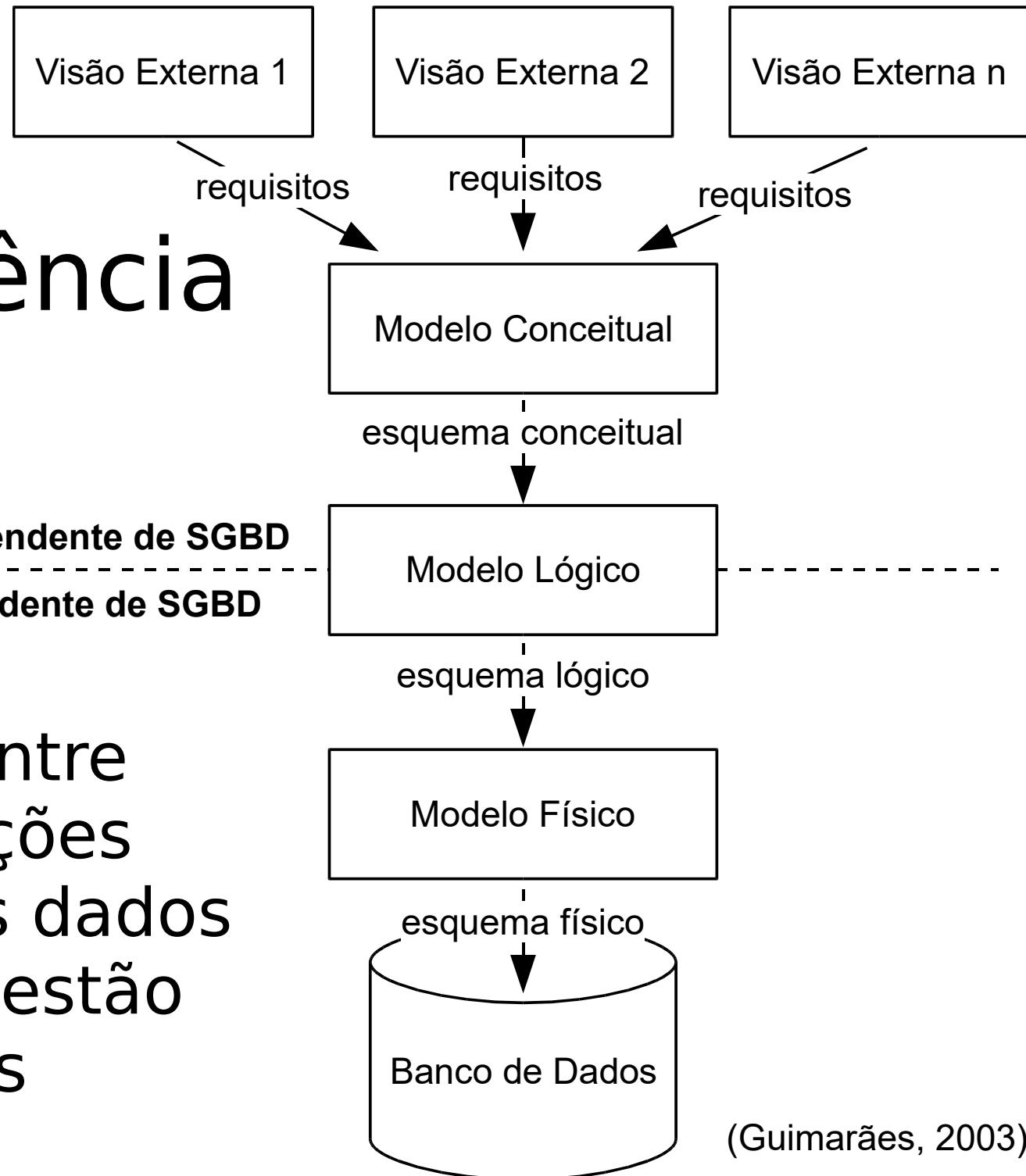
SGBD



Independência de Dados

Independente de SGBD
Dependente de SGBD

- Separação entre como aplicações enxergam os dados e como eles estão armazenados



(Guimarães, 2003)

Modelo Relacional

- Proposto por E. F. Codd em 1970 no artigo:
 - “A Relational Model for Large Shared Data Banks”
- Independência da estrutura interna
 - “Activities of users at terminals and most application programs should remain unaffected when the internal representation of data is changed [...]”
(Codd, 1970)

Modelo Relacional

- Modelo mais amplamente utilizado por SGBDs
- Maiores empresas de informática oferecem soluções: IBM, Microsoft, Oracle, SAP
- Grandes projetos Open Source: MySQL, PostgreSQL, SQLite
- De celulares à data centers
- Mercado de U\$24bi (2012)

Larry Ellison, ORACLE

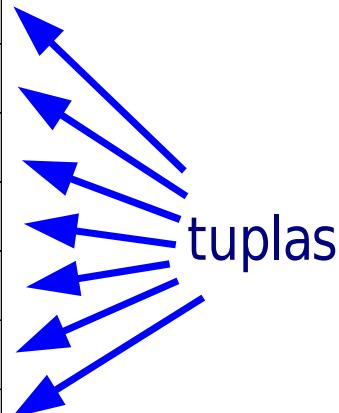


Relação (Tabela)

- Tupla = conjunto ordenado de atributos
- Valores de atributos são atômicos e monovalorados

Livro  relação

| 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 |



Consultas (SQL)

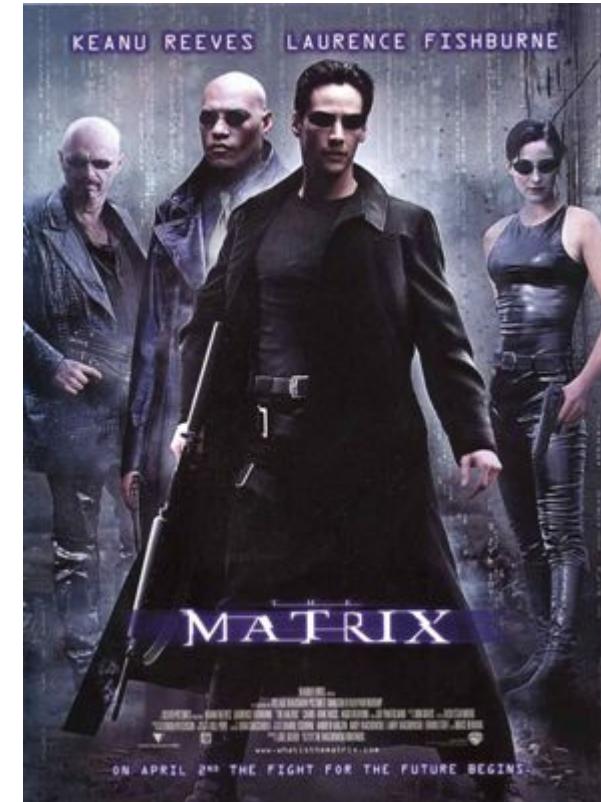
- Obtenha todos os livros de categoria “drama”
- `SELECT * FROM Livro
WHERE Categoria = “Drama”`

Livro

| 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 |

SGBDs (90's em diante)

- Modelo relacional continua dominando o mercado
- BDs relacionais são adequados para mais de 90% das aplicações
- Mas a Internet, Celulares, Redes Sociais, Sensores... Têm causado uma revolução!



BigData – Dilúvio de Informação



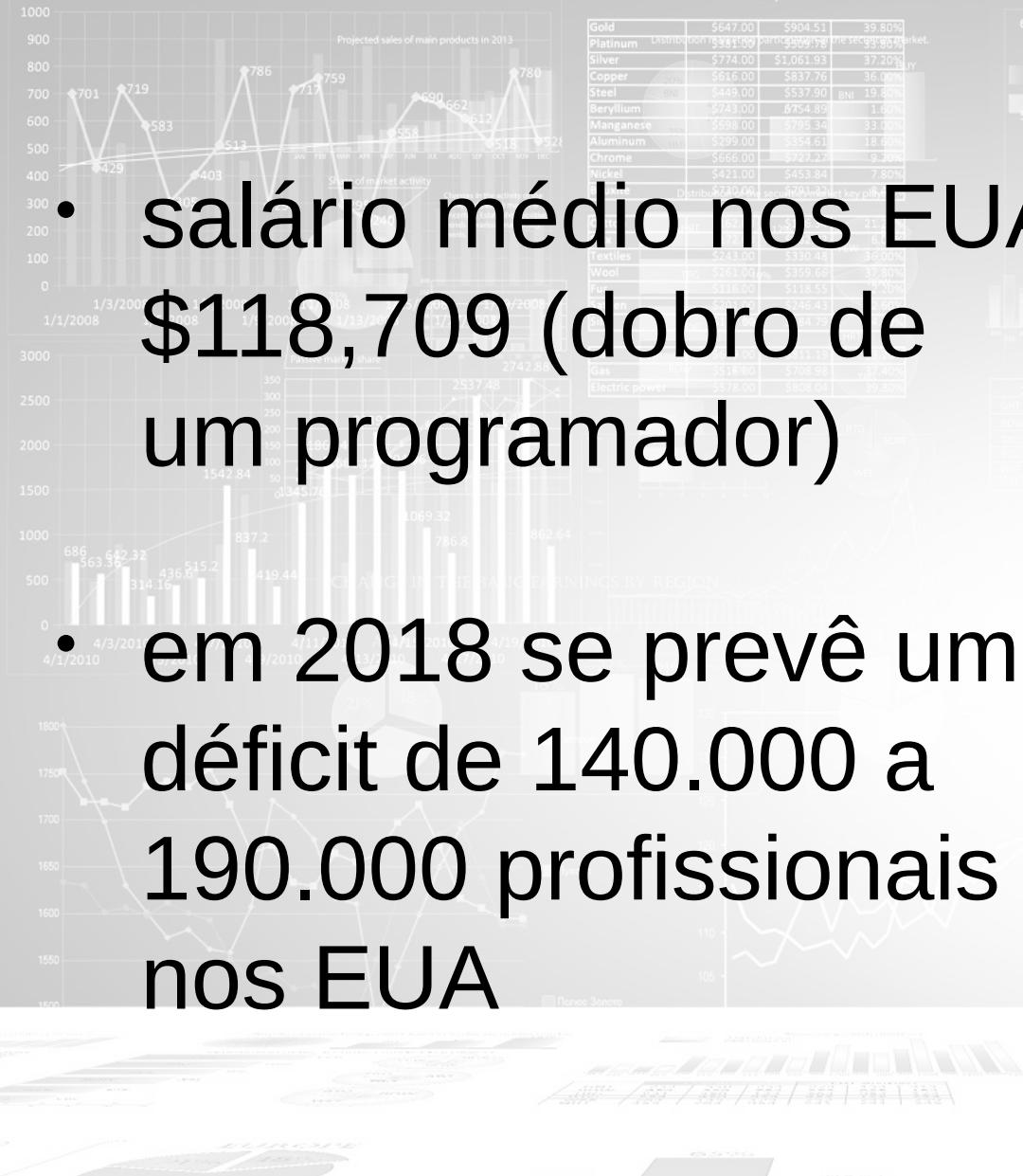
Dilúvio de Informação

- 1bi usuários conectados no facebook
(23/08/2015)
- 2bi smartphones no mundo
- 1b sites web
- 300 horas de vídeo no YouTube a cada minuto
- Google, Amazon, Microsoft and Facebook =
1,200 petabytes = 1.200.000.000.000.000
bytes = 5 pilhas de CDs até a Estação Espacial
Internacional



Cientista de Dados

- salário médio nos EUA \$118,709 (dobra de um programador)
 - em 2018 se prevê um déficit de 140.000 a 190.000 profissionais nos EUA



Para os cientistas de dados não há desemprego

Germano Lüders / EXAME



Sinhigno, do Itaú: em três anos, o número de cientistas de dados no banco passou de 15 para mais de 100 profissionais

Revista Exame

- Cerca de 2,5 bilhões de gigabytes de dados são criados diariamente, o suficiente para encher 156 milhões de iPads.
- ... É aí que entra a figura do cientista de dados, capaz de extrair informações úteis de onde impera o caos.
- “Esse profissional é alguém que consegue unir dados desconexos, apontar tendências e ajudar na tomada de decisões” (IBM Brasil)

Revista Exame

- A atual ênfase dada a esse profissional marca uma nova fase da indústria de TI.
- No Brasil, um profissional recém-formado pode ganhar até 6000 reais, enquanto os gerentes chegam a receber cerca de 20000 reais.
- “Hoje, a disputa por esses profissionais é global e deve persistir por pelo menos uma década”
- o déficit de profissionais pode chegar a 380000 vagas no mundo em 2018.

O que vamos ver

- Como obter e armazenar dados
- Como modelar os dados
- Como gerenciar os dados
- Como consultar e manipular os dados
- Como otimizar o desempenho do banco para atender às necessidades da aplicação
- E como resolver qualquer problema relacionado com dados no mundo e região!