

Banco de Dados II

Introdução

Diego Addan

DS330 - UFPR - 2023

Para hoje

Apresentação

Introdução a DS330

Referências

Introdução

Ementa e Objetivos

Site

https://www.inf.ufpr.br/dagoncalves/_disciplinas.html

Introdução

Porque BD é importante?

onde se insere na computação atual?

Buscas atuais na pesquisa

- Segurança (FER)





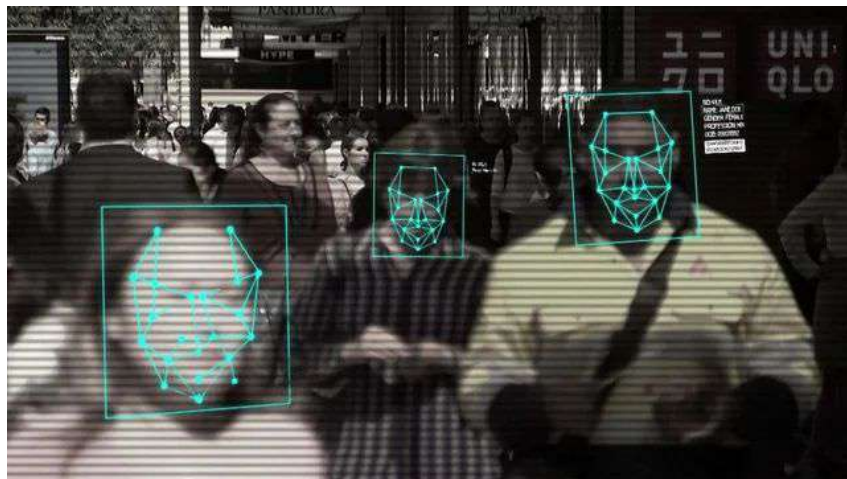
Procedia Computer Science

Volume 96, 2016, Pages 1571-1580



Face Recognition in the Wild ☆

Wajdi Dhifi, Abdoulaye Baniré Diallo  



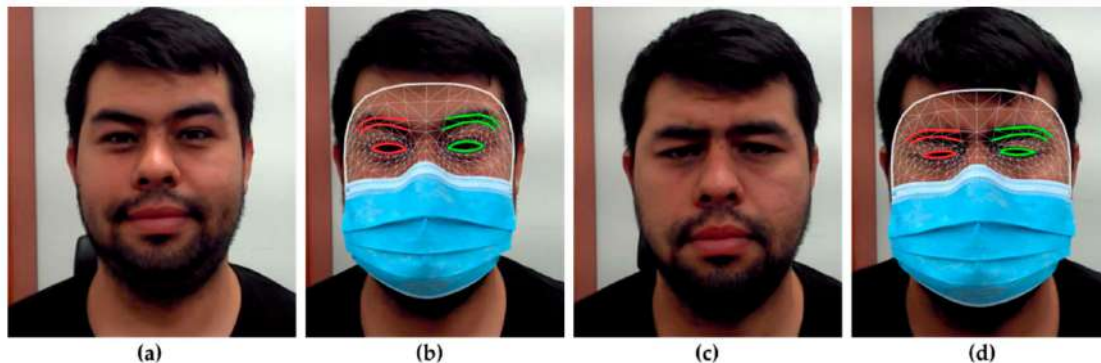
Introdução

- POI
- Displacement



3bi de variações

- Occlusion
- Morfologia



Introdução

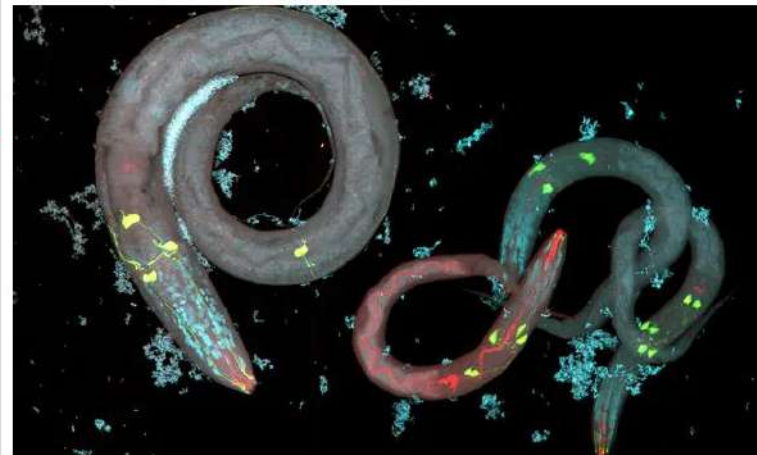
Porque BD é importante e onde se insere na computação atual

Buscas atuais na pesquisa

- Segurança (FER)
- Saúde (fonte: the guardian journal)

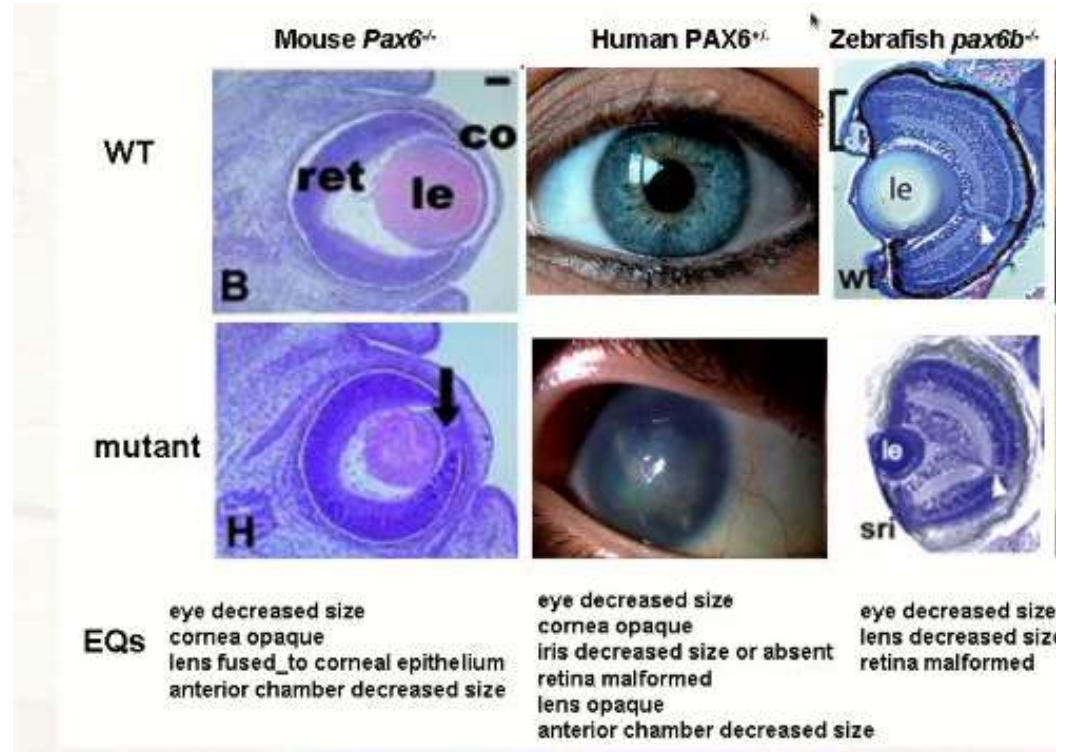
Scientists discover how to make youth last longer - in worms

Tests on roundworms show antidepressant drug can prolong life by more than 30%, but only extends young adulthood, and has no effect on later life stages



Introdução

- Doenças genéticas e genes relacionados
- Repositórios de informações



Introdução

Levantamento de dados conceituais
como:

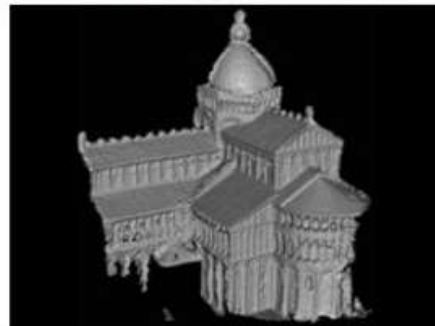
- características biológicas
- **Ações**
- Mapeamento geográfico

Relações x Volume

© Snavely, Seitz, Szeliski (2006)



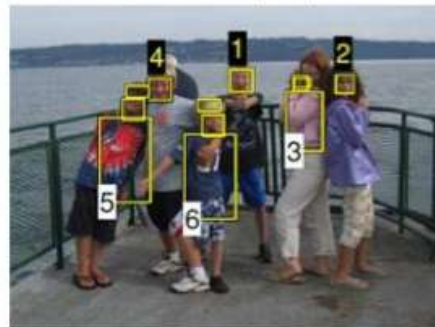
© Goesele, Snavely, Curless et al (2007)



© Sidenbladh, Black, Fleet (2000)



© Sivic, Zitnick, Szeliski (2006)



Exemplo de aplicações e algoritmos presentes em Computação Visual – SFM, Stereo Matching, Rastreamento, Reconhecimento

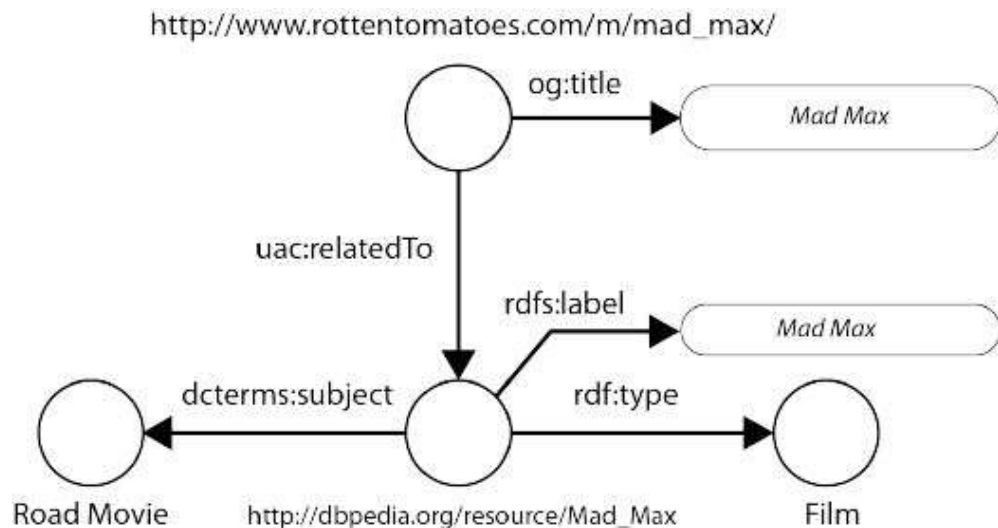
Introdução

Linked Data

Proposto por Tim Berners-Lee (2006) para estruturar e publicar dados conectados na Web por meio de hyperlinks e anotações, de maneira a permitir buscas semânticas.

Wikipedia - Infobox

DBpedia

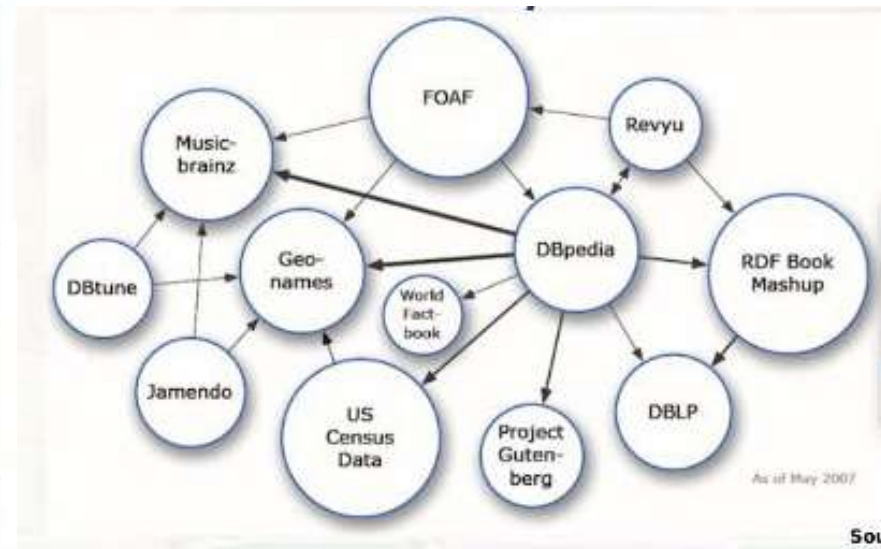


Introdução

DBpedia

Class	Instances
Resource (overall)	4,828,418
Place	967,491
Person	1,592,912
Work	552,115
Species	190,369
Organisation	317,867
Other	1,207,664

Linked Data

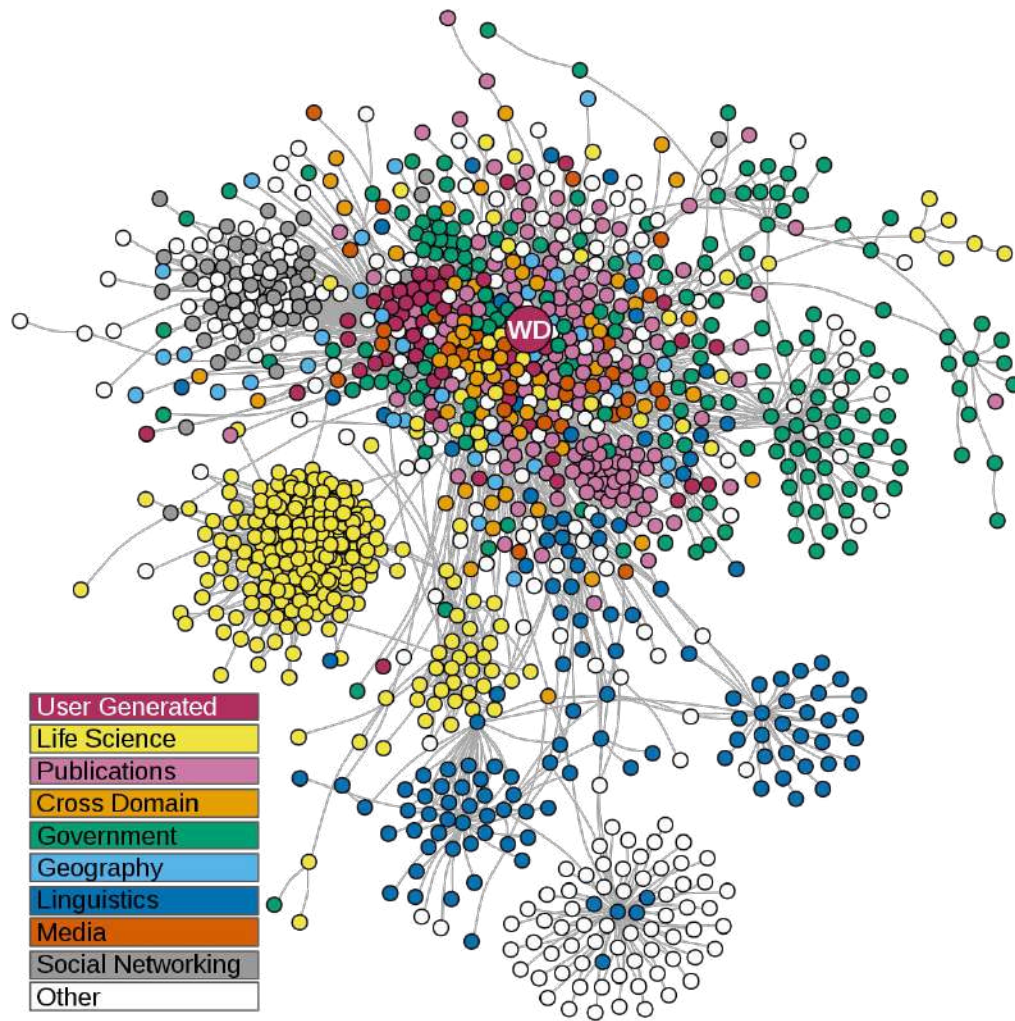


Introdução

Linked Data

2.203 datasets

19.389 links



Introdução

Bases fechadas:

ISCTE: AIM-Health

VRI: (AMR): 4 dataset

Phasic Dopamine Release

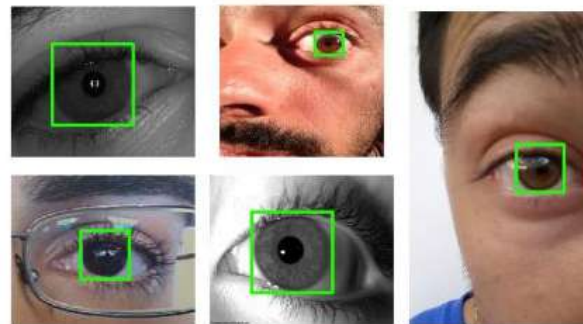
Entre **MUITAS** outras

Iris Location – Annotations

We provide the annotations (iris location) of 5 iris databases:

- BERC Mobile-iris – 500 images
- CASIA-IrisV3 Interval – 2639 images
- MobBIO Fake – 1600 images
- IIIT-Delhi Contact Lens Iris – 6570 images
- MICHE-I – 3191 images

We define the iris location task as the determination of the smallest squared bounding box of the entire region of the iris. The figure below provides some samples of these annotations.



Introdução

Apps,

Web Services

Streamings

Dinâmicas sociais



Ou seja

Aplicabilidade gigantesca:

IA, ML, contextos práticos, Federated Learning

Tratamento e análise

Poder de processamento x volume

Vamos praticar a modelagem de BD em projetos e práticas

Arquiteturas de SGBD

Desafios: Volume de dados

- Transações bancárias, interações em aplicativos, troca de dados WS

Universo gigantesco de dados e Heterogêneo

Requisições - - - **BD** - - - aplicações

SGBD: responsável pelo armazenamento, gerenciamento e disponibilização dos dados

Arquiteturas de SGBD

- 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)

Arquiteturas de SGBD

Contexto da Aplicação: OLTP x OLAP

■ OLTP

- Online Transaction Processing
- Bancos de dados “tradicionais”
- Operações de inserção, atualização e exclusão em pequenas partes do banco

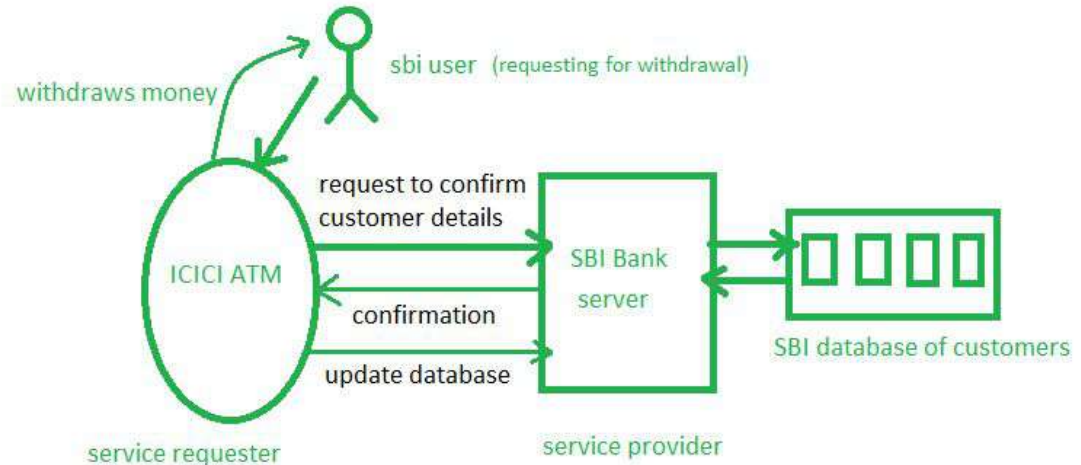
■ OLAP

- Online Analytical Processing
- Operações de extração, recuperação e análise de dados

Arquiteturas de SGBD

Exemplo de OLTP

- Gerenciamento de prontuários eletrônicos
- Sistema bancário
- Sistema bibliotecário



Arquiteturas de SGBD

Exemplo de OLAP

- **Cruzamento de dados de um survey**
- **Análise de dados em smartwatch**
- **Recomendações (Streaming, Amazon)**
- **Análise de dados para suporte a decisão**



Arquiteturas de SGBD

Dados de diferentes naturezas

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

SGBD Introdução

Aplicação exemplo

Gerenciamento de uma biblioteca

Serviços

- Cadastro de membros associados
- Registro de acervo
- Controle de empréstimos

SGBD Introdução

Aplicação exemplo

Gerenciamento de uma biblioteca

Serviços

- Cadastro de membros associados
- Registro de acervo
- Controle de empréstimos

Problemas (Aplicação cuidando diretamente de arquivos)

- Redundância não controlada
 - Repetição
 - Inconsistência
- Barreira na integração entre arquivos
- Acesso concorrente
- Integridade e recuperação

SGBD Introdução

Aplicação exemplo

- BD Unificado

Modelo que conhecemos: Relação entre os dados

Independente da aplicação - Genérico

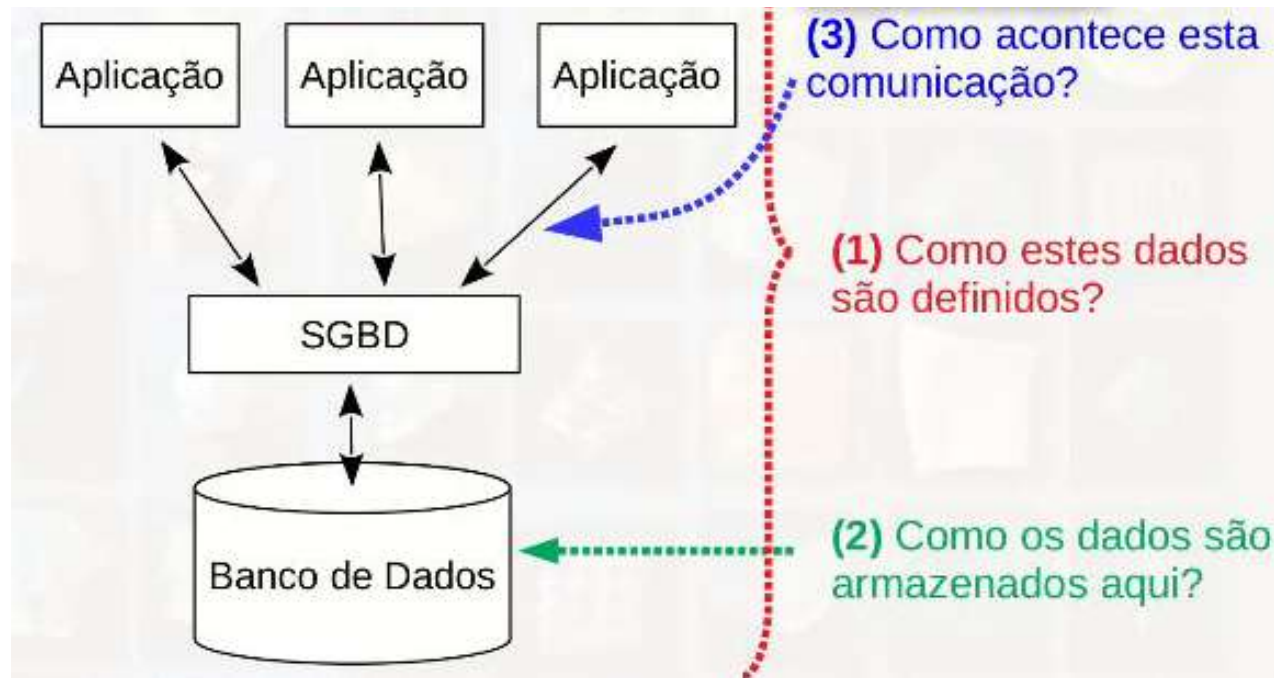
SGBD Introdução

Arquitetura de um SGBD

Modelo
Conceitual

Modelo Lógico

Modelo Físico



SGBD Introdução

Modelo Conceitual

Coleção de ferramentas conceituais

Descrevem: Dados, Relações, semântica e restrições

Universo de discurso: Representação de mundo/situação real (pode ser abstrato)

Abstração: Conceitos e relações - **UML**

SGBD: Modelo Conceitual

Análise de requisitos

- **Modelo Participativo**
- **Não computacional**

Etapa importante!

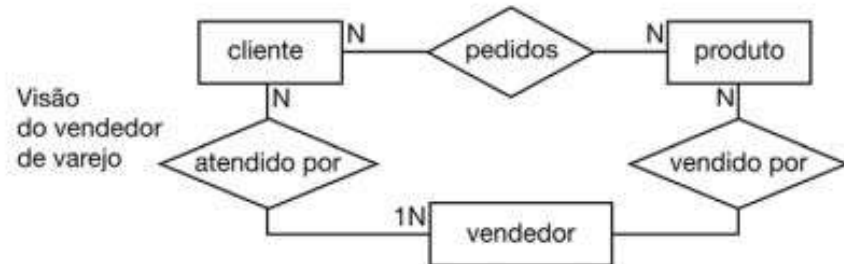
Database Life Cycle

Etapa 1 Análise de requisitos (realidade)



Etapa 2 Projeto lógico

Etapa 2(a) Modelagem de dados conceitual



SGBD: Modelo Lógico

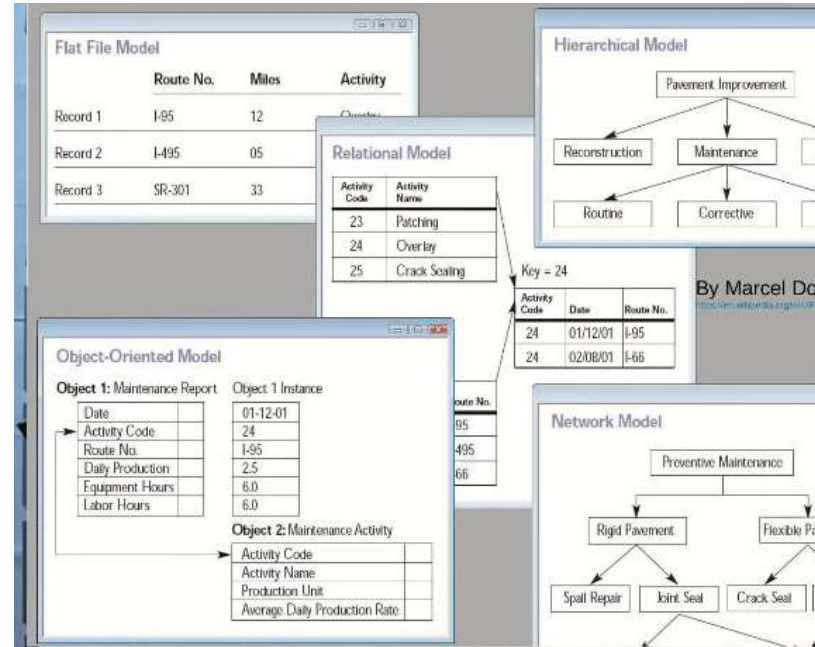
É possível estruturar a representação dos dados **independente da implementação?**

Estruturar do ponto de vista computacional

Grafo

Tabelas

Árvores



SGBD:

Modelo Lógico

Nota Fiscal



Green Leaf Design
111 Main Street
Pleasanton, CA 94566
555-555-4444
info@greenleafdesign.com

INVOICE

Invoice	00016
Date:	03/11/2009
Due Date:	04/10/2009
Balance Due \$:	1,937.01
Customer PO#:	

Bill To:
Aqua Hotel and Resort
1040 Boulevard
Anytown, CA 94558
555-555-2222 (Phone)
555-555-2225 (Fax)

Ship To:
Aqua Hotel and Resort
1040 Boulevard
Anytown, CA 94558
555-555-2222 (Phone)

Shipment Details:
Carrier: USPS
Method: Overnight Priority
Tracking#: w126a5s4321sas
Ship Date: 03/02/2009

Item	Price (\$)	Unit	Qty	Total (\$)	Tax
Design Services - Business System Includes logo, layout for letterhead, 2nd sheet, A10 envelope, and business card	3,500.00	project	1	3,500.00	8.25%
Design Services - Additional Concepts Includes thumbnail sketches for one additional concept	125.00	each	1	125.00	8.25%

Pre-tax Total: 3,625.00

Tax: 299.06

Shipping: 12.95

Total: 3,937.01

Payments: -2,000.00

Balance (\$): 1,937.01

Notes
Thanks for the work! Art hard copies and CD mailed to printer per instructions.

Terms and Conditions
Balance Due Net 30 days.

■ Considere o modelo de nota fiscal ao lado. Proponha um formato de armazenameto para representá-la.

Fonte: WorkPoint

<http://www.workingpoint.com/features/invoicing>

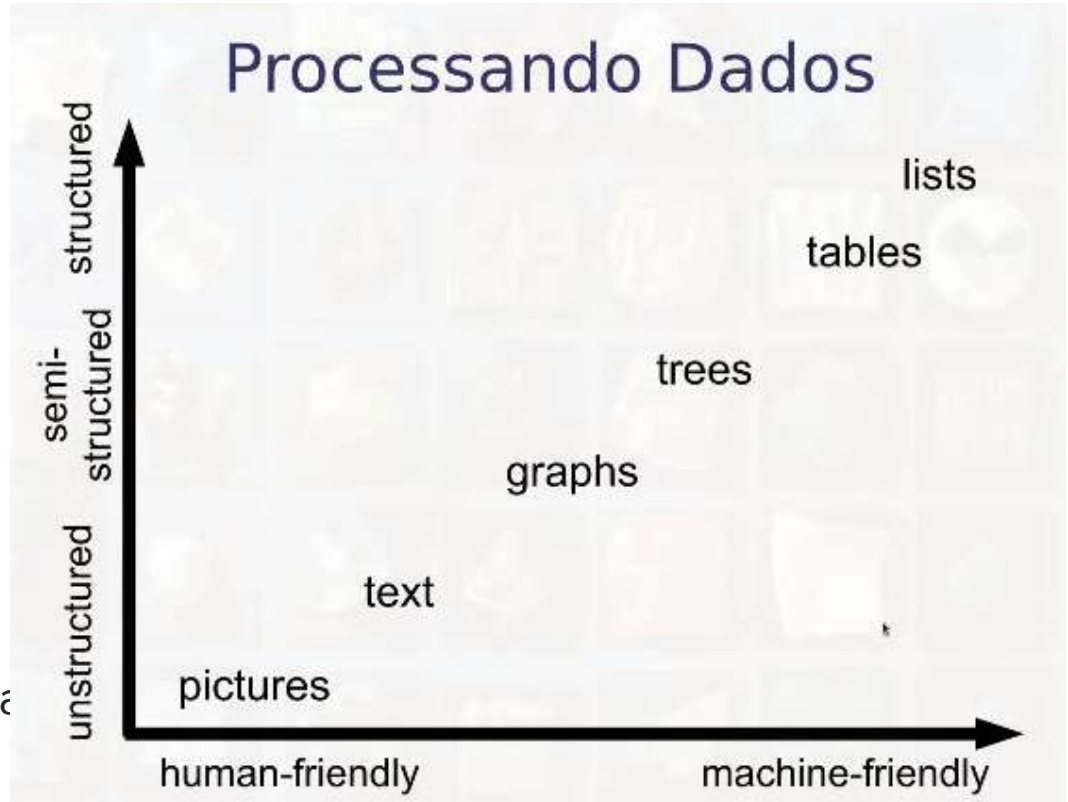
SGBD: Modelo Lógico

Estruturado

- Formato estrito (Modelo Relacional)
- Registro tem o mesmo formato

Semi-estruturado

- Itens podem ter estrutura Diversa
- Grupos de itens compartilham estruturas

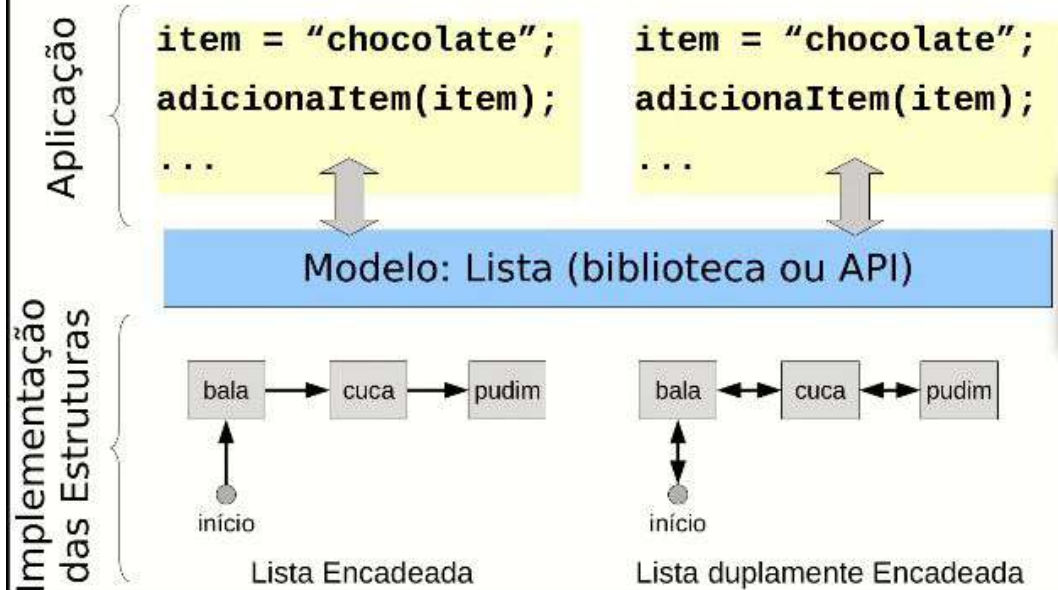


SGBD: Modelo Lógico

O **Modelo Lógico** vai permitir que pense nos dados independente da implementação.

- Modelo Lógico de Tabelas
- Modelo Lógico Hierárquico
- Modelo Lógico em Grafo

Importância dos modelos



SGBD: Modelo Lógico

Modelo Lógico de Tabelas



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

Instância

SGBD: Modelo Lógico

Previsíveis e organizadas

Ótimas para intercâmbio de dados (R, Pandas)

Ideias para análise

Tabelas são Ótimas para selecionar e recuperar itens

SEQN	RIAGENDR	RIDAGEYR	LBXIRN	LBXTIB	LBXSLDSI	LBXWBCSI	LBXLYPCT	LBXMOPCT	LBXNEPCT
31131	2	44	52	310	105	5,3	75,8	7,8	55,1
31133	2	16	55	398	101	6,6	30	5,7	63,1
31137	2	14	39	372	182	6,9	43,3	8,6	46,4
31148	2	16	103	472	136	4,8	31,3	6,7	58,5
31151	2	59	54	388	125	4,3	37,8	10,4	39,8
31152	2	27	85	449	88	9,7	17,2	3,8	76,6
31153	2	44	50	410	97	8,4	33,3	6,8	55,3
31156	2	43	84	385	134	5,3	41,8	5,6	50,9
31160	2	39	92	398	116	9	21,2	8,6	68,1
31171	2	12	61	393	190	10,4	31,3	6,1	62,1
31172	2	20	82	410	120	10,9	19,3	4,6	72,6
31177	2	18	0	0	0	9	35,3	5,2	57,7
31178	2	14	56	368	95	5,1	43,4	8,2	45,7
31183	2	33	130	293	100	6,1	20,2	12,7	61,4
31186	2	46	78	414	118	8,7	31	7,3	59
31187	2	22	56	402	116	4,9	32	5,8	61

SGBD: Modelo Lógico

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)

SGBD: Modelo Lógico

MR: Análise de dados segue com a relação entre características, valores e suas correlações.

Ex: Iris, Doenças Oftalmológicas, Reconhecimento Biométrico

No mundo real, nem todos os modelos de dados podem ser estruturados por tabelas.

SGBD: Modelo Lógico

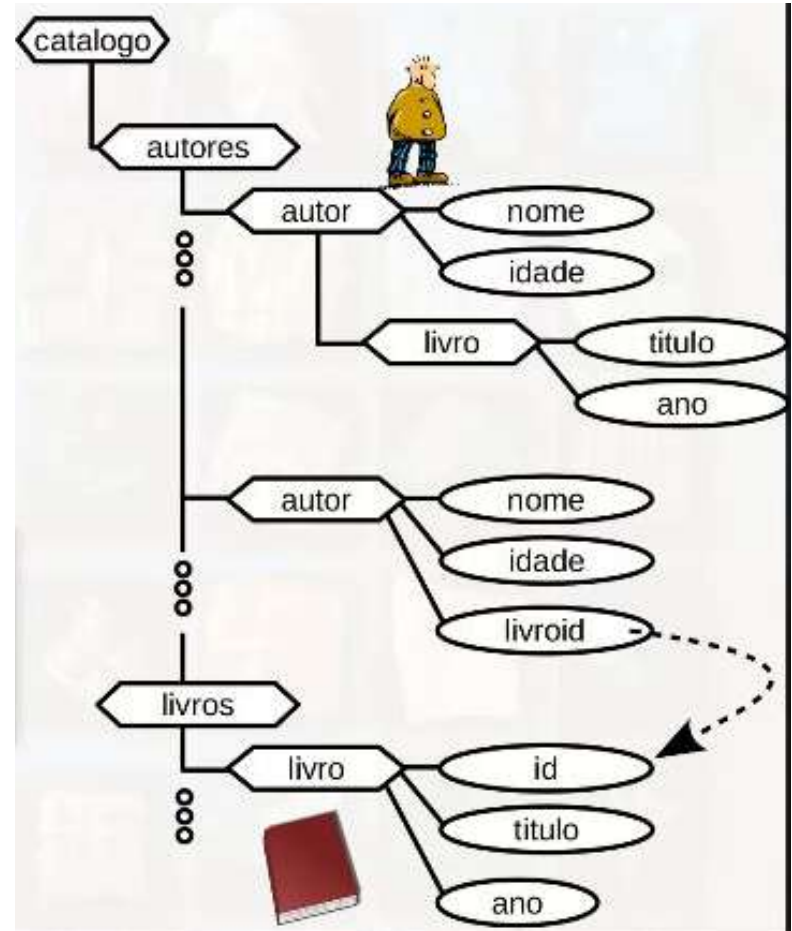
Modelo Hierárquico

Naturalmente aninhado

Nota Fiscal, XML

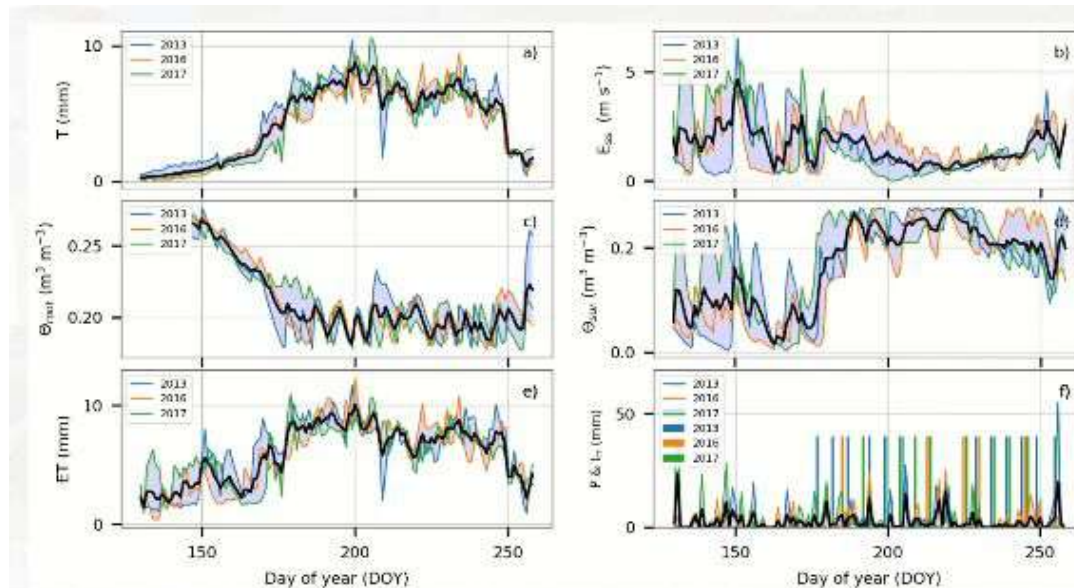
Prontuários

Semi-Estruturado: Partes do dado não necessariamente seguem um modelo



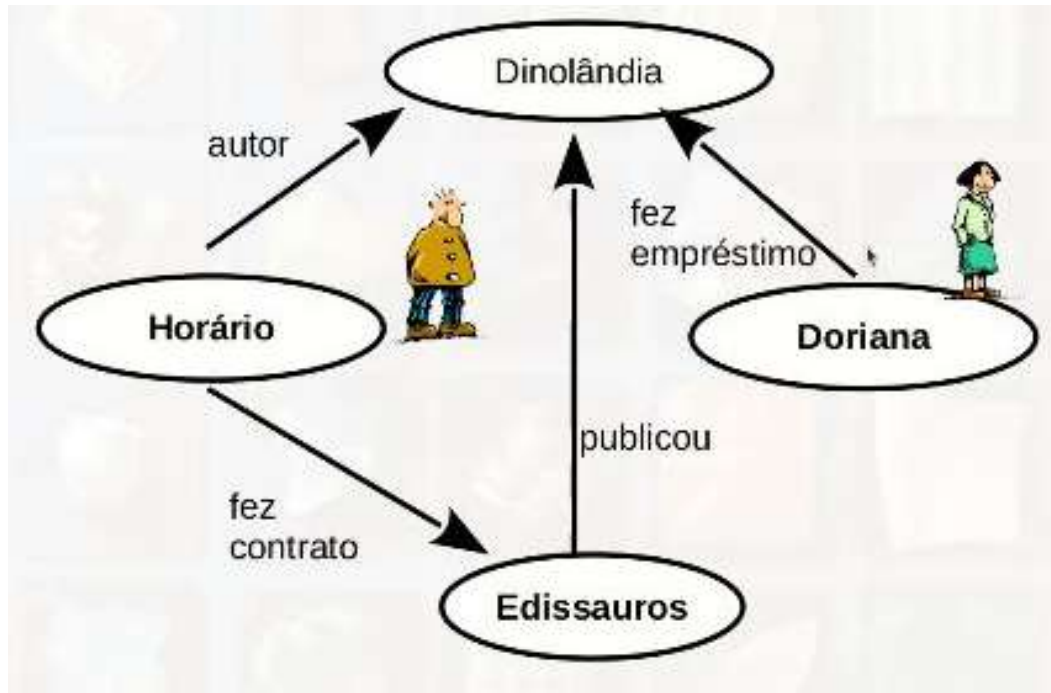
SGBD: Modelo Lógico

Modelo Hierárquico: Smartwatch, series temporais



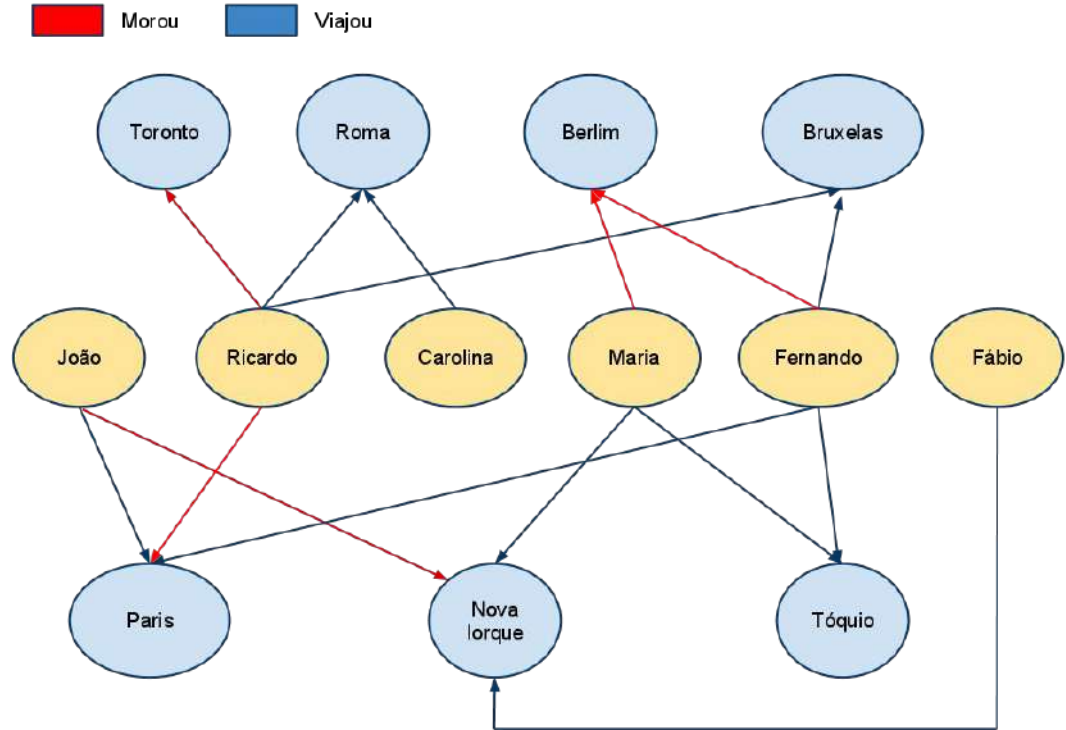
SGBD: Modelo Lógico

Modelo Lógico de Grafos



SGBD: Modelo Lógico

Modelo Lógico de Grafos



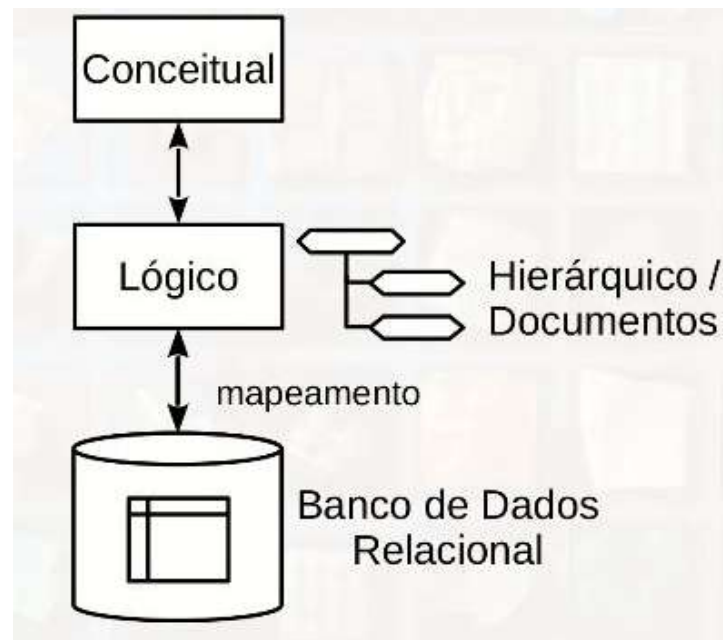
SGBD: Modelo Físico

Os sistemas tinham a preocupação do armazenamento:

Leitura, acesso (Bloco inteiro é acessado)

Mapeamento Lógico Relacional

MonetDB (Colunas)



SGBD: Modelo Físico

Otimização

Representações dos dados e armazenamento

Servidor Físico (VM)

Nuvem

SGBD: NoSQL

Sistemas com representações físicas diferentes do Relacional

SGBD: NoSQL

Sistemas com representações físicas diferentes do Relacional

Banco de dados de documentos /JSON



Amazon DocumentDB

SGBD: NoSQL

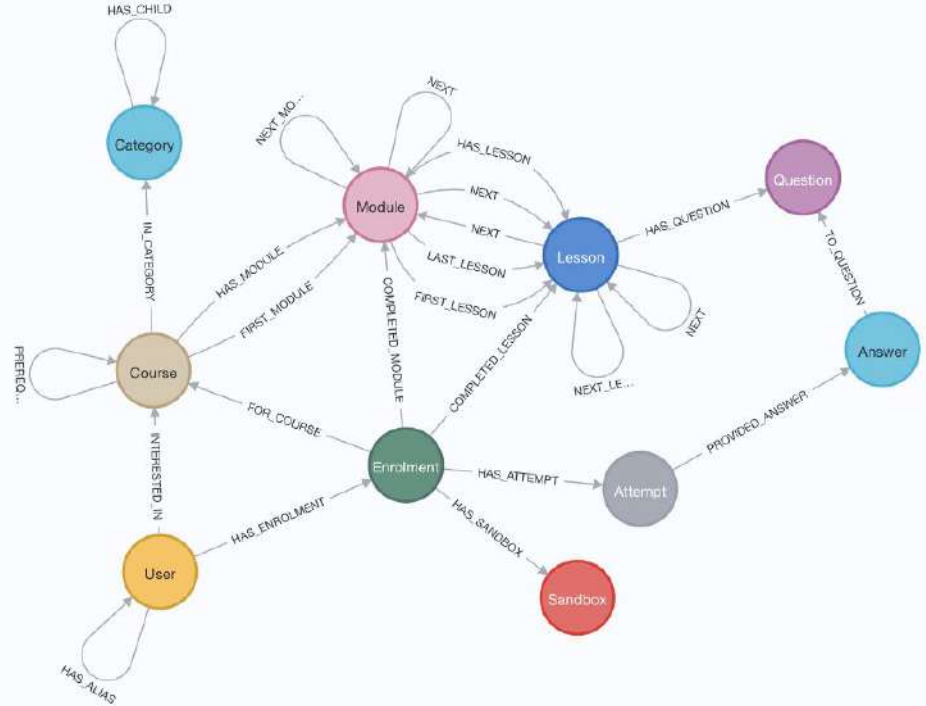
Semi-Estruturados

Estruturas mais flexíveis

Próximo do armazenamento web

SGBD: Banco de dados de Grafos

Semi-Estruturados

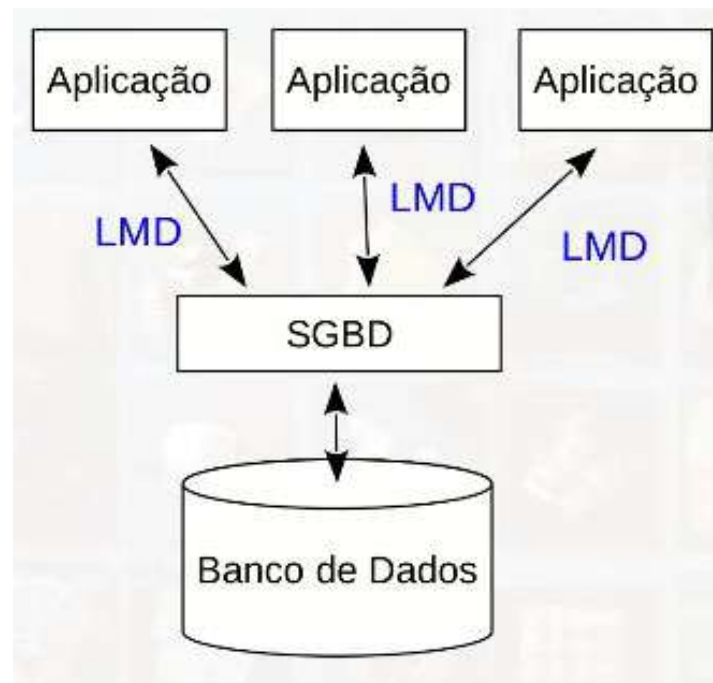


SGBD: Arquitetura

E como acontece a comunicação entre a aplicação e o SGBD?

Linguagem de Manipulação de Dados

“O quê” ao invés de “como”



SGBD: Arquitetura

Vantagens de um SGBD

- Independência de dados
- Acesso eficiente
- Tempo reduzido no desenvolvimento de aplicações
- Segurança e integridade de dados

(Ramakrishnan,



SGBD: Arquitetura

Tipos de arquitetura:

- **Local (VM)**
- **Cliente-Servidor (AWS)**
- **Duas Camadas (Ex. API Rest)**
- **Três Camadas (Loja Virtual)**
- **Paralelo ou Distribuido (FL)**

SGBD: Arquitetura

Ferramentas



Sumarizando

O volume de dados justifica a evolução dos SGBDs

Diferentes tipos de arquitetura

Bases estão mudando: Originalmente pensadas centralizadas agora são distribuídas

Modelagem semântica, ontológica e novas abordagens (noSql)

Referências

1. SILBERSCHATZ, A.; KORTH, H. F.; SUDARSHAN, S. Sistema de Banco de Dados. Rio de Janeiro: Elsevier, 2012.
2. ELMASRI, Ramez; NAVATHE, Shamkant B. Sistemas de Banco de Dados. 6. ed. São Paulo: Pearson, 2011.
3. BEAULIEU, Alan. Aprendendo SQL. Tradução Edgard Batista Damiani. São Paulo: Novatec Editora, 2010