



Universidade Federal
de Campina Grande



Banco de Dados I

Prof. Cláudio de Souza Baptista, Ph.D.
Laboratório de Sistemas de Informação – LSI
UFCG

Dados em todo o lugar

- “The world is increasingly **driven by data...**”
- Filosofia: **More data => More Value!**
- Este curso ensina como **manipular e gerenciar dados!!!!**

Introdução

- O que é um banco de dados?
- Porquê usar banco de dados?
- Pra quem banco de dados?

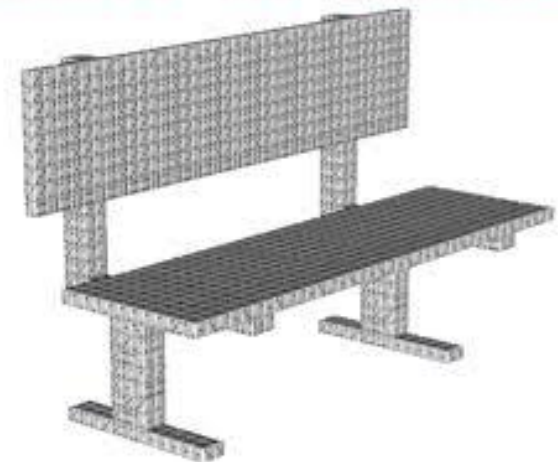


O que é um Banco de Dados?

- Conjunto de dados, inter-relacionados, normalmente gerenciados por um SGBD, que provê o armazenamento, indexação e recuperação eficiente (queries) dos dados.



Banco de Dados



Dados

- Estruturados
- Semi-estruturados
- Não estruturados



Dados Estruturados

- Esquema definido a priori
- Dados tipados e bem comportados

- Exemplo

```
CREATE TABLE CLIENTE (  
    cod int,  
    nome varchar (50),  
    endereco varchar (100),  
    primary key (cod)  
)
```

Dados Não Estruturados

- São dados em que não há nenhuma estrutura os definindo.
- São normalmente armazenados em arquivos ou em colunas do tipo BLOB ou CLOB em SGBDs
- Exemplos:
 - um texto em pdf
 - uma imagem em jpg
 - um video ou áudio

Dados Semi-Estruturados

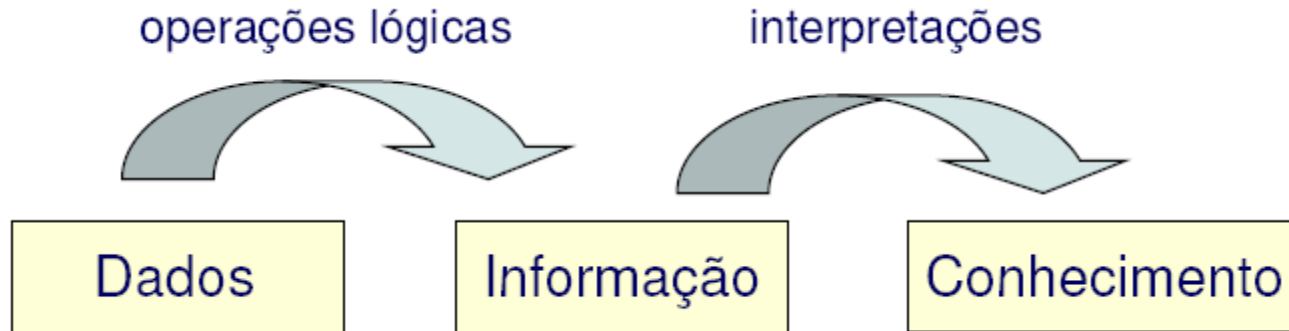
- Dados heterogêneos e irregulares, auto-descritivos
 - São dados em que uma parte há estrutura e em outra não há nenhuma estrutura definida.
 - O esquema da parte estruturada está contido junto com o dado
- Ex.: Um email
 - Parte estruturada:
<to>, <subject>, <date>, etc.
 - Parte não estruturada: o corpo da mensagem

Dados Semi-Estruturados

■ JSON:

```
{
  "nome": {
    "primeiro nome": "João",
    "sobrenome": "da Silva",
    "título": "Dr."
  },
  "endereço": {
    "rua": "Rua das Araucárias",
    "bairro": "Bairro do Salgado",
    "num": "201",
    "cidade": "Campina Grande",
    "complemento": "apt 1406"
  }
}
```

Dado, Informação, Conhecimento



- **Dado** é algo bruto; é a matéria-prima da qual podemos extrair informação.
- **Informação** é o dado processado, com significado e contexto bem definido.
- **Conhecimento** é o uso inteligente da informação; é a informação contextualizada e utilizada na prática.

Modelo DIKUW

Wisdom - Buy John a dog bowl for his birthday and he'll be very happy

Understanding - John's birthday is on April 27th. If John Smith likes Dogs then he probably has one

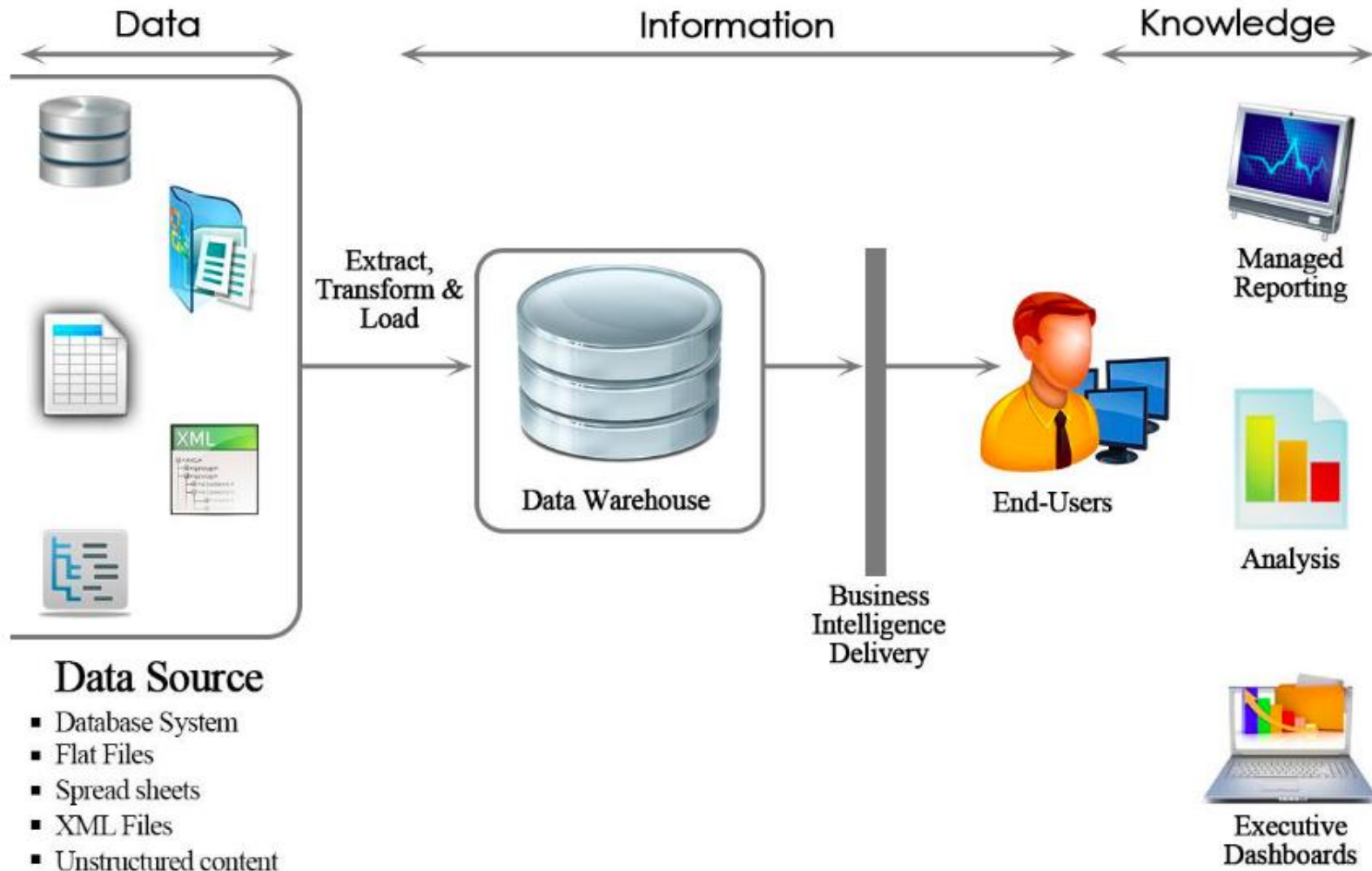
Knowledge - 1979-04-27 is John Smith's date of birth, and John Smith likes Dogs

Information - 1979-04-27 is a Date, John Smith is a Person, Dog is an Animal (data in context)

Data - "19790427", "John Smith", "Dog" (raw groups of symbols)



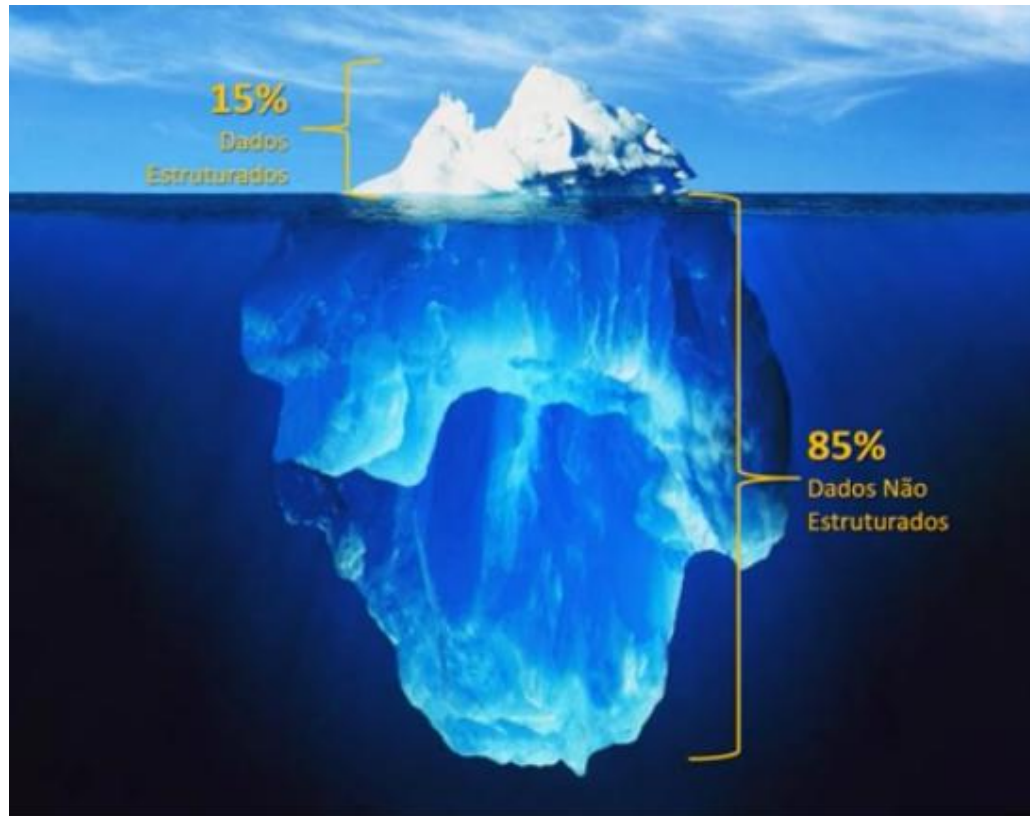
Dado, Informação, Conhecimento



Os 5 C's dos Dados

- **Clean**: limpo
- **Consistent**: correto
- **Compliant**: de acordo com a regras de negócio e leis
- **Current**: atualizado
- **Comprehensive**: completude / escopo da empresa / organização / negócio

Situação dos dados atualmente



O Grande desafio

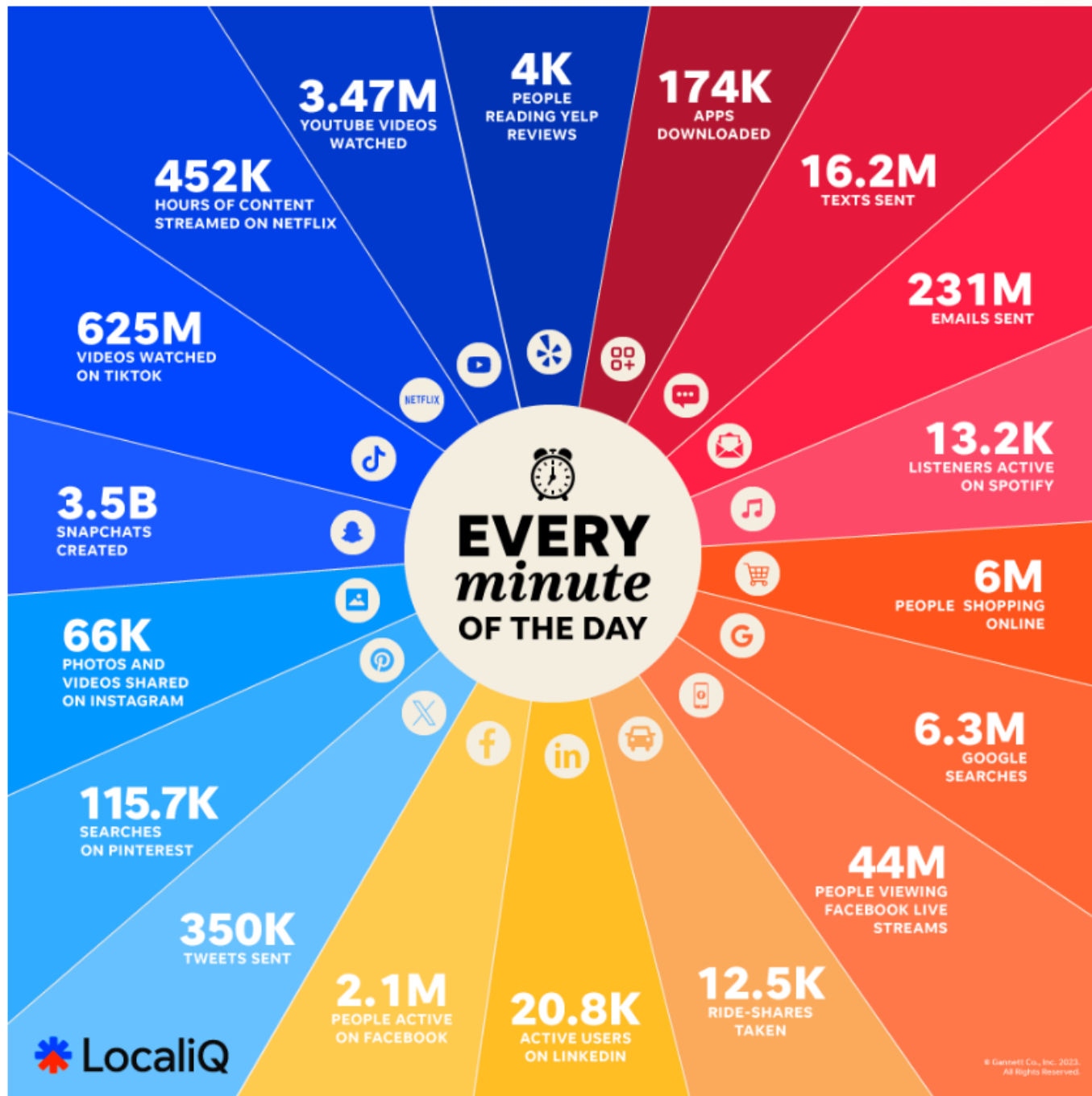


Big Numbers

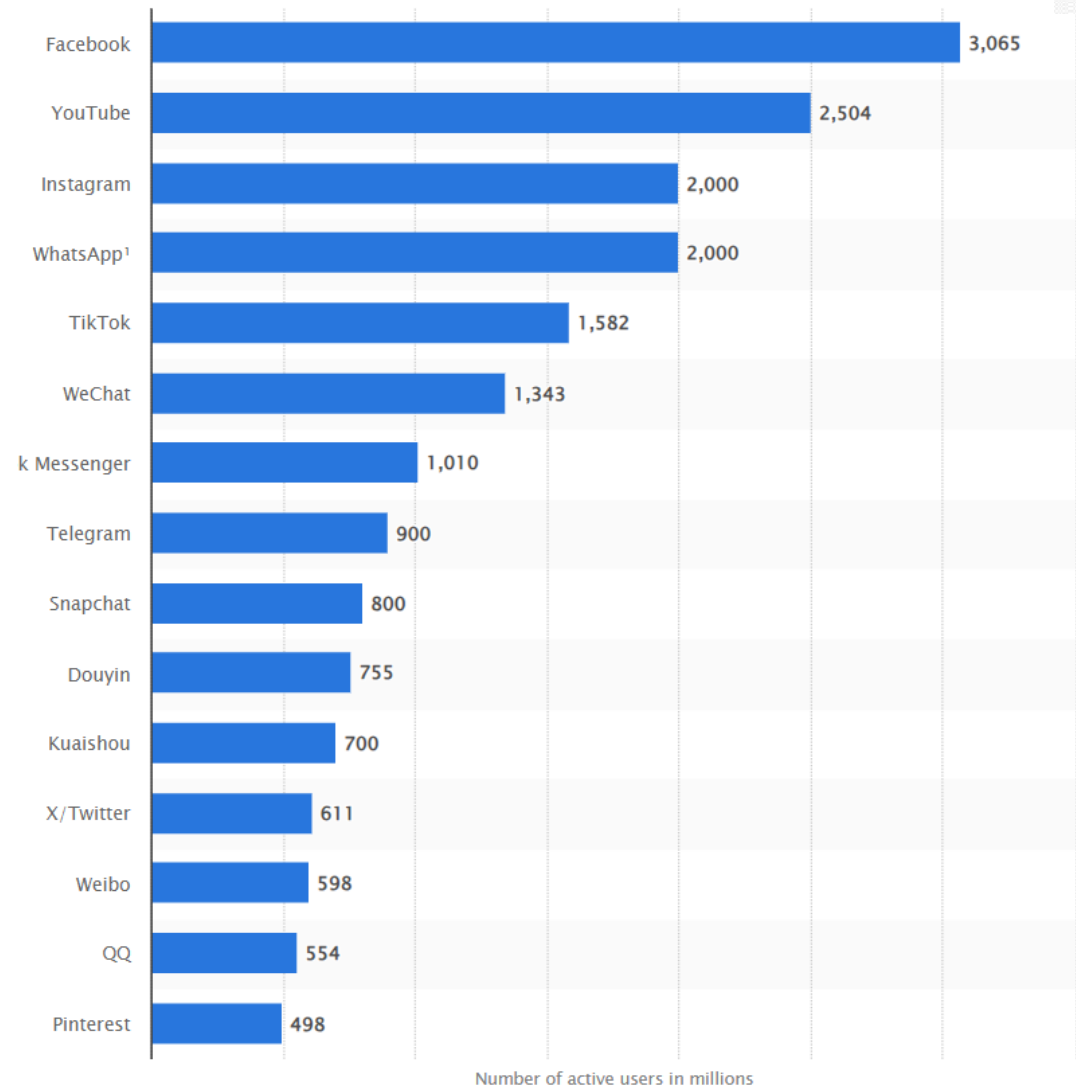


For Starters..

Quantities of Bytes		
BIT	=	A BINARY DIGIT SET TO EITHER A 1 OR 0
BYTE	=	8 BITS
KB KILOBYTE	=	1,000 BYTES
MB MEGABYTE	=	1,000,000 BYTES
GB GIGABYTE	=	1,000,000,000 BYTES
TB TERABYTE	=	1,000,000,000,000 BYTES
PB PETABYTE	=	1,000,000,000,000,000 BYTES
EB EXABYTE	=	1,000,000,000,000,000,000 BYTES
ZB ZETTABYTE	=	1,000,000,000,000,000,000,000 BYTES
YB YOTTABYTE	=	1,000,000,000,000,000,000,000,000 BYTES

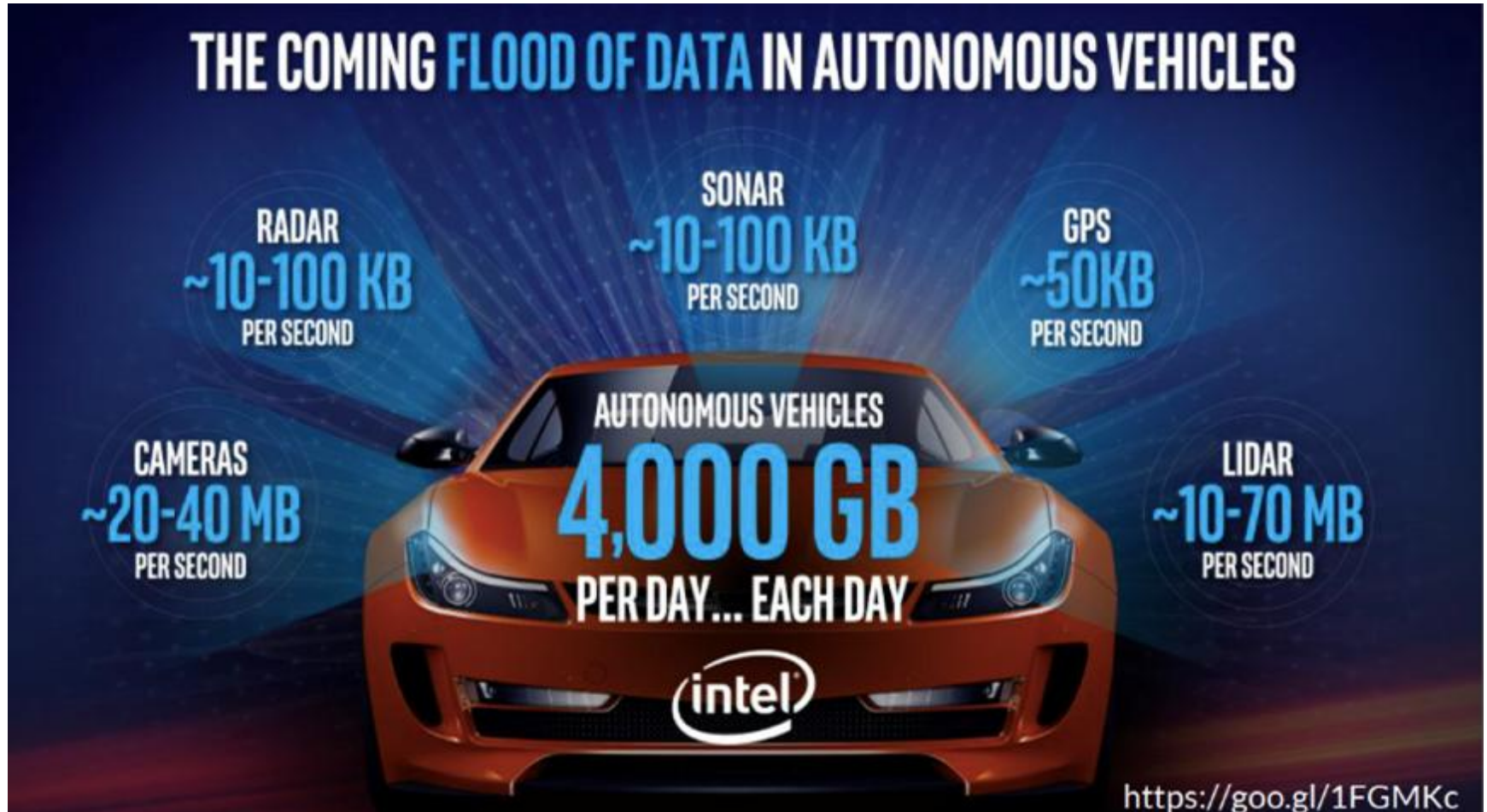


Mídias Sociais



Usuários Ativos de **Redes Sociais** -
jan/2024 em Milhões

IOT: sensors and sensors ...



Forças que guiam o crescimento dos dados

- Sensores e relatórios ubíquos:
 - Câmeras, mobile computing, blogs ...
- Alta colaboração em projetos de ciência
- Filosofia:
 - Mais dados
 - Mais valor (\$\$\$\$)
- Tecnologia
 - barata, escalável



Previsão de crescimento

■ Segundo a IDC:

- 175 zettabytes de dados serão criados até 2025
- A produção de dados tem crescido 10x na última década
- 90% dos dados de todo o mundo foram criados nos últimos 10 anos



Problemas

- 2/3 dos dados nunca são analisados
- Cerca de 70% das organizações não são capazes de agregar valor tangível e mensurável dos dados
- 90% dos empregados declaram que seus trabalhos atrasam porquê os dados não são confiáveis

Evolução da Tecnologia



IoT – Internet das Coisas



Mais de 50 Bilhões de dispositivos conectados

Outros fatores para Big Data

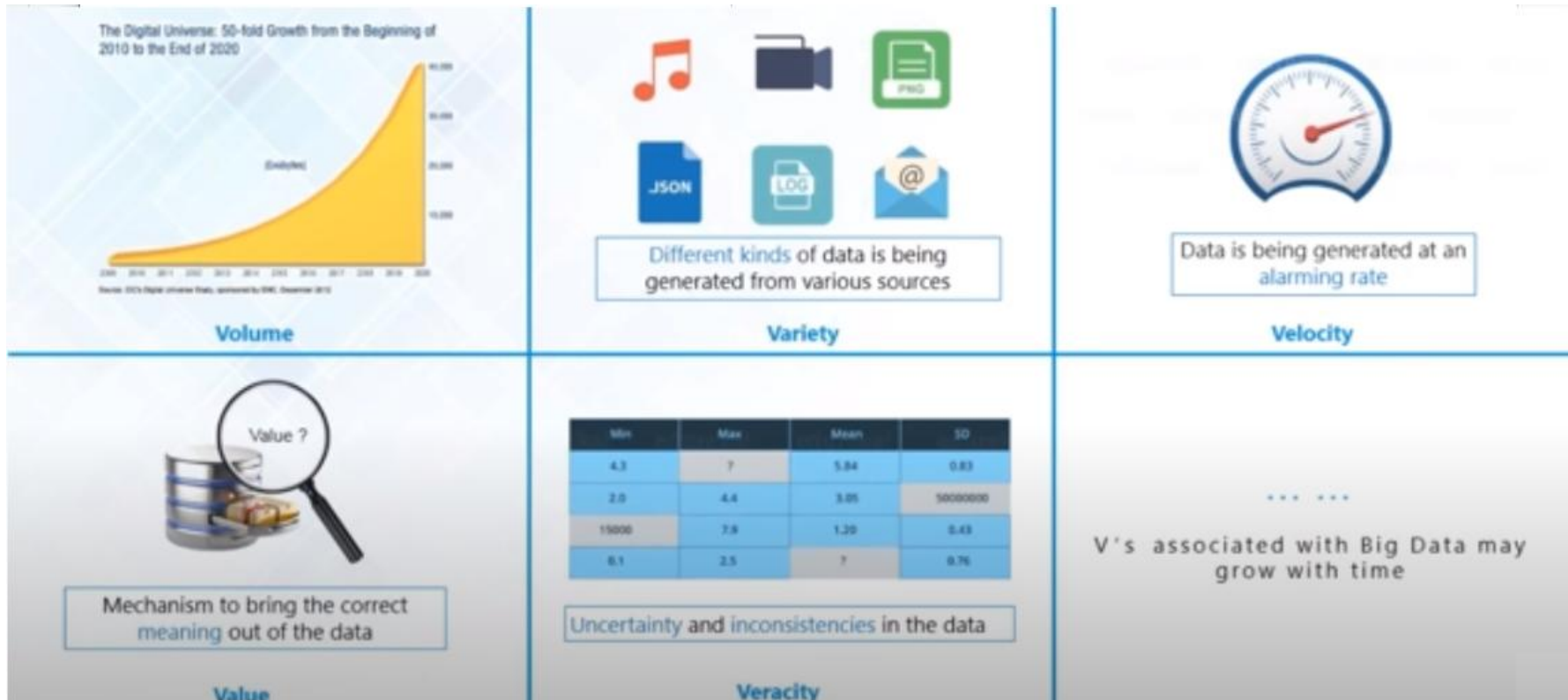


Big Data - Definição

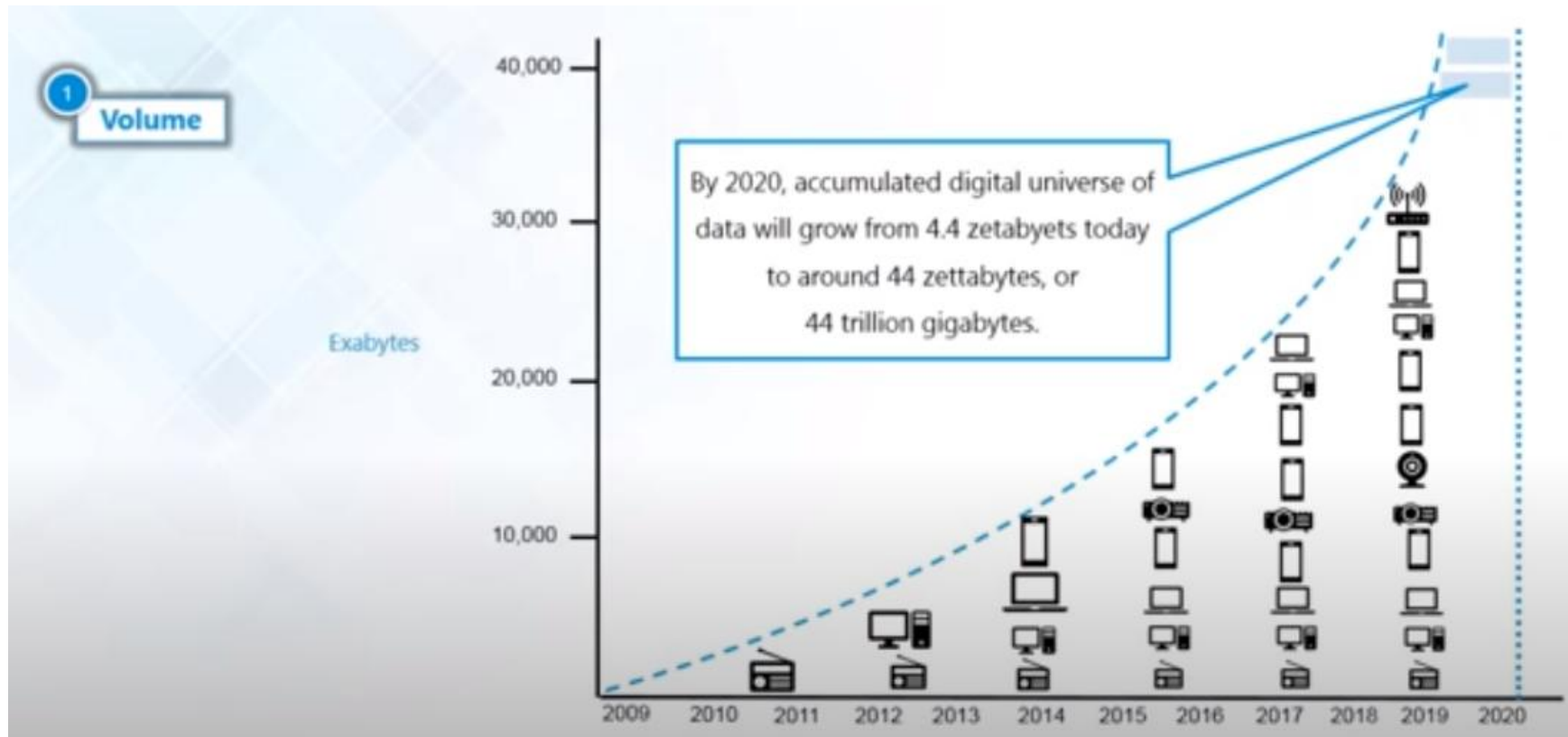
- Coleção de dados heterogêneos, em grande volume e complexo que é difícil de ser processado em aplicações e banco de dados tradicionais



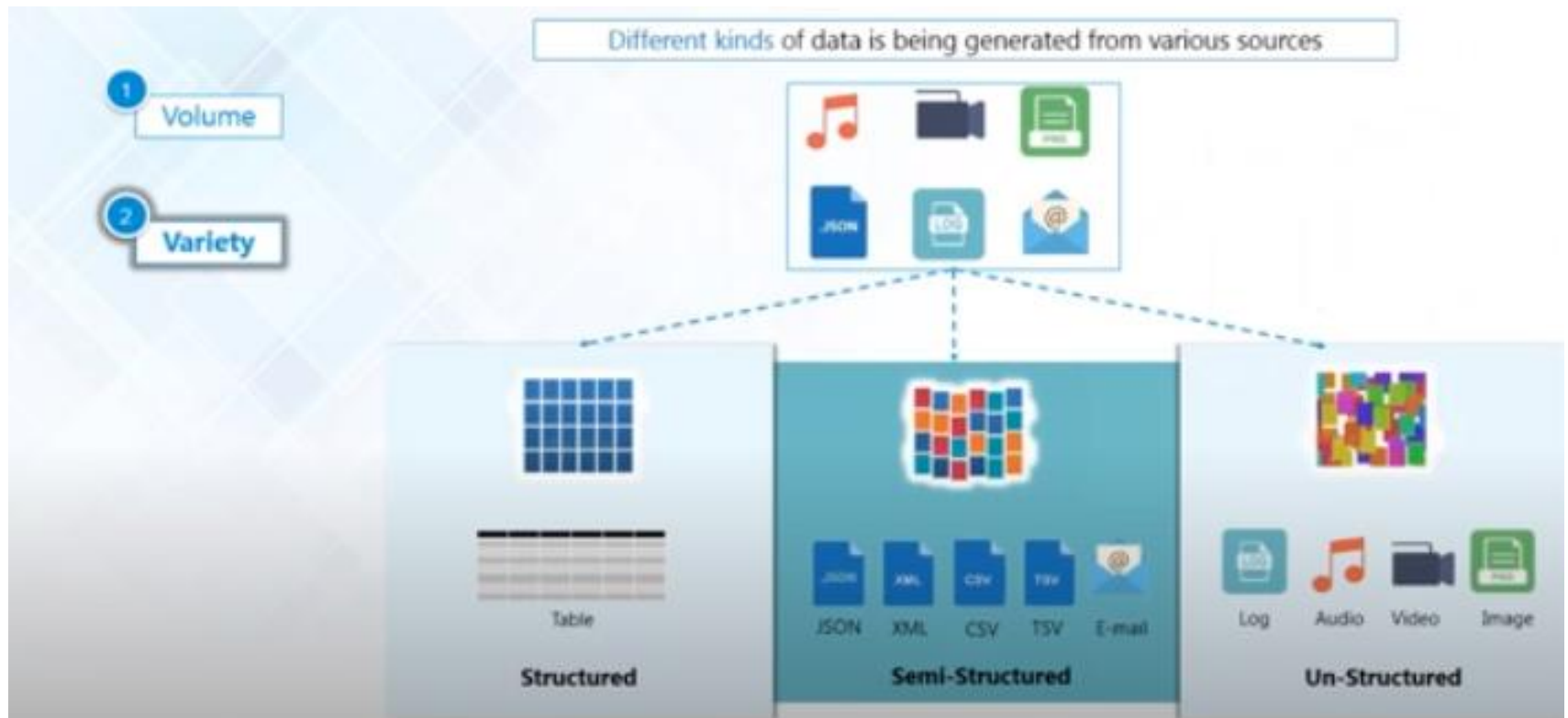
Big Data – 5 V's



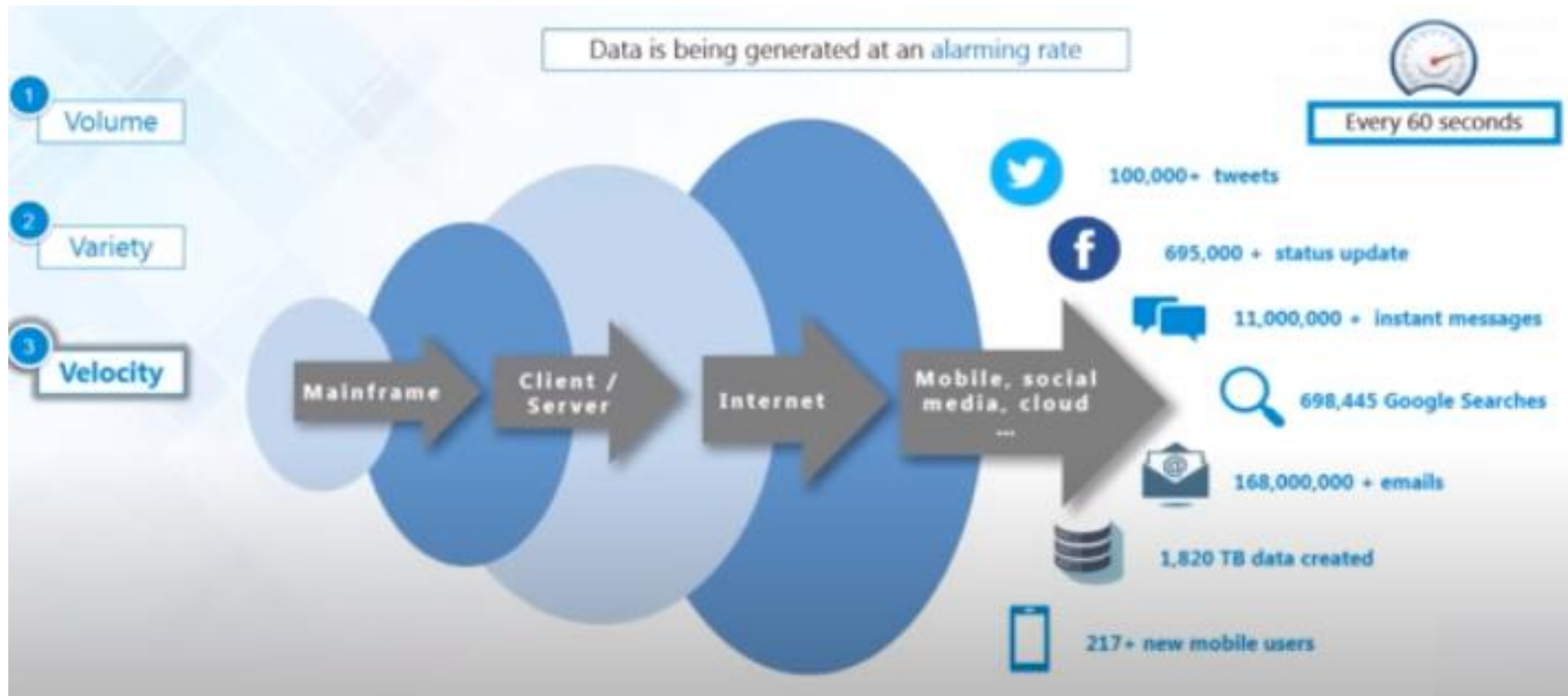
Volume



Variety



Velocity



Value

- 1 Volume
- 2 Variety
- 3 Velocity
- 4 Value

Mechanism to bring the correct meaning out of the data



Veracity

1	Volume				
2	Variety				
3	Velocity				
4	Value				
5	Veracity				

Min	Max	Mean	SD
4.3	?	5.84	0.83
2.0	4.4	3.05	50000000
15000	7.9	1.20	0.43
0.1	2.5	?	0.76

Uncertainty and inconsistencies in the data

Ex. IBM Smart Meter



Managing the large volume and velocity of information generated by short-interval reads of smart meter data can overwhelm existing IT resources

96 million reads per day
for every million meters

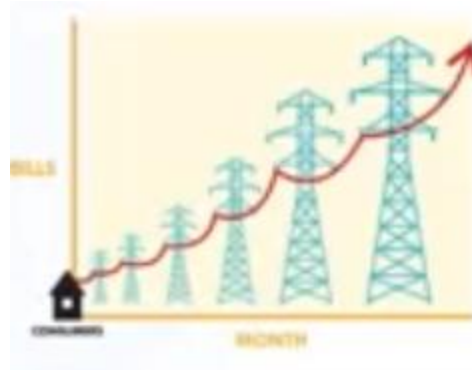


Big Data generated
by Smart Meter



Ex. IBM Smart Meter: Como Big Data pode ser analisado

Antes de analisar Big data



Depois de analisar Big data



Alto consume durante
horário de pico



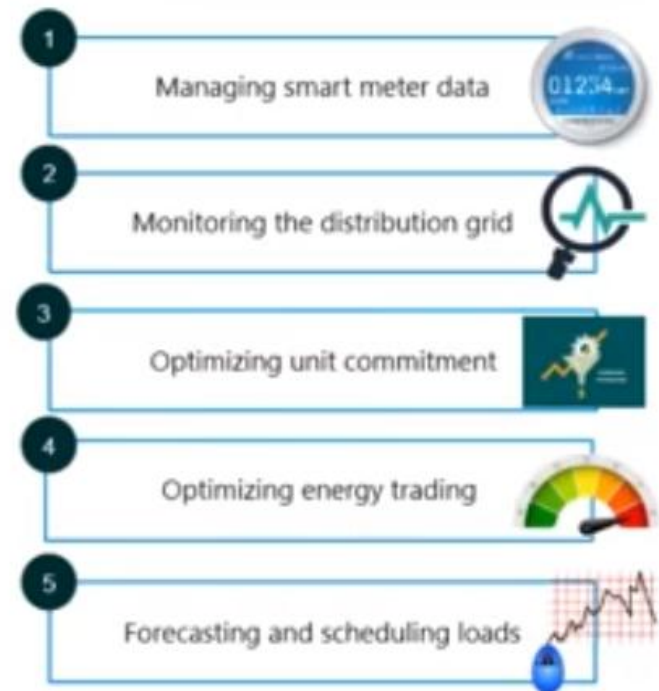
baixo consumo for a do
horário de pico

Insight: colocar máquinas de alto consume de energia para trabalharem for a do horário de pico

Ex. IBM Smart Meter Solution



IBM Solution



Big Data

- Google: a empresa boazinha !?!?!? Será?

- Busca de graça
- Email de graça: gmail
- Fotos de graça:
- Vídeos de graça: youtube
- Navegação de graça: Mapas de graça
- Armazenamento de graça (google drive)
- Plataforma de comunicação de graça
- Serviços de Escritório de graça (doc, planilha, apresentação)
- etc., etc. Etc.



Big Data

- Google: a empresa que será?
 - Busca de graça
 - Email de graça
 - Fotos de graça
 - Vídeos de graça
 - Navegação de graça
 - Armazenamento de drive
 - Plataforma de graça
 - Serviço de tradução (doc, planilha, apresentação)
 - etc., etc.



Big Data



Acre 9 de março de 2021

Mudanças na mobilidade

COVID-19: Relatório de mobilidade da comunidade

Varejo e lazer

+20%

comparado a valor base



Mercados e farmácias

+72%

comparado a valor base



Parques

+14%

comparado a valor base



Estações de transp. público

-11%

comparado a valor base



Locais de trabalho

+35%

comparado a valor base



Residencial

+3%

comparado a valor base



<https://www.google.com/covid19/mobility/>



SGBD

- Sistema de Gestão de Banco de Dados provê:
 - armazenamento e manipulação de forma eficiente, robusta, conveniente, segura, e concorrente de quantidades massivas de dados.

Porquê usar SGBD?

- Armazenar e consultar dados em **sistema de arquivos** tem várias limitações:
 - Redundância de dados e inconsistência
 - Formatos de arquivos múltiplos
 - Duplicação de informação em arquivos diferentes
 - Dificuldade em acessar os dados
 - Precisa codificar para acessar os dados
 - Falta de indexação



Porquê usar SGBD?

- Problemas com integridade
 - Modificar um arquivo (ou campo), e não mudar os campos/arquivos dependentes
 - Restrições de integridade não garantidas (saldo da conta ≥ 0)
- Acesso concorrente a múltiplos usuários
 - Vários usuários precisam acessar/atualizar os dados ao mesmo tempo
 - Acesso concorrente sem controle pode levar a inconsistências (Ex. 2 pessoas acessando a mesma conta corrente ao mesmo tempo)



Porquê usar SGBD?

- Problemas de segurança
 - Difícil prover acesso a uma parte dos dados e não a tudo
- Recuperação de falhas
 - Enquanto os dados são modificados o sistema pode cair, levando à inconsistência
- Problemas de manutenção
 - Evolução de esquema difícil



Pra quem? Aplicações Tradicionais de SGBD

- Sistema bancário, reservas de passagens, lojas, hotéis, restaurantes
- Características
 - Dados simples e bem estruturados
 - Sem operações ou relacionamentos complexos
 - Tipos de dados simples
 - Consultas e relatórios simples

Pra quem? Aplicações Tradicionais de SGBD

- Sistema bancário, reservas de passagens, lojas, hotéis, restaurantes
- Características
 - Dados simples e bem estruturados
 - Sem operações ou relacionamentos complexos
 - Tipos de dados simples
 - Consultas e relatórios simples
- Dados os ingredientes acima RDBMS é a ferramenta perfeita!

SGBD tradicionais

- Propriedades essenciais a uma Transação



Um pouco de história (1)

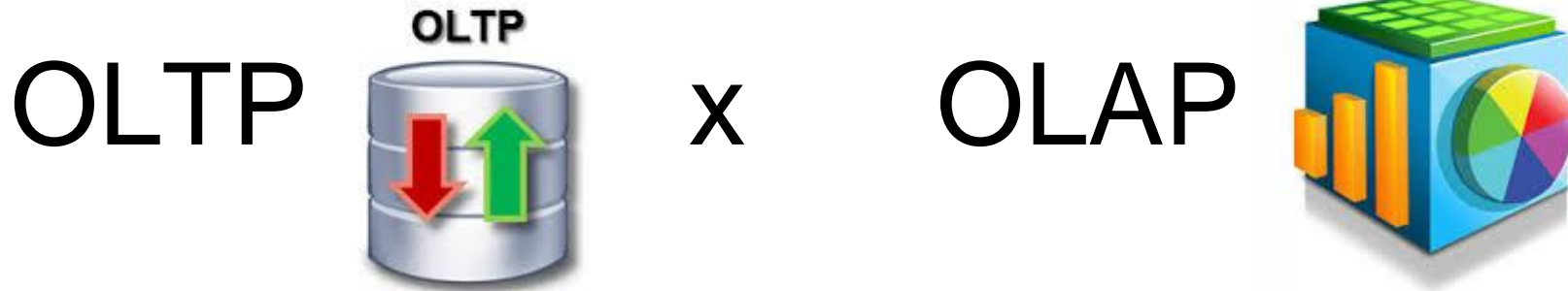
- 60's:
 - Sistemas de arquivo
 - Banco de Dados em Rede (TOTAL)
 - Banco de Dados Hierárquico (IMS)
- 70's:
 - Banco de Dados Relacional (R, Ingres)

Um pouco de história (2)

- 80's:
 - Oracle, DB2, Sybase, Informix
 - Padronização do SQL
 - BD Distribuídos
 - SGBD OO
- 90's:
 - SGBDOR
 - Produtos universais e caros
 - Conectores para Internet; dados espacial, temporal, e multimedia; bd ativos
 - Paralelismo

Um pouco de história (3)

- 2000's:
 - Oracle, IBM e Microsoft são os Top 3
 - Surgimento de SGBD para XML (Tamino, MarkLogic)
- 2010's:
 - Open source – BD para todos (Postgresql, MySQL)
 - NoSQL – sem ACID + estruturas complexas + sem esquema
 - NewSQL – SQL + NoSQL
 - Novas características: GraphDB, In memory DB, Array DB, MPP, Distributed DB, Cloud DB (VM / DBaaS), Autonomic DBMS
- 2020's: VectorDB, massivo uso de CloudDB



- Online Transaction Processing (OLTP):
 - Transações de curta duração
 - Operações repetitivas
 - Consultas exatas
- Online Analytical Processing (OLAP):
 - Consultas de longa duração
 - Análise multidimensional
 - Joins complexos
 - Consultas exploratórias

OLTP x OLAP

- BD tradicionais são projetados e ajustados para suportar operações do dia-a-dia:
 - Assegurar acesso concorrente e rápido aos dados
 - Processamento de transações e controle de concorrência
 - Foco na consistência dos dados (update online)
 - Conhecidos como BD Operacional ou Online Transaction Processing (OLTP)
- Características de OLTP DB:
 - Dados detalhados
 - Não incluem dados históricos
 - Altamente normalizados
 - Desempenho pobre em consultas complexas incluindo joins e agregação

OLTP x OLAP

- Análise de dados requer um novo paradigma: Online Analytical Processing (OLAP)
 - Consulta OLTP típica: pedidos pendentes do cliente Pedro
 - Consulta OLAP típica: total de vendas por produto e por cliente

OLTP x OLAP

■ Características OLAP:

- OLTP foca em transações, OLAP foca em consultas analíticas
- Normalização não é bom para consultas analíticas, reconstrução dos dados requer muitos joins
- BD OLAP suporta alta carga de consulta
- Técnicas de indexação OLTP não são eficientes para OLAP:
 - OLTP é orientado a acesso de poucos registros
- A necessidade de um novo modelo de BD para suportar OLAP levou ao DW
- DW: grande repositórios que consolidam dados de diferentes fontes (internas e externas à empresa), são atualizados offline, seguem o modelo multidimensional, projetado e otimizado para eficientemente suportar OLAP

OLTP x OLAP

	OLTP	OLAP
Target	operational needs	business analysis
Data	small, operational data	large, historical data
Model	normalized	denormalized/ multidimensional
Query language	SQL	not unified – but MDX is used by many
Queries	small	large
Updates	frequent and small	infrequent and batch
Transactional recovery	necessary	not necessary
Optimized for	update operations	query operations

Data Analytics

- Datawarehouse
 - SQL Server – SISS e SASS
- Data Lake
 - AWS Athena + S3
 - Databricks
- Data Lakewarehouse
 - Databricks
 - Snowflake
- Estudamos BIA (Business Intelligence & Analytics) na disciplina SAD (Sistemas de Apoio à Decisão)

Porque BI é importante?

Principais postos de trabalho com mais criação de vagas

1. Especialistas em IA e aprendizagem de máquina
2. Especialista em sustentabilidade
3. Analista em inteligência de negócios
4. Analista de segurança da informação
5. Engenharia de fintechs
6. Cientistas e analistas de dados
7. Engenharia de robótica
8. Especialista em big data
9. Operadores de equipamentos agrícolas
10. Especialistas em transformação digital

Fonte: <https://www.uol.com.br/tilt/noticias/redacao/2023/05/15/especialista-em-ia-eles-ganham-ate-r-35-mil-por-mes-criando-ias.htm>

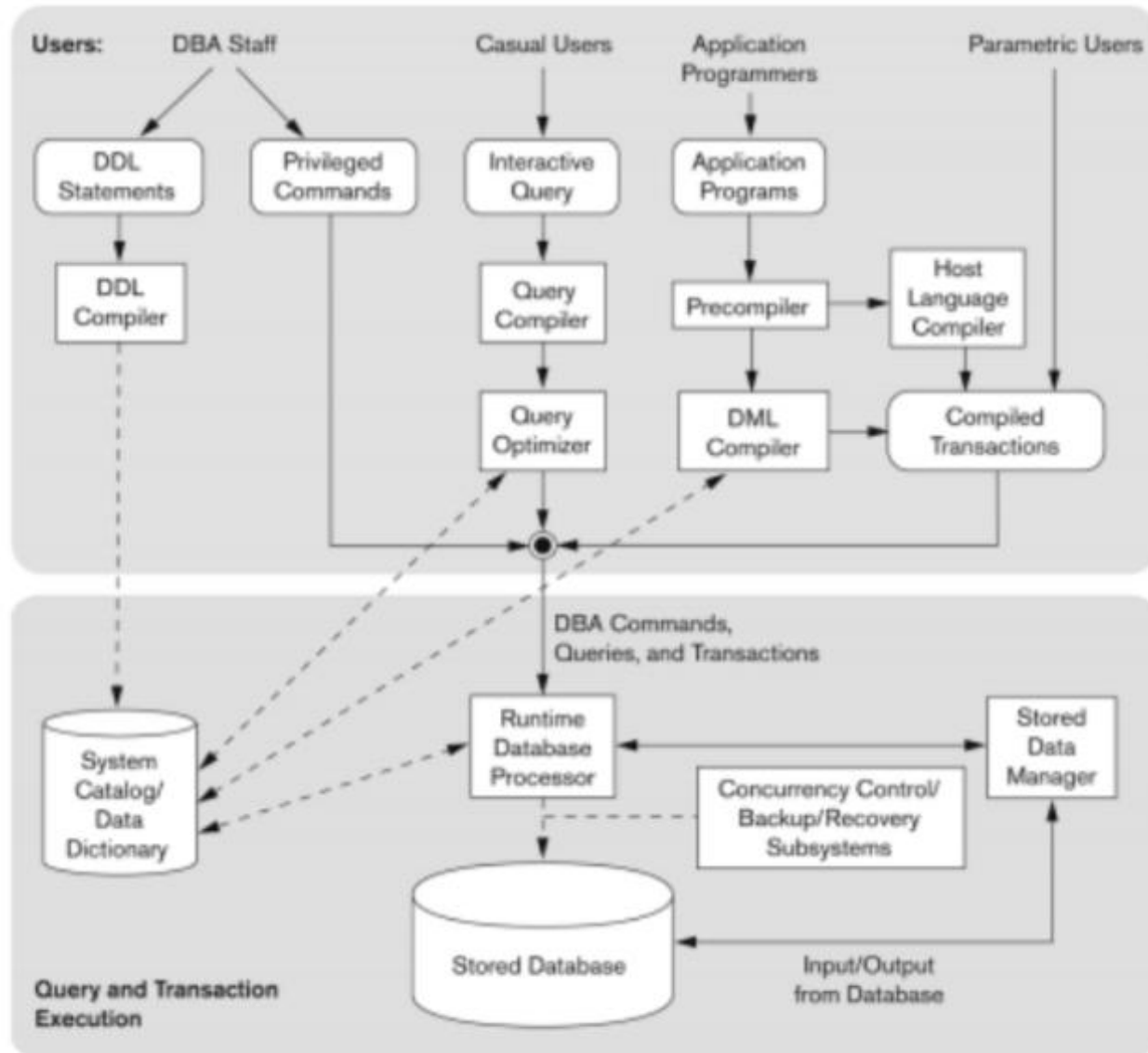
Em 15/05/2023

Serviços de um SGBD

- Concorrência
- Tolerância a falhas
- Controle de Integridade
- Segurança
- Otimização de consultas
- Auto-administração (Autonomic Computing)

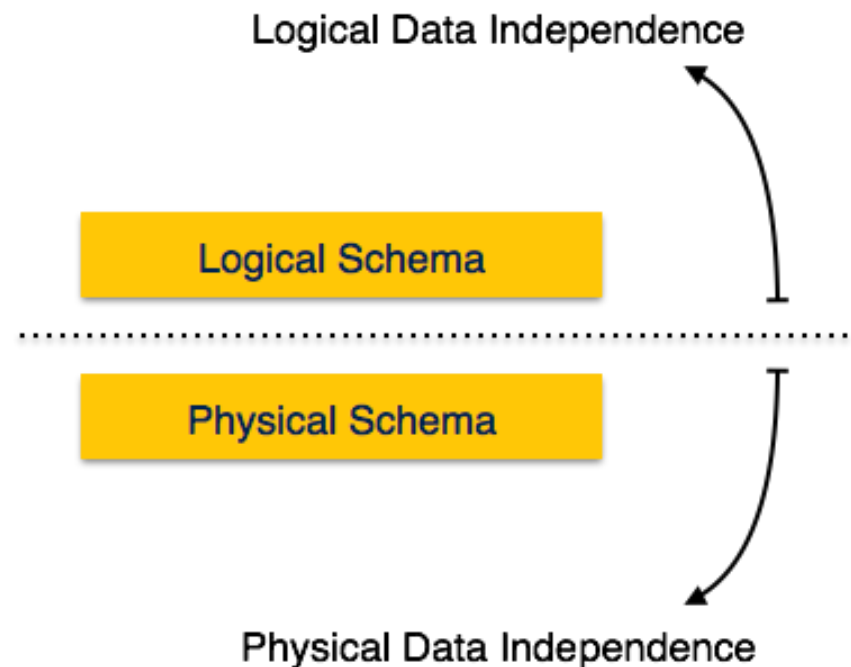


SGBD Relacional: Arquitetura



BD Relacional

- A grande sacada: INDEPENDÊNCIA de DADOS



Sucesso do SGBD Relacional

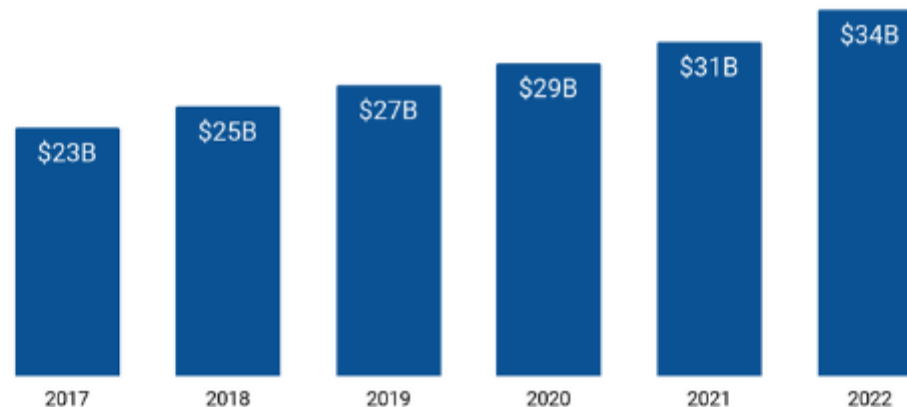
- Vários serviços oferecidos numa única ferramenta
- Simplicidade do modelo
- Formalismo baseado na teoria das relações
- Padronização: SQL
- Bom desempenho

Mercado SGBDR (\$B)

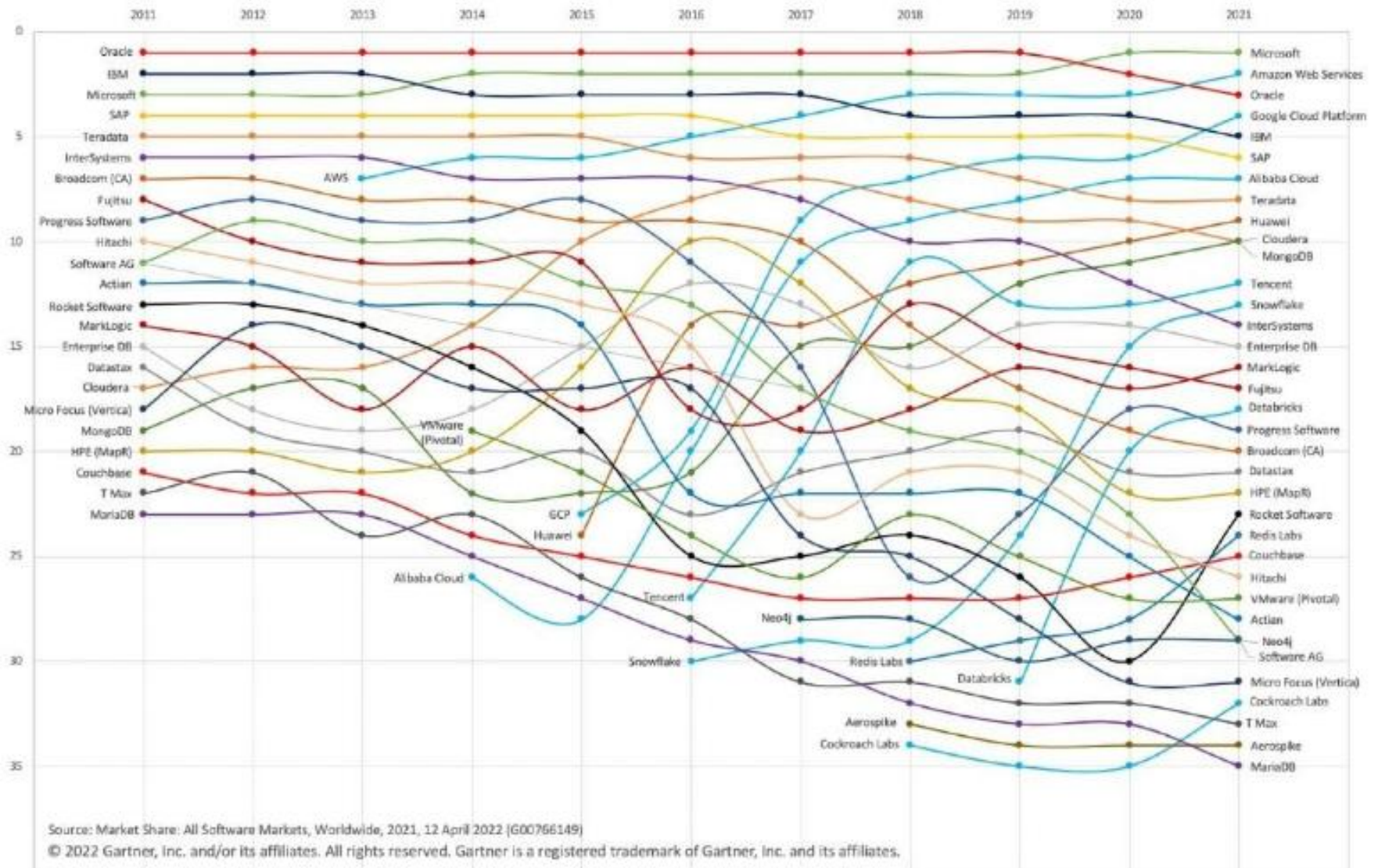
Worldwide Relational Database Market Systems Market Size, 2017-2022 (Excluding Data Warehousing)

Worldwide Relational Database Market Systems Market Size,
2017-2022 (Excluding Data Warehousing)

www.T4.ai



Gartner DBMS Market Share Ranks: 2011-2021



DB-Engines Ranking

The DB-Engines Ranking ranks database management systems according to their popularity. The ranking is updated monthly.

Read more about the [method](#) of calculating the scores.



424 systems in ranking, June 2025

Rank			DBMS	Database Model	Score		
Jun 2025	May 2025	Jun 2024			Jun 2025	May 2025	Jun 2024
1.	1.	1.	Oracle	Relational, Multi-model ⓘ	1230.38	+3.82	-13.70
2.	2.	2.	MySQL	Relational, Multi-model ⓘ	953.57	-11.41	-107.77
3.	3.	3.	Microsoft SQL Server	Relational, Multi-model ⓘ	776.75	+1.86	-44.81
4.	4.	4.	PostgreSQL +	Relational, Multi-model ⓘ	680.65	+6.34	+44.41
5.	5.	5.	MongoDB +	Document, Multi-model ⓘ	402.85	+0.33	-18.23
6.	6.	↑8.	Snowflake	Relational	174.49	+2.48	+44.13
7.	7.	↓6.	Redis	Key-value, Multi-model ⓘ	151.72	-0.47	-4.22
8.	8.	↑9.	IBM Db2	Relational, Multi-model ⓘ	125.13	-1.27	-0.77
9.	9.	↓7.	Elasticsearch	Multi-model ⓘ	121.28	-2.53	-11.55
10.	10.	10.	SQLite	Relational	117.03	-0.74	+5.63
11.	11.	↑12.	Apache Cassandra	Wide column, Multi-model ⓘ	108.27	+0.22	+9.44
12.	12.	↑15.	Databricks	Multi-model ⓘ	104.67	+2.02	+23.59
13.	↑14.	13.	MariaDB +	Relational, Multi-model ⓘ	94.54	+0.93	+3.50
14.	↓13.	↓11.	Microsoft Access	Relational	88.28	-7.63	-12.88
15.	15.	↑17.	Amazon DynamoDB	Multi-model ⓘ	83.34	+4.04	+8.90
16.	16.	↑18.	Apache Hive	Relational	76.68	+0.91	+16.92

DB-Engines Ranking - Trend Popularity

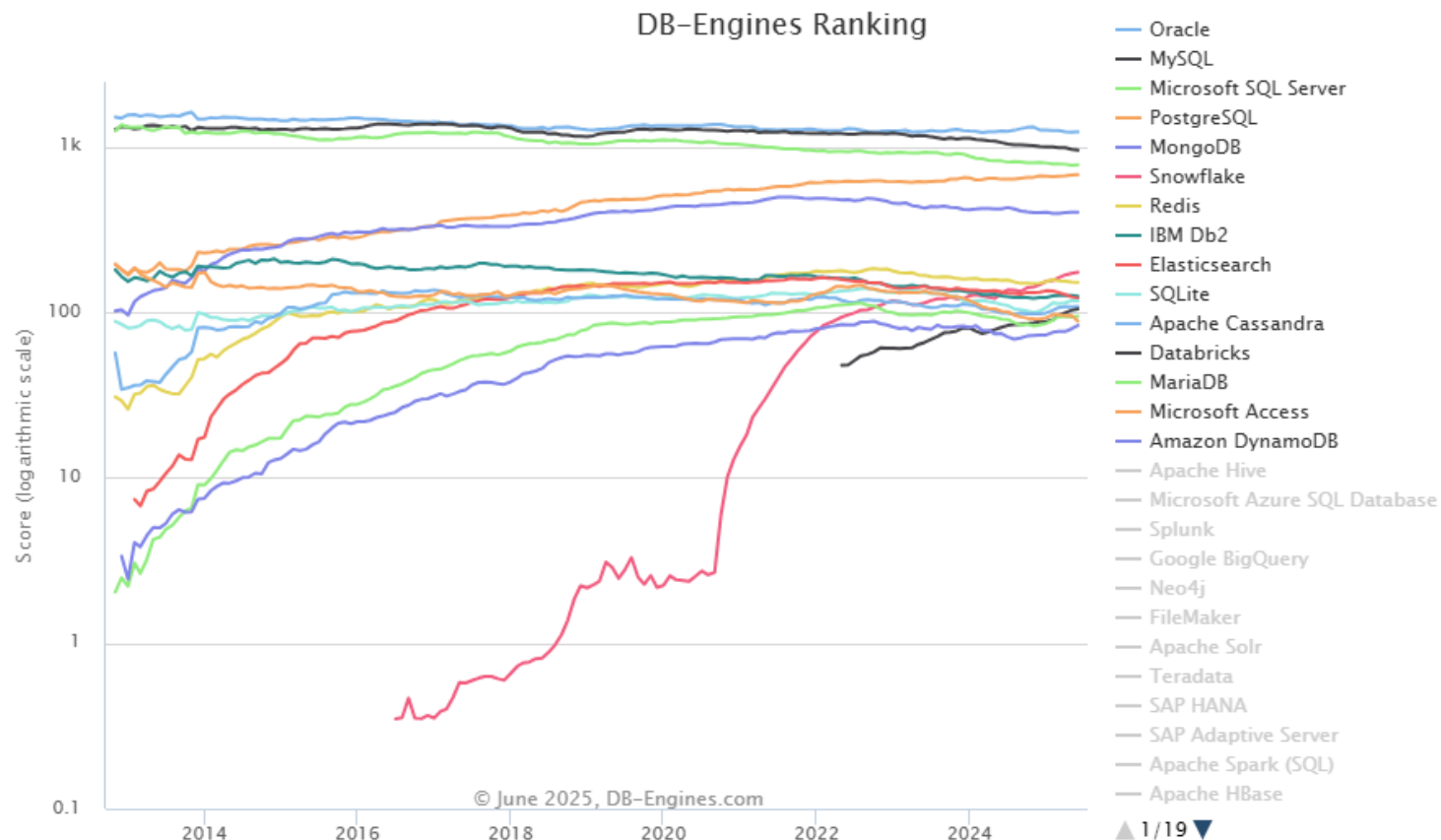
The DB-Engines Ranking ranks database management systems according to their popularity.

Read more about the [method](#) of calculating the scores.

Rank	Trend	System	Score	Change
1		Oracle	1560	+ 27
2	▲	MySQL	1342	+ 47
3	▼	SQL Server	1278	- 40
4		PostgreSQL	174	- 3
5		MS Access	161	- 8
6		DB2	155	- 4

ranking table

June 2025



Relatório Gartner OpDBMS (OLTP)

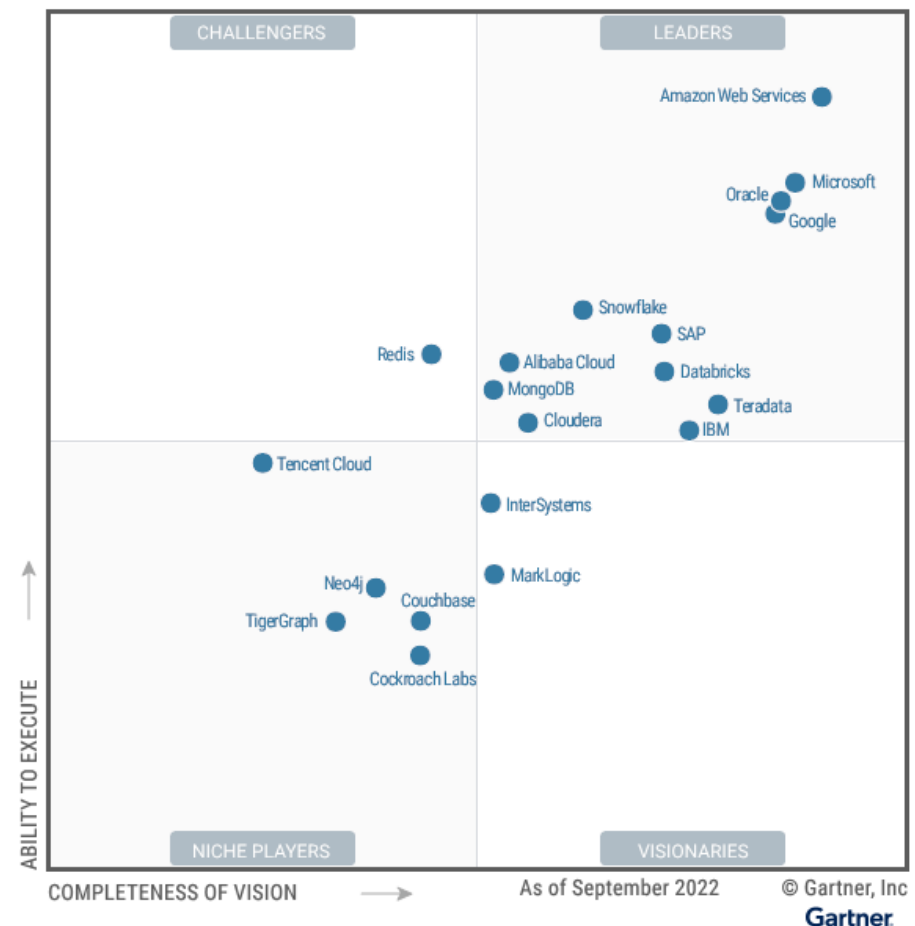
■ Market Definition/Description

- O mercado de operational database management system (OPDBMS) é definido pelos produtos de SGBD relational e nonrelational voltados para transações tradicionais usadas para dar suporte aos processos de negócio.
- Inclui ERP e CRM como também custom-made transactional systems

Gartner Quadrante Mágico: Cloud DBMS - 2022

Magic Quadrant

Figure 1: Magic Quadrant for Cloud Database Management Systems



Source: Gartner (December 2022)

Quadrante Mágico CloudDBMS: 2024



Banco de Dados 1

- Exemplo Motivacional 1:
- Imagine que você é contratado para construir uma aplicação do ZERO em Java para gerenciar o banco NuDinheiro, SEM usar um banco de dados.
- Meta: gerenciar clientes, contas, contas conjuntas, transferências, taxas de juro.
- Quais são os aspectos que você precisa se preocupar para desenvolver esta aplicação?

Banco de Dados 1

- Exemplo 1: Resposta
 - Lidar com um grande volume de dados

Banco de Dados 1

- Exemplo 1: Resposta
 - Lidar com um grande volume de dados
 - Ser rápido

Banco de Dados 1

- Exemplo 1: Resposta
 - Lidar com um grande volume de dados
 - Ser rápido
 - Não perder dados

Banco de Dados 1

- Exemplo 1: Resposta
 - Lidar com um grande volume de dados
 - Ser rápido
 - Não perder dados
 - Permitir múltiplos usuários ao mesmo tempo

Banco de Dados 1

- Exemplo 1: Resposta
 - Lidar com um grande volume de dados
 - Ser rápido
 - Não perder dados
 - Permitir múltiplos usuários ao mesmo tempo
 - Ter consistência nos dados

Banco de Dados 1

- Exemplo 1: Resposta
 - Lidar com um grande volume de dados
 - Ser rápido
 - Não perder dados
 - Permitir múltiplos usuários ao mesmo tempo
 - Ter consistência nos dados
 - Fácil de usar (UI/UX)

Banco de Dados 1

- Exemplo 1: Resposta
 - Lidar com um grande volume de dados
 - Ser rápido
 - Não perder dados
 - Permitir múltiplos usuários ao mesmo tempo
 - Ter consistência nos dados
 - Fácil de usar
 - Ser seguro

Banco de Dados 1

- Exemplo 1: Resposta
 - Lidar com um grande volume de dados
 - Ser rápido
 - Não perder dados
 - Permitir múltiplos usuários ao mesmo tempo
 - Ter consistência nos dados
 - Fácil de usar
 - Ser seguro
 - Acesso via plataforma móvel

Banco de Dados 1

- Exemplo 1: Resposta
 - Lidar com um grande volume de dados
 - Ser rápido
 - Não perder dados
 - Permitir múltiplos usuários ao mesmo tempo
 - Ter consistência nos dados
 - Fácil de usar
 - Ser seguro
 - Acesso via plataforma móvel
 - Gestão de Logística

Banco de Dados 1

- Exemplo1: Resposta
 - Lidar com um grande volume de dados:
 - **indexação**
 - Ser rápido
 - **otimização**
 - Não perder dados
 - **persistência**
 - Permitir múltiplos usuários ao mesmo tempo
 - **concorrência**
 - Ter consistência nos dados
 - **integridade**
 - Fácil de usar
 - **SQL: o que você quer, ao invés de como**
 - Ser seguro
 - **segurança**
 - Acesso via plataforma móvel (web banking)

Sistema de Informação

- é a expressão utilizada para descrever um sistema informacional computadorizado
 - Abrange:
 - pessoas
 - máquinas e/ou
 - métodos organizados
 - Visando:
 - Colectionar
 - Armazenar
 - Processar
 - Transmitir e
 - Disseminar dados de uma organização



Case 2: Mercadinho VendeTudo

- Imagine que você vai abrir um mercadinho
 - 5 funcionários
 - quantos produtos distintos são comercializados?
 - 20 fornecedores
 - Infra:
 - 2 caixas (checkout)
 - 30 gôndolas
 - 1 veículo pick-up
 - 2 motos
 - Quais Sistemas de Informação você precisará?



Sistema de Informação

- Alguns sistemas de informação do mercadinho VendeTudo:
 - Gestão de Vendas

Sistema de Informação

- Alguns sistemas de informação do mercadinho VendeTudo:
 - Gestão de Vendas
 - Controle de estoque

Sistema de Informação

- Alguns sistemas de informação do mercadinho VendeTudo:
 - Gestão de Vendas
 - Controle de estoque
 - Financeiro (Contas a pagar e a receber)

Sistema de Informação

- Alguns sistemas de informação do mercadinho VendeTudo:
 - Gestão de Vendas
 - Controle de estoque
 - Financeiro (Contas a pagar e a receber)
 - Contábil

Sistema de Informação

- Alguns sistemas de informação do mercadinho VendeTudo:
 - Gestão de Vendas
 - Controle de estoque
 - Financeiro (Contas a pagar e a receber)
 - Contábil
 - Fiscal

Sistema de Informação

- Alguns sistemas de informação do mercadinho VendeTudo:
 - Gestão de Vendas
 - Controle de estoque
 - Financeiro (Contas a pagar e a receber)
 - Contábil
 - Fiscal
 - RH (folha de pagamento, férias, licença, etc)

Sistema de Informação

- Alguns sistemas de informação do mercadinho VendeTudo:
 - Gestão de Vendas
 - Controle de estoque
 - Financeiro (Contas a pagar e a receber)
 - Contábil
 - Fiscal
 - RH (folha de pagamento, férias, licença, etc)
 - Cadastro de Clientes

Sistema de Informação

- Alguns sistemas de informação do mercadinho VendeTudo:
 - Gestão de Vendas
 - Controle de estoque
 - Financeiro (Contas a pagar e a receber)
 - Contábil
 - Fiscal
 - RH (folha de pagamento, férias, licença, etc)
 - Cadastro de clientes
 - Cadastro de Fornecedores

Sistema de Informação

- Alguns sistemas de informação do mercadinho VendeTudo:
 - Gestão de Vendas
 - Controle de estoque
 - Financeiro (Contas a pagar e a receber)
 - Contábil
 - Fiscal
 - RH (folha de pagamento, férias, licença, etc)
 - Cadastro de clientes
 - Cadastro de fornecedores
 - Controle Patrimonial (Despesas com máquinas)

Sistema de Informação

- Alguns sistemas de informação do mercadinho VendeTudo:
 - Gestão de Vendas
 - Controle de estoque
 - Financeiro (Contas a pagar e a receber)
 - Contábil
 - Fiscal
 - RH (folha de pagamento, férias, licença, etc)
 - Cadastro de clientes
 - Cadastro de fornecedores
 - Controle Patrimonial (Despesas com máquinas)
 - Sistema para Marketing (Redes sociais/Tv/impressos)

Sistema de Informação

- Alguns sistemas de informação do mercadinho VendeTudo:
 - Gestão de Vendas
 - Controle de estoque
 - Financeiro (Contas a pagar e a receber)
 - Contábil
 - Fiscal
 - RH (folha de pagamento, férias, licença, etc)
 - Cadastro de clientes
 - Cadastro de fornecedores
 - Controle Patrimonial (Despesas com máquinas)
 - Sistema para Marketing (Redes sociais/Tv/impressos)
 - Administrativo (Automação de Escritório)
 - Sistema de Logística

Tipos de Sistema de Informação (alguns)

- ERP
- SCM
- CRM
- TMS

Tipos de Sistema de Informação (alguns)

- ERP – Enterprise Resource Processing
- SCM – Supply Chain Management
- CRM – Customer Relationship Management
- TMS - Transportation Management System

ERP

- sistema de informação que interliga vários dados e processos de uma organização em um único sistema
- Módulos básicos:
 - Emissão de notas fiscais eletrônicas de produtos e serviços
 - Controle financeiro
 - Produção
 - Compras
 - Gerenciamento de estoque
 - PDV online

ERP



SCM

- Gerenciamento de Cadeia de Suprimento
- sistema no qual organizações e empresas entregam seus produtos e serviços aos seus consumidores, numa rede de organizações interligadas e lida com problemas de planejamento e execução envolvidos no gerenciamento de uma cadeia de suprimentos.
- Ex. Fábrica de automóveis
- Ex. Fábrica de Queijo



CRM

- Sistema de relacionament ocom os clientes



TMS

- Sistema para Gestão de Transporte e Logística
- Aprimora a qualidade do processo de distribuição de bens e produtos



Case 3: VendeTudo e-commerce

- E se seu mercadinho VendeTudo se expandisse e com a pandemia da COVID-19 você precisasse fazer vendas online?
- O que mudaria na sua solução?



Ex. Arquitetura de um e-Commerce

– Sistema Integrado Supply Chain



PLATAFORMA

- Cadastro de Produtos
- Desenvolvimento do Layout do site
- Emails Marketing
- Contato com Cliente
- Armazenamento de dados de clientes
- Cadastro de produtos
- Integração com os Marketplaces
- Gestão de pedidos



ERP

- Cadastro de Produtos
- Emissão de documentação fiscal (NF)
- Controle financeiro
- Controle de estoque (crossdocking)
- Notas de entrada
- Cadastro de Fornecedores
- Pedidos de compra



ADM FINANCEIRA

- Análise de Fraude
- Transações Financeiras
- Integração com redes de cartões (Master Card, Visa etc)



INTEGRADORA

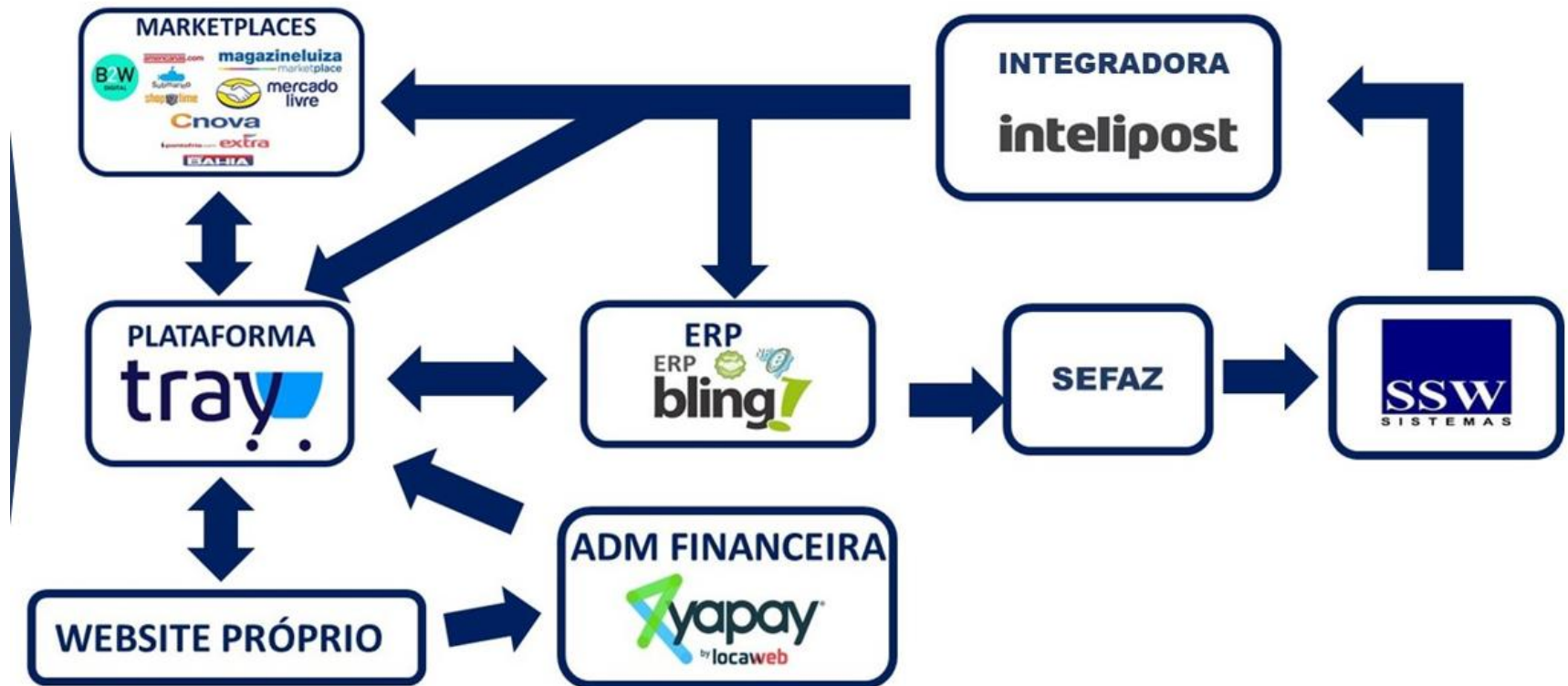
- Cadastro de tabelas de frete
- Leilão de Frete, escolha da melhor opção de transportadora
- Atualização de status automaticamente
- Baixa de entregas
- Controle de faturas



TMS

- Gestão de Fretes
- Emissão de Ctes, MDFes, Subcontratos
- Gestão Financeira
- Fiscal
- Análise de desempenho
- Baixa e acompanhamento de entregas
- Transação com SEFAZ

Ex. Arquitetura de um e-Commerce APC – Automação do Processo de Compras



Por que Banco de Dados?

- Qual a relação de todos estes sistemas do Mercadinho VendeTudo com Banco de Dados?

Video Motivacional

- Porque Banco de Dados?

Por que Banco de Dados?

- Qual a relação de todos estes sistemas do Mercadinho VendeTudo com Banco de Dados?
- Todos estes sistemas que discutimos utilizam um SGBD (sistema de Gestão de Banco de Dados) no backend

Por que Banco de Dados?

- Todas estas soluções precisam:
 - Transações ACID – Atomicidade, Consistência, Isolação e Durabilidade
 - Indexação
 - Facilidade de consulta – SQL
 - Controle de concorrência
 - Tolerância à falhas
 - Armazenamento com **Independência de Dados**
 - Controle de Integridade
 - Controle de Segurança

BD - Novas Arquiteturas

■ NOSQL

- Banco de dados sem esquema
- Vários tipos : orientado a coluna, key-value, grafo, orientado a documento
- API proprietária

■ NEWSQL

- Nova geração de SGBD relacional que pode escalar como um sistema NOSQL, mas sem esquecer SQL e transações

BD NOSQL

■ NOSQL

- Banco de dados sem esquema
- Vários tipos : orientado a coluna, key-value, grafo, orientado a documento
- API proprietária e programável

NOSQL

Key/Value



redis



riak



FOUNDATIONDB



membase



amazon
DynamoDB

Column-Family



cassandra

APACHE
HBASE



HYPERTABLE

Documents

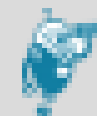


mongoDB

Couchbase



CouchDB
relax



RethinkDB

NOSQL

■ Desvantagens

- Desenvolvedores escrevem código para manipular dados eventualmente consistentes, falta de transações
- Nem todas as aplicações podem renunciar à semântica transacional
- Problemas com maturidade ainda, ver problema de ataques ao MongoDB reportados em 2017

NEWSQL



NewSQL

- Controle de concorrência distribuído
- Armazenamento em memória principal
- Arquitetura híbrida
 - Suporte à OLTP e OLAP num mesmo SGBD

Cloud Database

- SGBD que roda numa plataforma de cloud (Amazon AWS, Google Cloud, Salesforce, Azure, Oracle Cloud)
- Modelo de deployment
 - Virtualização e Containers
 - Database as a service (DBaaS)

Cloud Database: porque agora?

- Experiência com grandes data centers
 - Economia de escala sem precedente
 - Transferência de risco
- Fatores tecnológicos:
 - Largura de banda
 - Maturidade em virtualização

Cloud Database: porque agora?

- Experiência com grandes data centers
 - Economia de escala sem precedente
 - Transferência de risco
- Fatores tecnológicos:
 - Largura de banda
 - Maturidade em virtualização
- Fatores de negócio:
 - Mínimo investimento de capital
 - modelo de pagamento: Pay-as-you-go

AWS Cloud Database Systems

- Amazon Relational Database Service (RDS)
- Amazon Aurora (relacional: MySQL - Postgresql)
- Amazon DynamoDB (chave-valor)
- Amazon Neptune (grafo)
- Amazon DocumentDB
- Amazon KeySpaces (for Apache Cassandra)
- Amazon MemoryDB (for Redis)

Google Cloud Database Systems

- Google Cloud SQL
- Cloud Spanner (relacional distribuído)
- Cloud Bigtable (colunar e chave-valor)
- BigQuery (datawarehouse)
- Dataproc (datalake com hadoop/spark/flink)
- Cloud Firestore (NOSQL document)
- Firebase Realtime Database (realtime/json)
- BigLake (data lake + DW)

Microsoft Cloud Database Systems

- Azure SQL
- Azure Database for PostgreSQL
- Azure Database for MySQL
- Azure Database for MariaDB
- Azure Cache for Redis
- Azure Managed Instance for Apache Cassandra
- Azure Cosmos DB.

Aplicações Emergentes

- SGBDs são o lugar natural para dados
- Porém, as aplicações estão se tornando cada vez mais complexas
- SGBD precisam mudar e expandir para dar suporte aos novos requisitos e desafios
- Novas direções de pesquisa
 - Novos modelos e formatos de dados
 - Novos métodos de acesso
 - Novas arquiteturas
 - Novas técnicas de otimização e processamento de consultas

Aplicações Emergentes

- Bancos de Dados Vetoriais
 - SGBDs projetados para armazenar vetores
 - Apropriados para Machine Learning e NLP (LLM)
 - Exemplos:
 - Milvus
 - Chroma
 - Margo
 - Vespa

Ex. de Aplicações Emergentes

- Sistema de Gestão de Fluxo de Dados (Data Stream Management System)
 - Dados ficam continuamente chegando (sem persistência)
 - Balanceamento de carga
 - Exemplos:
 - Network monitoring and traffic engineering,
 - Healthcare monitoring,
 - Financial applications,
 - Sensor networks,
 - Manufacturing processes,
 - Web logs and click streams

Ex. de Aplicações Emergentes

- Objetos Móveis e Aplicações Espaço-Temporal
 - Fluxo contínuo de objetos móveis
 - Dados bi-dimensionais (espaço e tempo)
 - Tipos de consultas espaciais: point queries, range queries, KNN queries, topológicas
 - Tipos de dados espaciais: point, line, polygon

Ex. de Aplicações Emergentes

■ Gestão de Dados Científicos

- Ex.: biologia, química, meteorologia, etc
- Tipos de Dados complexos (arrays, imagens, sequências, estruturas)
- Metadados, anotações e comentários sobre os dados
- Processamento complexo e workflow
- Informação sobre Proveniência e Linhagem

Ex. de Aplicações Emergentes

- Analytics de dados em larga escala e processamento distribuído
 - Processamento de dados em escala massiva (terabytes e petabytes)
 - Processamento altamente distribuído e paralelo
 - Nova infraestrutura e paradigmas de computação
 - SGBD distribuído e uso dos frameworks Spark Hadoop/MapReduce com Pig e Hive
 - Snowflake
 - Google Big Query
 - Amazon Redshift
 - Microsoft Azure Synapse Analytics

Mais pesquisas em BD

- ❑ Banco de Dados Multimídia
- ❑ Bancos de Dados Textuais e Recuperação da Informação
- ❑ BD e Web
 - ❑ Hidden Web
 - ❑ Semântica, Ontologias, OWL, RDF, SPARQL, YAGO

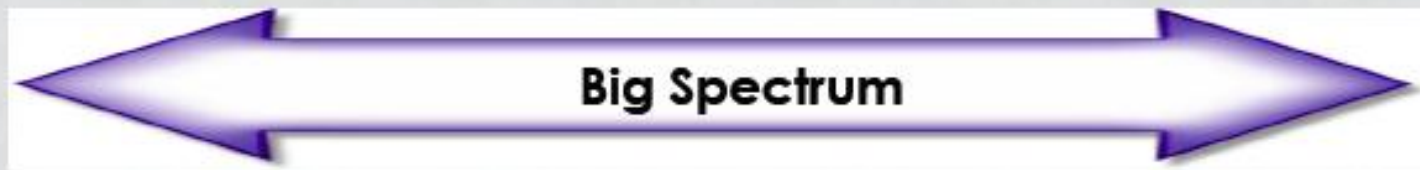
Profissionais de Dados

- **1. Data engineer:** projetam e constroem sistemas para coletar e analisar dados. Usam SQL para consultar e gerenciar dados
- **2. Database administrator / DBA:** agem como suporte técnico para banco de dados, assegurando desempenho, segurança, backup, migração e integração de dados
- **3. Data architect:** analisam a infraestrutura de uma organização para planejar ou implantar BD e SGBD que melhorem a eficiência do fluxo de dados.

Profissionais de Dados

- **3. Data architect:** analisam a infraestrutura de uma organização para planejar ou implantar BD e SGBD que melhorem a eficiência do fluxo de dados.
- **4. Analista de Negócios / BI:** usam dados de databases para limpá-los e interpretá-los visando resolver um business problem, com storytelling, KPI, dashboards e reports.
- **5. Cientista de Dados / Data scientist:** usam dados para encontrar padrões e tendências, usando algoritmos de predição. Uso massivo de técnicas de machine learning.
- **6. Analista de Segurança de Dados (DPO- Data Protection Officer):** trabalha na segurança da informação de olho em compliance e LGPD.

Conclusão: Várias aplicações



Banking



Physics



Streaming



Retail Sys



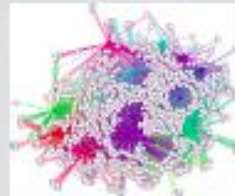
Biology



Social Media



Airlines



Graph Data



Spatio-Temporal



Big Data

Traditional

**Scientific and
More advanced**

Big Data

Conclusão: Várias ferramentas



Conclusões

- Existe um mar de oportunidades para quem investir na área de Banco de Dados nos próximos anos



Conclusões

- Existe um mar de oportunidades para quem investir na área de Banco de Dados nos próximos anos
- Mercado em expansão para perfil de profissional:
 - Banco de Dados + BI + Aprendizagem de Máquina

