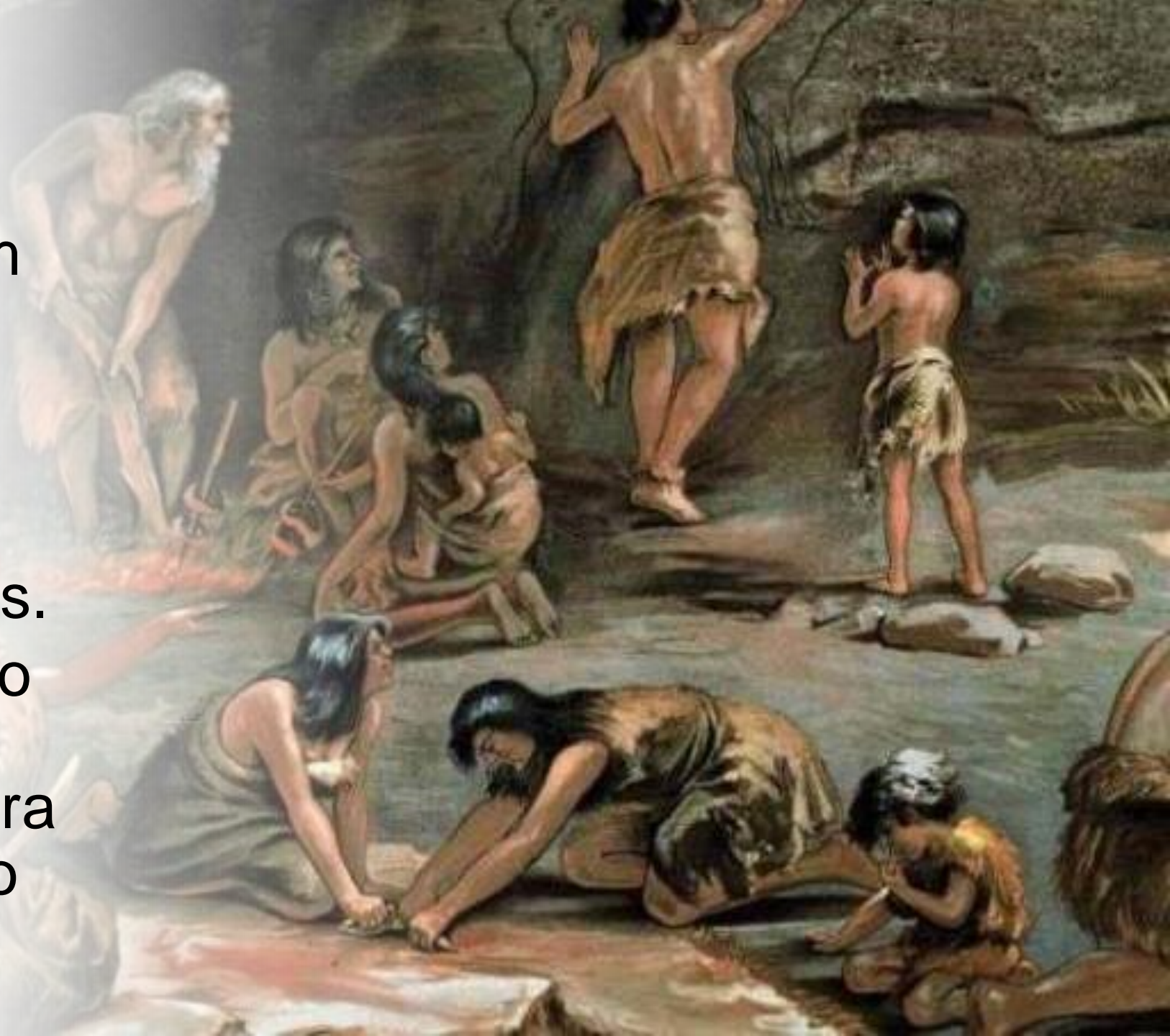




Banco de Dados Histórico até as Tecnologias Atuais

Prof. Dr. Vladimir Costa Alencar
LANA - UEPB
valencar.com

- Há 2,5 Milhões de anos, nossos ancestrais já talhavam a pedra para obter alimentos.
- Dessa forma, ele conseguia caçar, quebrar ossos e nozes.
- Também dominavam o fogo e sabiam usar plantas medicinais para combater parasitas do intestino.



- Há 200 mil anos, os homens se tornaram anatomicamente parecidos conosco: **Homo sapiens**.

- Ficaram mais inteligentes e solidários.

- Com o crescimento do cérebro, as fêmeas precisaram comer mais proteínas para alimentar o feto.



A necessidade de registrar informações
(44.000 a.C)



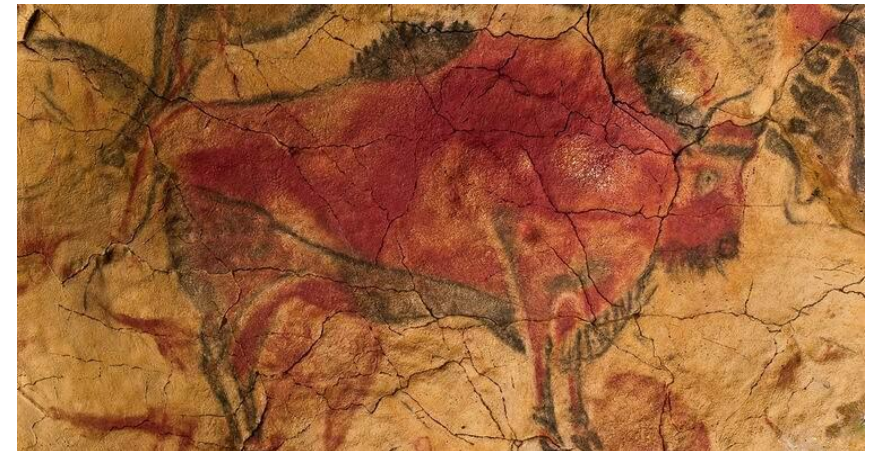
O uso de Inscrições rupestres, barro, etc



Bisão na Caverna de Chauvet, na França, 41.000 a.C.



A pintura rupestre de um javali, Indonésia. 44.000 a.C.



Bisão na Caverna de Altamira, Espanha. 32.000 a.C.



Caverna da Serra da Capivara, Piauí, 29.000 anos

Registros Rupestres Brasil (31.000 a.C.)



Itacoatiras de Ingá, Paraíba, 6.000 anos

- A escrita surgiu por volta de 3.400 a.C. para resolver um problema de natureza bem prática dos sumérios: contar mercadorias

- Uruk foi a primeira cidade construída na face da Terra, nascida por volta de 3.500 aC na Mesopotâmia, ao sul do que hoje seria o Iraque.

- Nela surgiram os primeiros ordenamentos do território, os primeiros escritos e a primeira arquitetura monumental, entre outros.

- No apogeu estima-se que ela teve uma população entre 50.000 e 80.000 habitantes e uma área murada de até seis quilômetros



Os primeiros achados de Escrita Sistematizada, datam de 3.200 a.C.



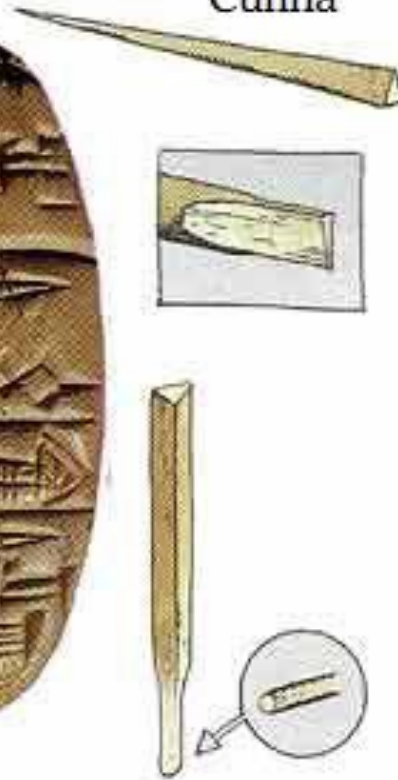
Escrita cuneiforme na
Cidade de Uruk,
Mesopotâmia
(Sumérios, Iraque e o Kuwait.)



Tábua de argila



Cunha

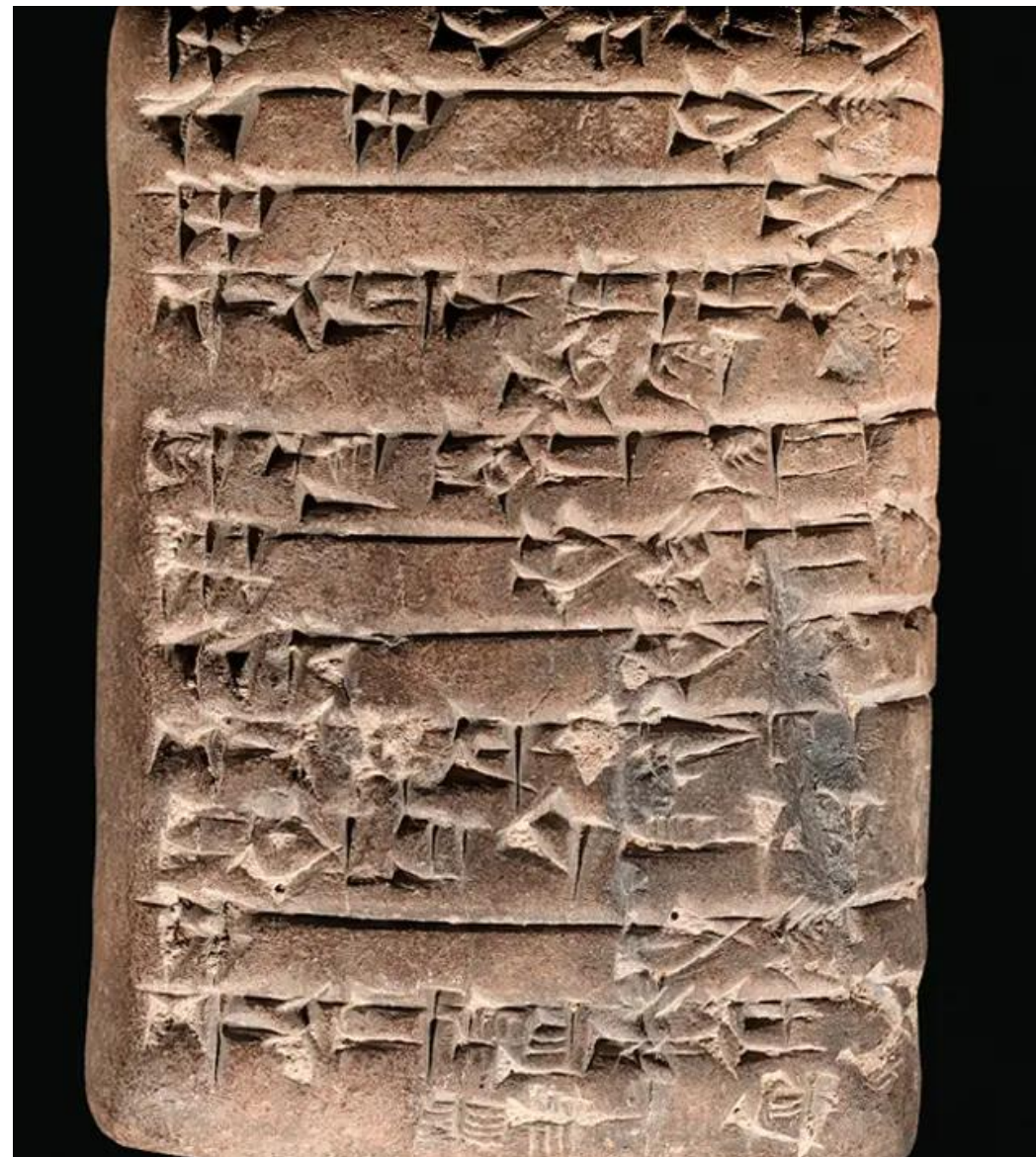


Uma escrita sistematizada
3.200 a.C.



Escrita cuneiforme na
Mesopotâmia
(Sumérios - Iraque e o
Kuwait.)

Tábua com Registro de transação de gado, na Cidade de Uruk, na Mesopotâmia, 3.200 a.C.





Uma escrita sistematizada
3.200 a.C.



Escrita cuneiforme na
Mesopotâmia (Sumérios -
Iraque e o Kuwait.)



Uso de Pinturas



Esculturas



Pinturas Egípcias, 2.686 a.C.

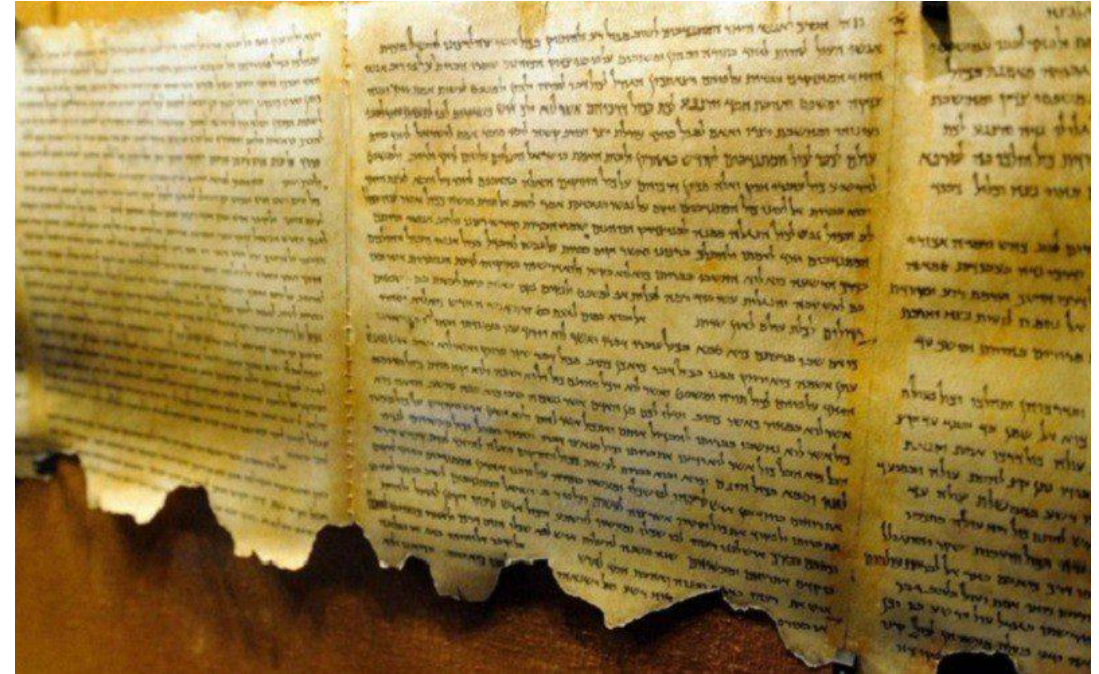


Código de Hamurabi,
Babilônia (1.750 a.C.)



Papiro

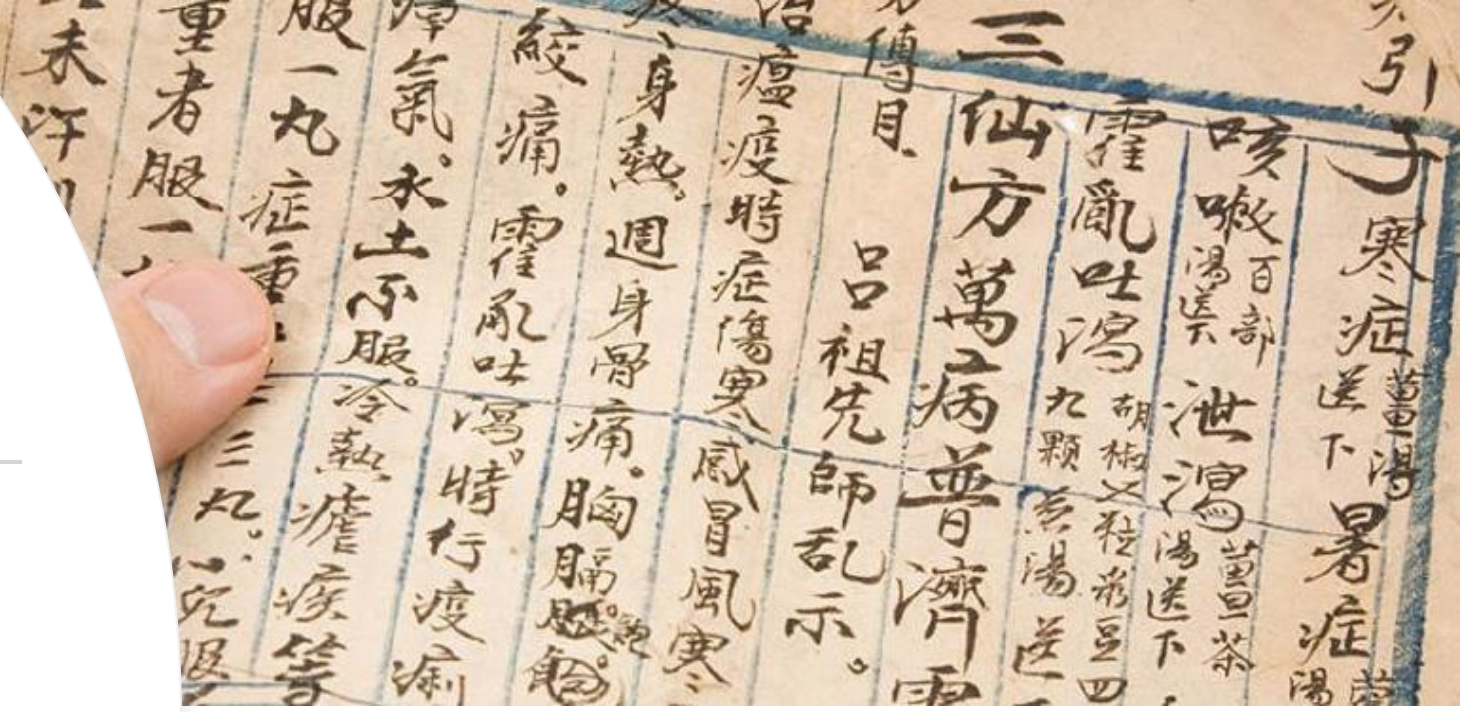
Produzidos por meio de uma planta também chamada de papiro Egito, 3.000 a.C.



Pergaminho

Surgiu no 200 a.C., em Pérgamo, uma cidade na zona de influência da cultura grega.

Era produzido com a pele de animais como ovelhas, carneiros, cabras

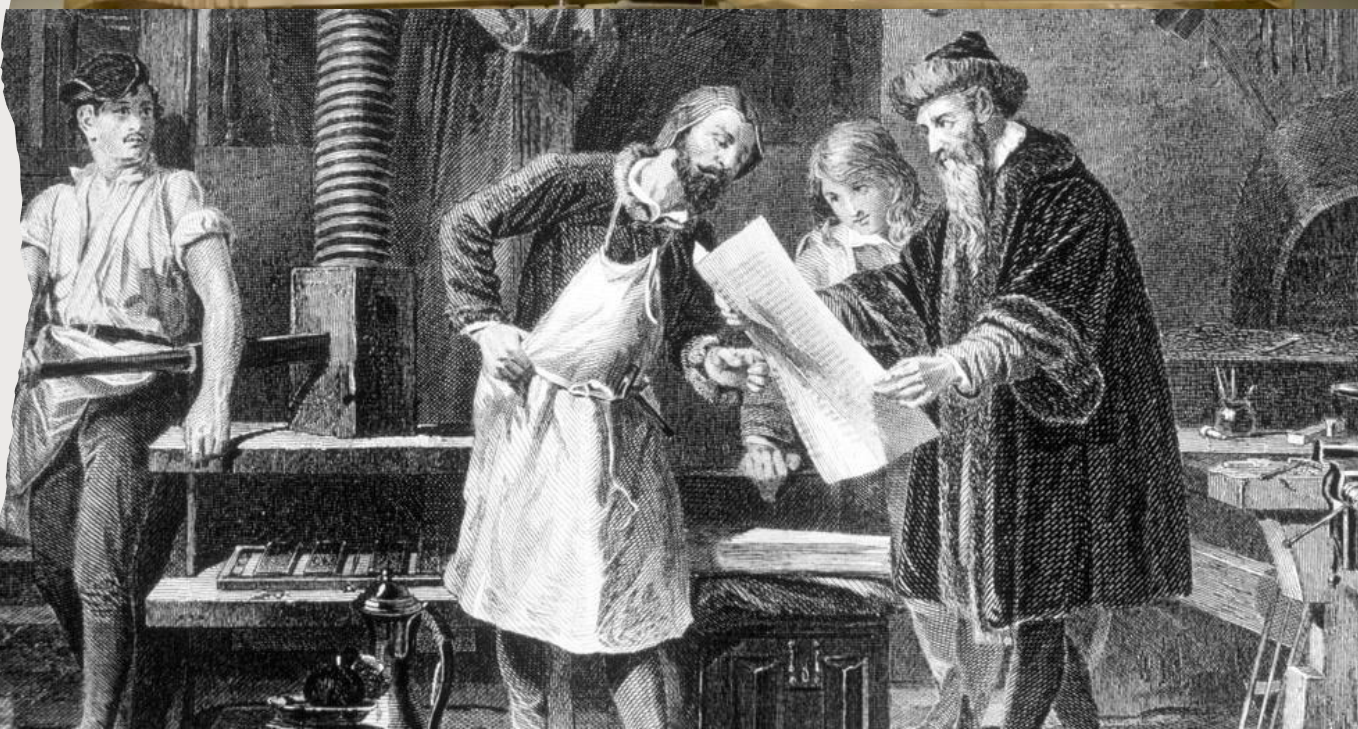


- O papel foi inventado na China, 105 d.C., por **T'sai Lun**.

- Ele fez uma mistura umedecida de casca de amoreira, restos de roupas e outros produtos que contivesse fonte de fibras vegetais



- Criada na China, a técnica da prensa tipográfica revolucionou a sociedade local antes de ser aprimorada pelo alemão
- A Bíblia impressa por Johannes Gutenberg em 1.455 d.C. marcou o início de seu uso no Ocidente, mas o processo de impressão já existia no Oriente.
- O primeiro livro impresso com a técnica de Gutenberg é chamado de **Bíblia em Latim**. O livro tinha **641 páginas** que foram forjadas em letras em chumbo e arranjadas manualmente.
- A Bíblia de Gutenberg foi impressa com estilo de escrita gótica. Foram feitos cerca de **300 exemplares do livro**, todos com detalhes minuciosos desenvolvidos pelo alemão Johannes Gutenberg.

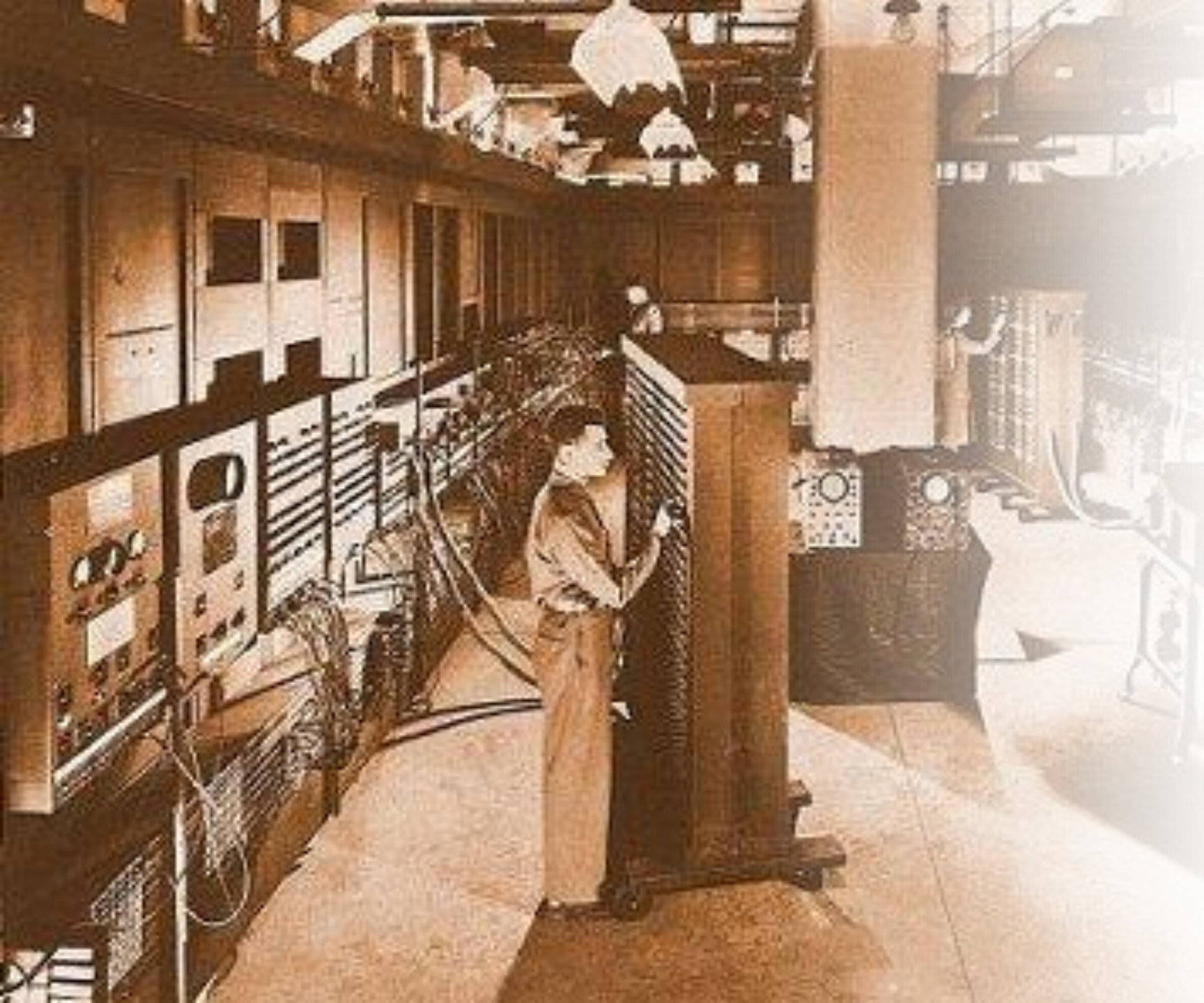


- Em 1867, o americano Christopher Latham Sholes, com auxílio de Carlos Glidden e Samuel Soulé, conseguiu desenvolver a primeira máquina de escrever que realmente atendia às necessidades dos escritórios.

- Para que as hastes de metal do teclado não se prendessem umas nas outras, Sholes reconfigurou o teclado, separando as letras mais usadas no alfabeto inglês.

- Nascia, assim, o teclado QWERTY (as seis primeiras teclas da primeira linha do teclado), padrão até hoje nos computadores.





ENIAC (Electronic Numerical Integrator and Computer)

- **Desenvolvedores:** John Presper Eckert e John Mauchly na Universidade da Pensilvânia.
- **Data de conclusão:** 1945.
- **Descrição:** O ENIAC é amplamente reconhecido como o **primeiro computador digital eletrônico de uso geral totalmente funcional**.
- Ele foi projetado inicialmente para **calcular tabelas de tiro** para a artilharia do Exército dos EUA.
- O ENIAC usava cerca de **18.000 válvulas** (tubos a vácuo) e pesava cerca de **30 toneladas**.
- Ele era programável, mas exigia que operadores reconectassem manualmente cabos e ajustassem interruptores para programar novos cálculos.
- Ele podia realizar **5.000 adições** ou subtrações **por segundo**, uma velocidade impressionante para a época.

Histórico – (Década de 50 e início de 60)



IBM com 8 MB de memória RAM
capacidade da Fita Magnética **2 MB**
velocidade de 7.500 caracteres por segundo

Os programas de banco de dados foram os programas pioneiros na computação (1950-60)

Surgiram para atender para atender equipamentos de grande porte (mainframes)

Histórico

Computador IBM 1130 (UFCG), foi o primeiro a ser instalado nas regiões Norte/Nordeste do Brasil

Terceiro mainframe a ser instalado no Brasil, 1968



Histórico

O setor prestava serviços de processamento de dados para diversas empresas do Nordeste a exemplo da:

CAGEPA (Companhia de Água e Esgotos da Paraíba), CAERNE (RN) , BNB (Banco do Nordeste do Brasil), SUDENE (PE), Indústria Wallig Fogões.

O Escocês Ian foi quem implantou os primeiros softwares e treinamentos na UFCG.



Histórico (Década de 50 e início de 60)



Fitas magnéticas foram desenvolvidas para o armazenamento de dados

Tarefas de processamento de dados, como folha de pagamento, eram automatizadas com dados armazenados em fitas

As fitas podiam ser lidas apenas sequencialmente

O processamento consistia em ler os dados e armazenar em uma nova fita (Fita Mestre)

Histórico – Arquitetura nos primórdios

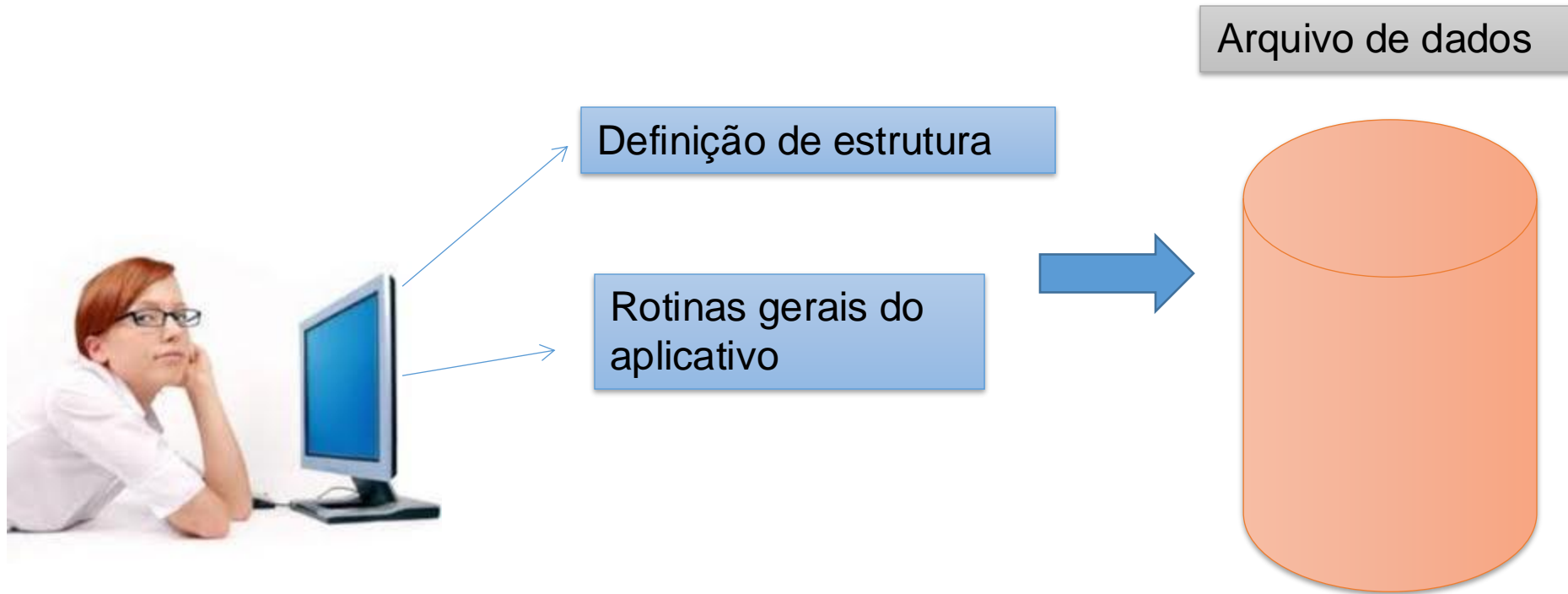
- Não era possível executar mais de um processo ao mesmo tempo num arquivo de dados
- A estrutura do arquivo estava inserida no próprio código do aplicativo
- --> Ruim para manutenção

O programas controlavam o acesso ao banco de dados

- Inconsistência, redundância, dificuldade de acesso, isolamento dos dados
- Problemas com Segurança

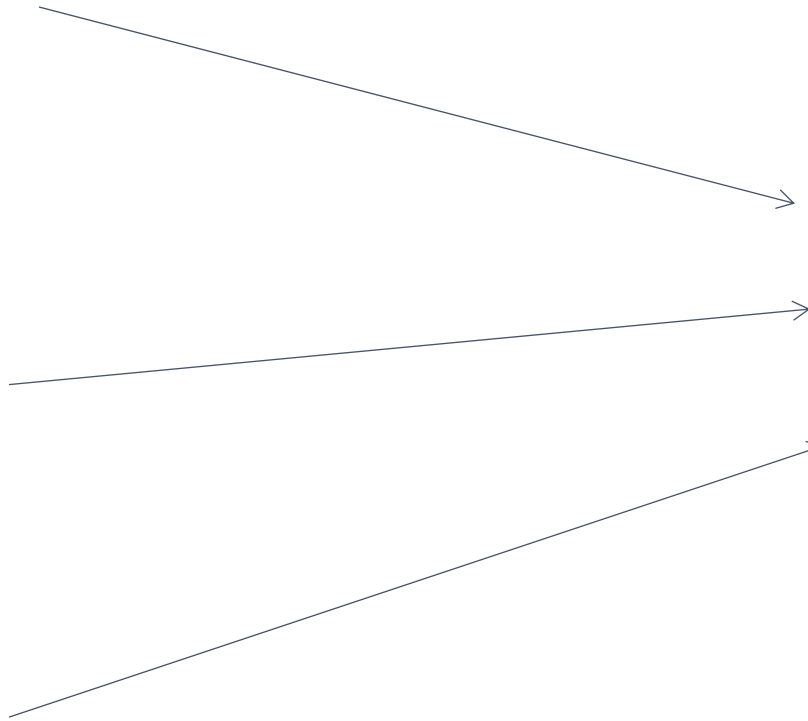
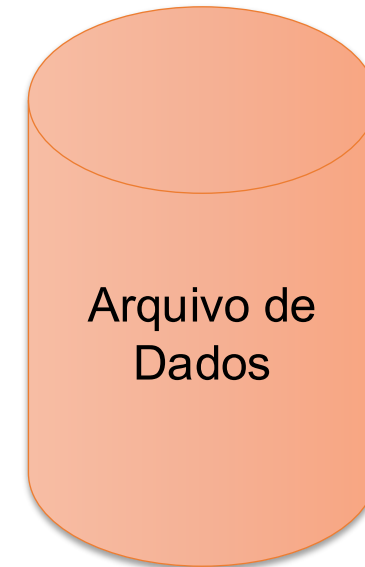


Histórico – Arquitetura nos primórdios



Histórico – Arquitetura nos primórdios

O acesso ao Arquivo era feito de forma direta



Histórico (1956 - 1970)

Cenário mudou drasticamente com o uso dos discos rígidos

Acesso direto ao dados em qualquer lugar do disco

Os bancos de dados poderiam armazenar estruturas de dados mais complexas (como Listas e Árvores)

Um documento revolucionário de Codd (1970) definiu o modelo relacional e os métodos procedurais para consultar os dados

==> Origem os **Bancos de Dados Relacionais**

Computador Lançado em **1956**
Disco com capacidade de **4MB**
Pensando mais de **1 Tonelada**



Histórico (Década de 80)

Os bancos de dados Relacionais não podiam se igualar aos banco de dados de Rede e Hierárquico, devido as suas desvantagens de desempenho.

Um projeto inovador da IBM (System R), levou ao primeiro produto de banco de dados relacional da IBM, o SQL/DS.

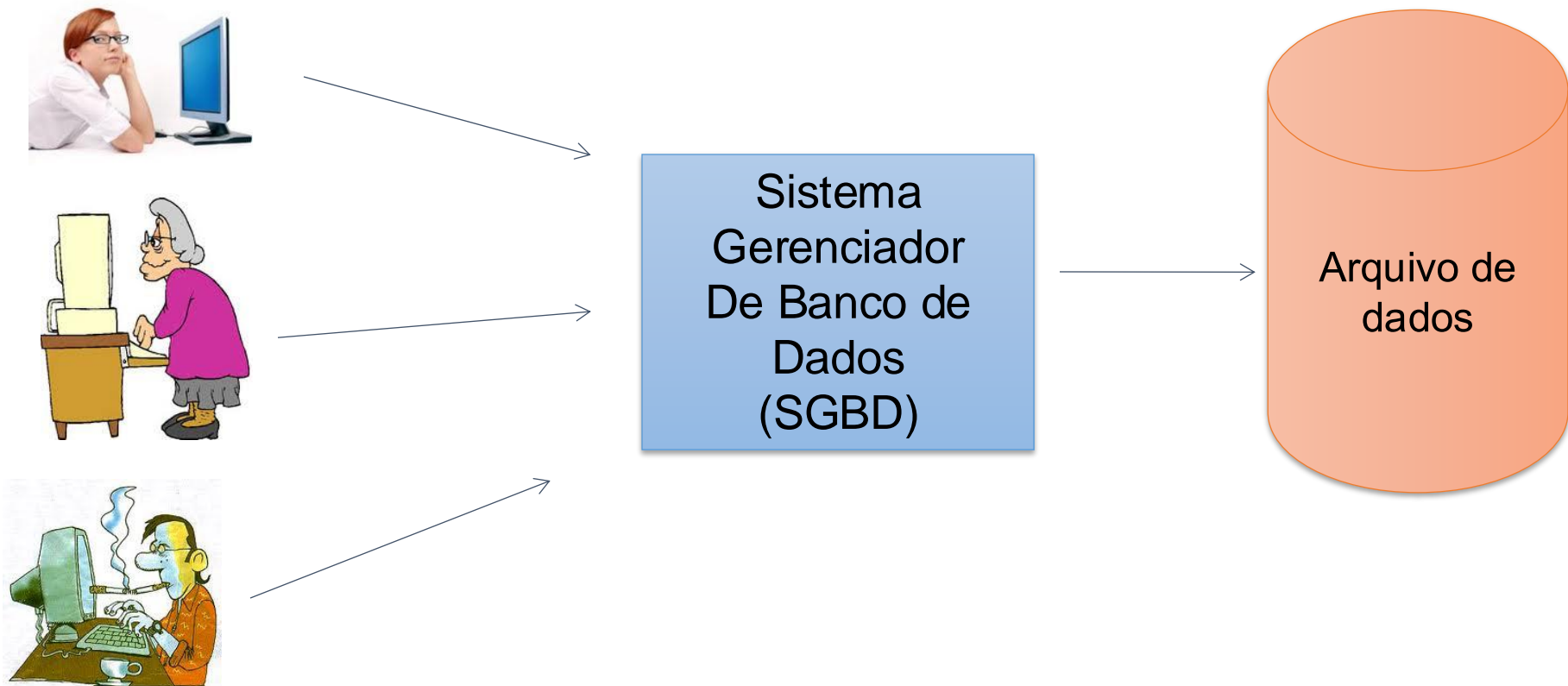
Os primeiros SGBD's relacionais comerciais, como o System R, Oracle, Ingres e DEC RDB, desempenharam papéis fundamentais no melhoramento da performance.

O SGBD's relacionais, por sua facilidade de manipular, começaram a reinar.



Histórico – Arquitetura dos SGBD (1970 a 1980)

- Os aplicativos não tem conhecimento sobre os métodos de Gravação e Leitura física dos dados
- Eles recuperam ou armazenam registro



Histórico (Início da década de 90)

- A linguagem SQL foi projetada para aplicações de suporte à decisão (concentradas na consulta)
- Os bancos de dados da década de 1980 tinham foco nas aplicações de processamento de transação (atualização de dados)
- As ferramentas para analisar grandes quantidades de dados tiveram um elevado crescimento de uso



computador 386 do início dos anos 90

Histórico (Final da década de 90)

Com a explosão de Web, os SGBD's foram mais exigidos:

Propiciar altas taxas de processamento de transação

Alta confiabilidade

Disponibilidade 24x7 (24 horas, 7 dias por semana)

Também precisavam aceitar interfaces Web para
Dados



A amazon.com foi uma das primeiras empresas que começou a usar comércio eletrônico para vender livros na internet (1998) usando SGBD's 24x7



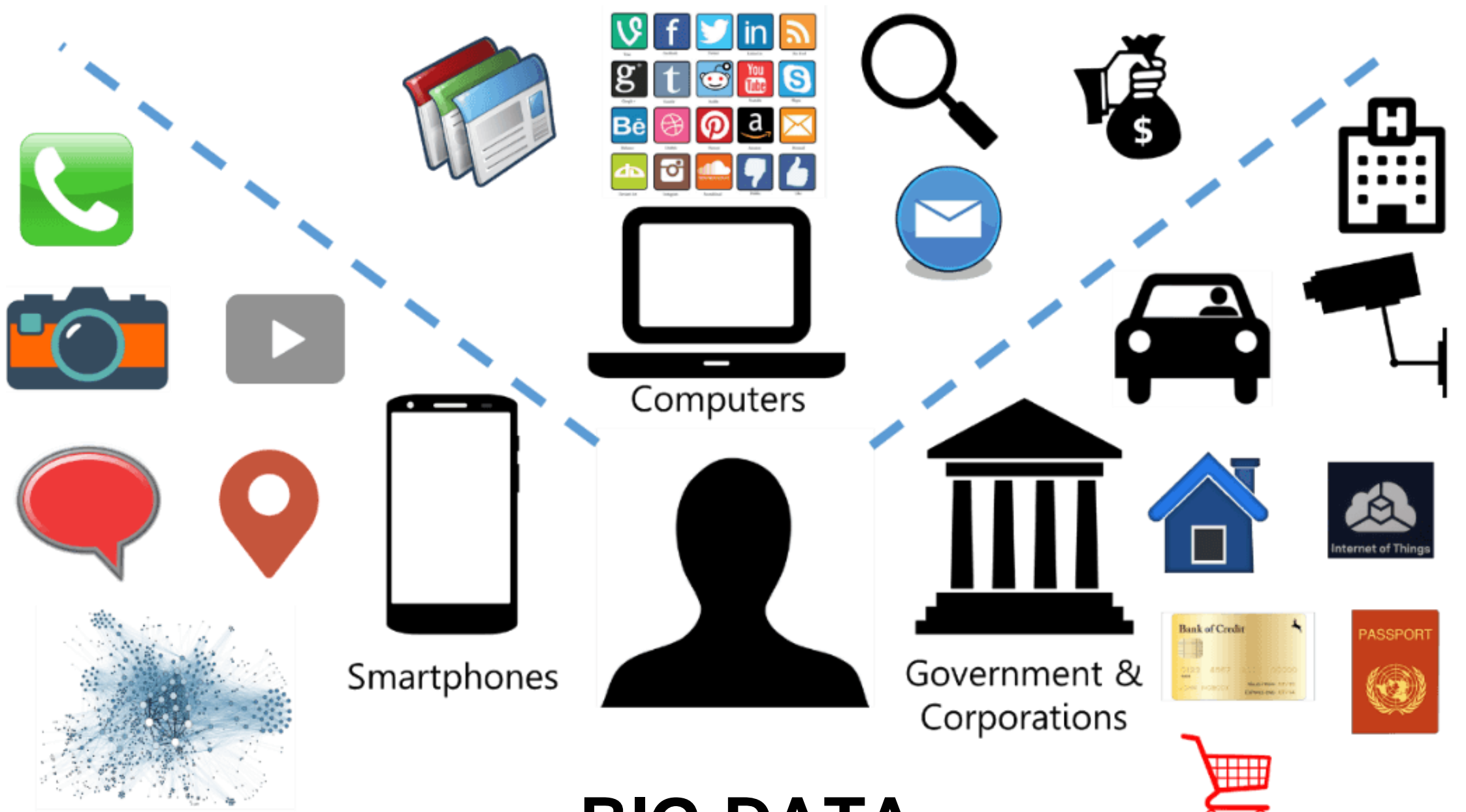
Histórico – Banco de Dados



Recentemente uma grande parte do papel está sendo substituída por armazenamento em mídias, como: SSD, Pendrive, HD, Fitas, CD/DVD, etc .

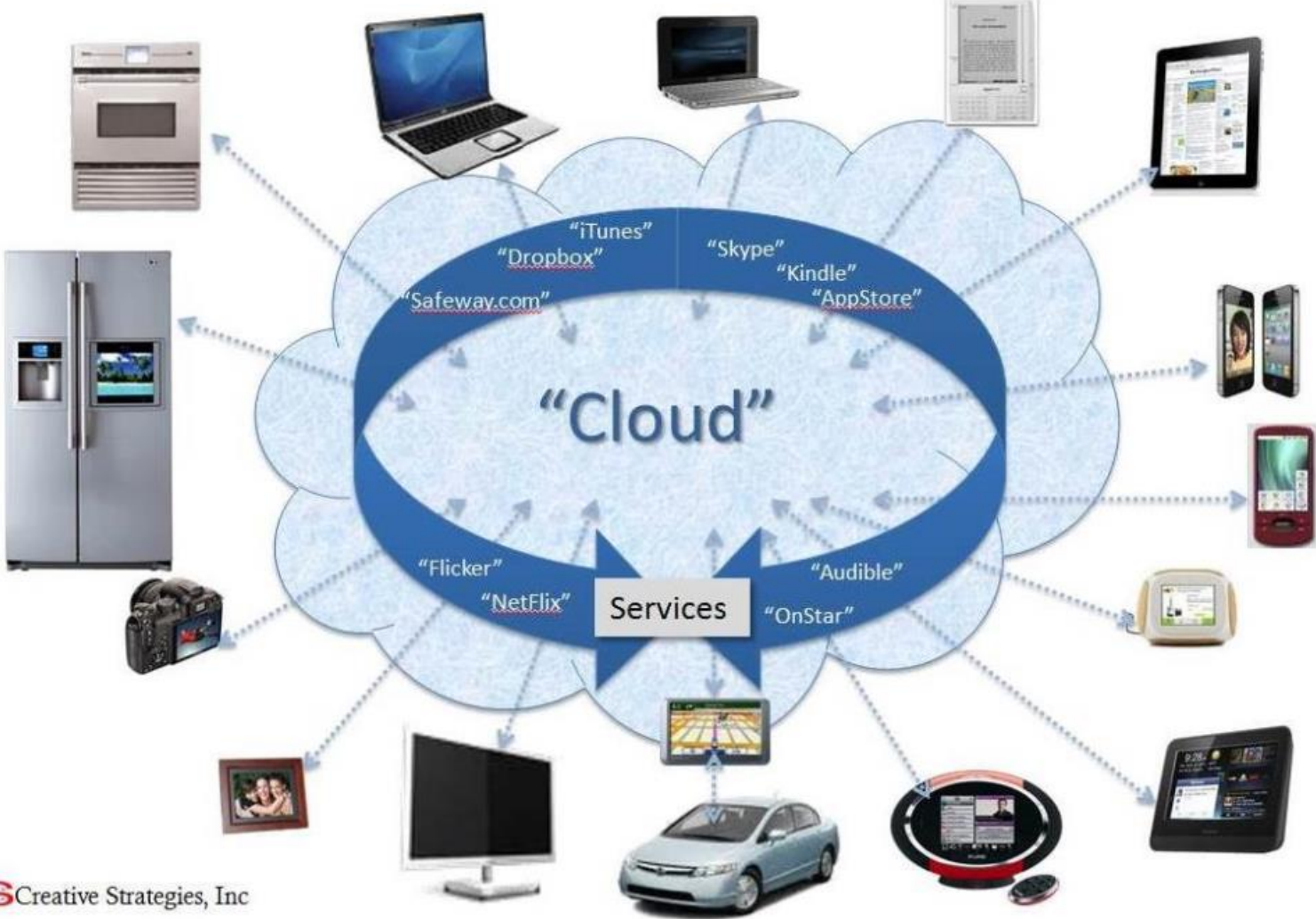


O Volume de dados tem aumentado exponencialmente pela geração de dados de mídias sociais, sensores, armazenamento em nuvem, fotos, áudio, vídeo, etc.



BIG DATA

Internet of Things



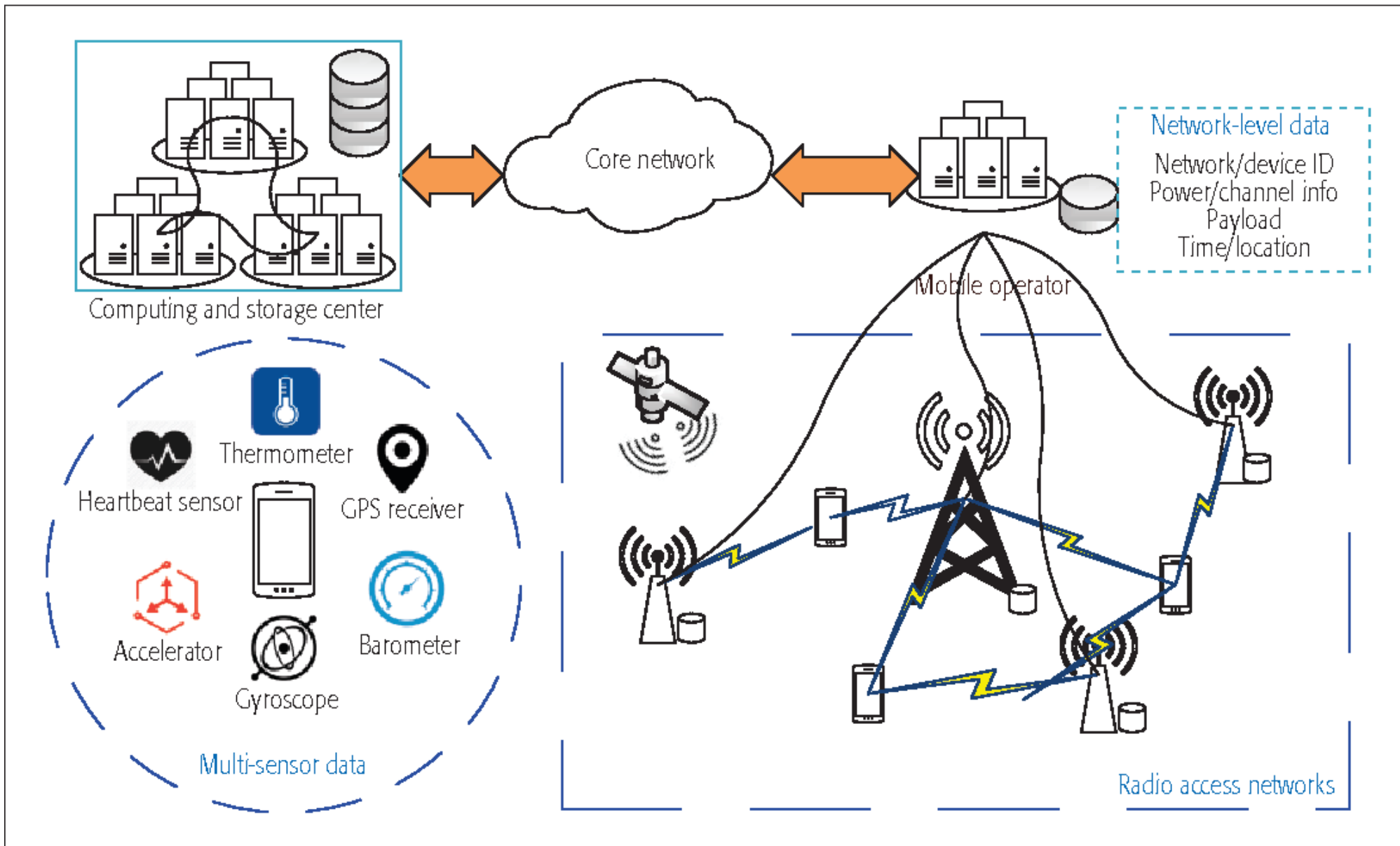
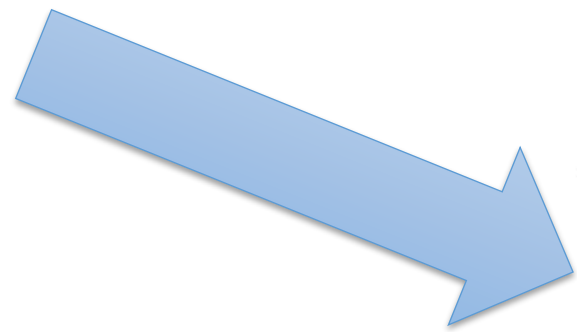


FIGURE1. Overview of mobile big data.

Sistema de Recomendação



Passa o mouse para ampliar a imagem

Sobre este item

- 165Hz de Taxa de Atualização: Desfrute de uma experiência ultra-suave com a impressionante taxa de atualização. Experimente gráficos super fluidos e responsivos no próximo nível!
- 1ms de Tempo de Resposta (MPRT): Seus adversários com um tempo de movimento é instantâneo, proporcionando a vantagem para vencer.
- Tecnologia AMD FreeSync: Com a tecnologia AMD FreeSync, o jogo flui de maneira perfeita, eliminando o tearing e a imagem. Tenha uma experiência visual suave e sem interrupções.
- Base ergonômica ajustável: Jogue confortavelmente com a base ergonômica ajustável. Ajuste a altura, o ângulo e a posição para obter a posição de visualização ideal sem desconforto.
- Modo Mira e Shadow Control: Aprimore sua precisão com as soluções de software Agon Eye e Shadow Control. O Shadow Control funciona como um filtro de luz e ilumina as áreas escuras, garantindo uma visualização detalhada em qualquer situação.
- Tecnologia Flicker Free: Proteja seus olhos com a tecnologia Flicker Free, que elimina o incômodo causado por piscadas mais tempo com conforto e sem cansaço.

[Relatar um problema com este produto](#)

Monitor Dell Widescreen SE2216H
4.8 ★★★★★ 1.717
R\$1.408,59 **prime**

Patrocinado

Frequentemente comprados juntos



Este item: Monitor Gamer AOC Hero 24" 165Hz 1ms Ajuste de Altura 24G2S/BK
R\$882³⁴

Clamper Energia 5 Tomadas
R\$81⁸¹

ELG F80N - Suporte Ergonômico Articulado de Mesa com Pistão a Gás, Design Compacto e Retrátil...
R\$204⁶³

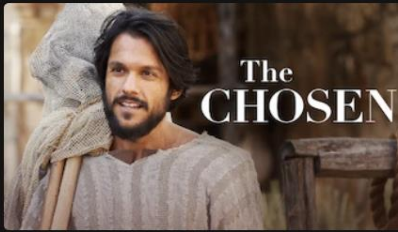
Preço total: R\$ 1.168,78

Adicionar todos os 3 ao carrinho

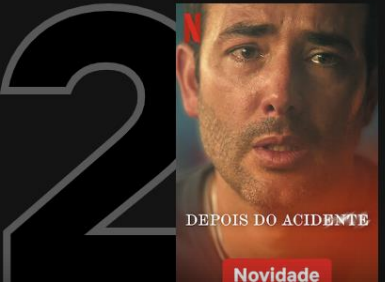
Descubra suas próximas histórias



Séries dos EUA



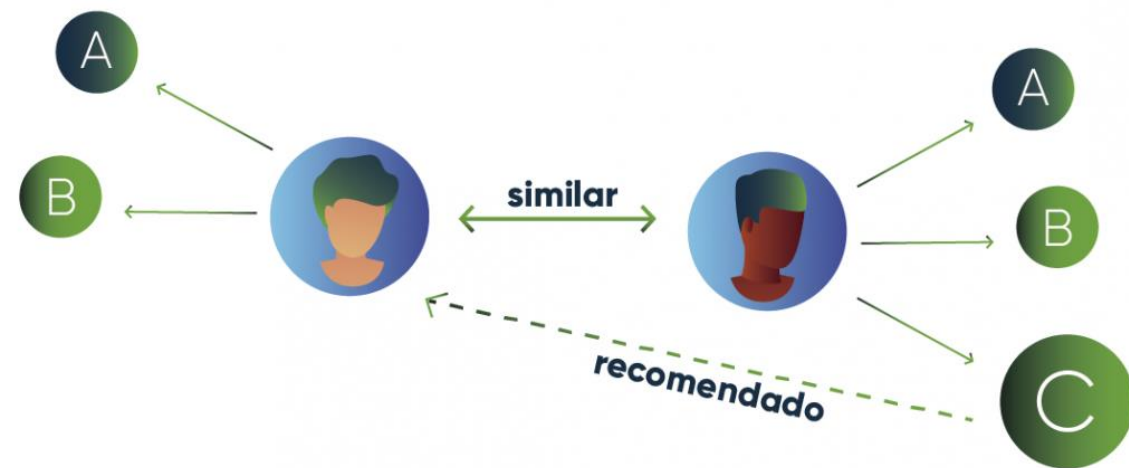
Brasil: top 10 em séries hoje



Sistema de Recomendação

O Paradoxo da Escolha:
"Riqueza de Informação => Pobreza de Atenção"

Sistema de recomendação Filtragem colaborativa



Histórico (Ano 2.000 - Hoje)

Machine Learning – Deep Learning

Big Data

Cloud Computing

Data Science

NoSQL (SGBD não Relacionais)

Engenharia de Dados (Hadoop, Spark)



Era do Big Data - Aprendizado de Máquina

Grande Volume de Dados
Evolução das Técnicas Analíticas
Análise de Dados em *Real Time*
Desenvolvimento de Aplicações
Inteligentes



Tecnologia Atual

O avanço da automação

Avanço de tecnologias que chamamos de "Machine Learning" e "Deep Learning"

Impacto da Robotização

Dispositivos vestíveis

Carros Autônomos

Visão Computacional (Reconhecimento de objetos, pessoas)

Tradução automática

Reconhecimento de voz e de texto,

Grandes Modelos de Linguagem - LLM (ChatGPT, etc.)

Computação Quântica

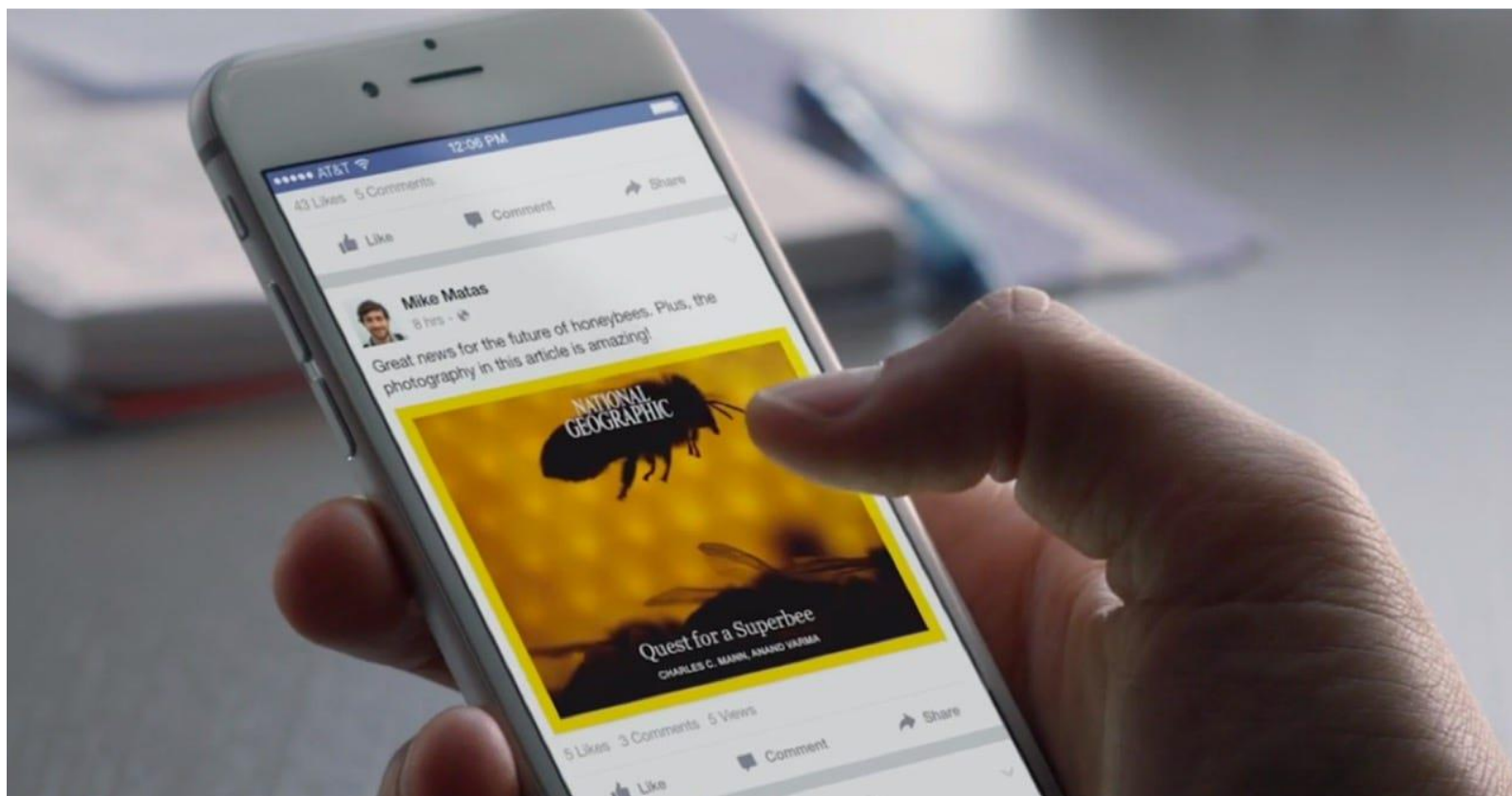
Aprendizado de Máquina (Machine Learning)

- Usamos algoritmos de aprendizado de máquina várias vezes por dia sem saber
- Busca em sites como o [google.com](https://www.google.com), e a sua razão de funcionarem tão bem está fundamentada em algoritmos de aprendizado de máquina
- Sites de serviços como amazon.com e Netflix usam sistemas de recomendação e aprendem suas preferências a cada clique ou compra que fizermos

Aprendizado de Máquina (Machine Learning)

- Quando o filtro de spam do seu email filtra automaticamente toneladas de mensagens indesejadas, isso também é um algoritmo de machine learning.





Machine Learning

Facebook usa para conseguir entender o conteúdo das fotos que você posta por lá.



Personalização do feed de notícias

Reconhecimento de imagem e vídeo

Tradução de idiomas

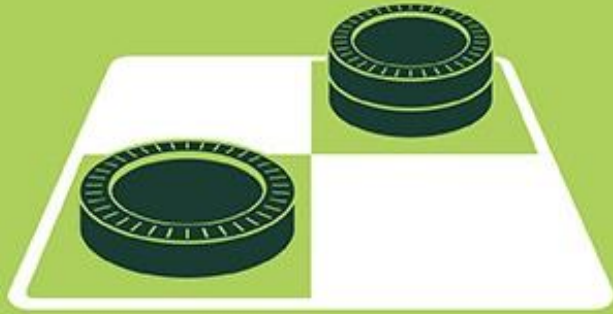
Detecção de spam

Detecção de notícias falsas

Análise Preditiva

ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



MACHINE LEARNING

Machine learning begins to flourish.



DEEP LEARNING

Deep learning breakthroughs drive AI boom.



1950's

1960's

1970's

1980's

1990's

2000's

2010's



Aprendizado de Máquina (Machine Learning)

Machine Learning é a utilização de algoritmos para extrair informações de dados brutos e representá-los através de algum tipo de modelo matemático.

Usamos então este modelo para fazer inferências a partir de outros conjuntos de dados.

- É um subcampo da Inteligência Artificial que permite dar aos computadores a habilidade de aprender sem que sejam explicitamente programados para isso



Aprendizado Profundo (Deep Learning)

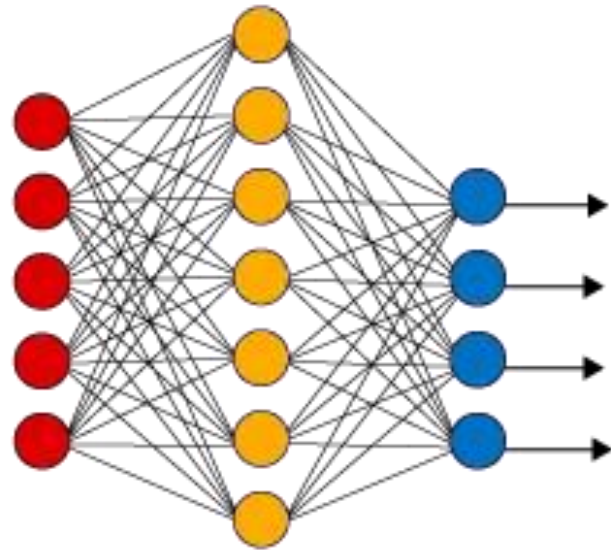
- é uma sub-área da Aprendizagem de Máquina, que emprega algoritmos para processar dados e imitar o processamento feito pelo cérebro humano

Deep Learning usa camadas de neurônios matemáticos para processar dados, compreender a fala humana e reconhecer objetos visualmente.

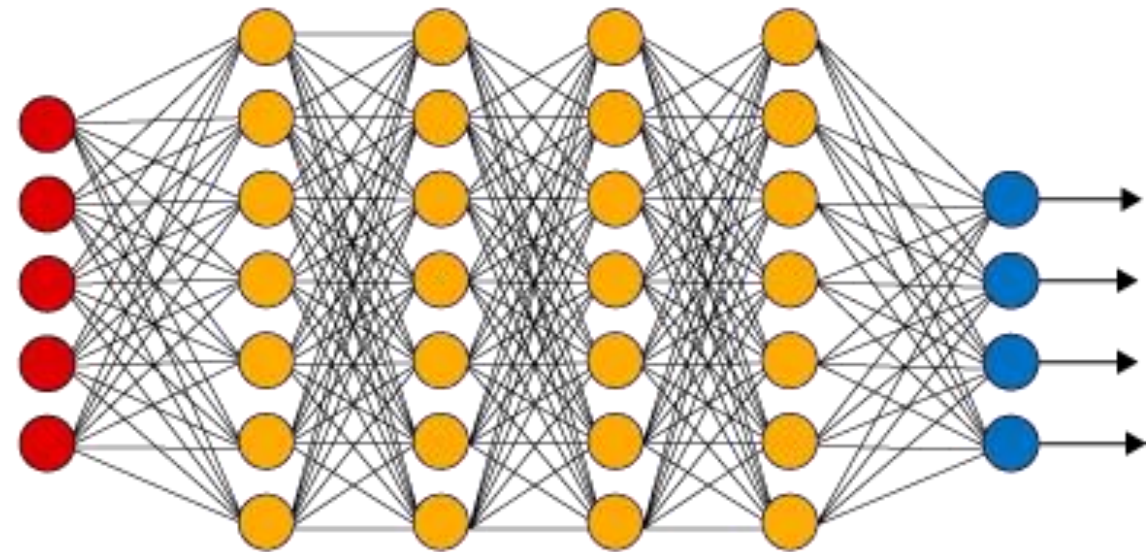
A informação é passada através de cada camada, com a saída da camada anterior fornecendo entrada para a próxima camada.

Aprendizado Profundo (Deep Learning)

Simple Neural Network



Deep Learning Neural Network



● Input Layer

● Hidden Layer

● Output Layer

Aprendizado Profundo (Deep Learning)

- A aprendizagem profunda é responsável por avanços recentes em visão computacional, reconhecimento de fala, processamento de linguagem natural e reconhecimento de áudio.
- O aprendizado profundo é baseado no conceito de redes neurais artificiais, ou sistemas computacionais que imitam a maneira como o cérebro humano funciona.
- A extração de recursos é outro aspecto da Aprendizagem Profunda. A extração de recursos usa um algoritmo para construir automaticamente “recursos” significativos dos dados para fins de treinamento, aprendizado e compreensão.
- Normalmente, o Cientista de Dados, ou Engenheiro de IA, é responsável pela extração de recursos.



Speech Recognition



Facial Recognition

**MACHINE
LEARNING**



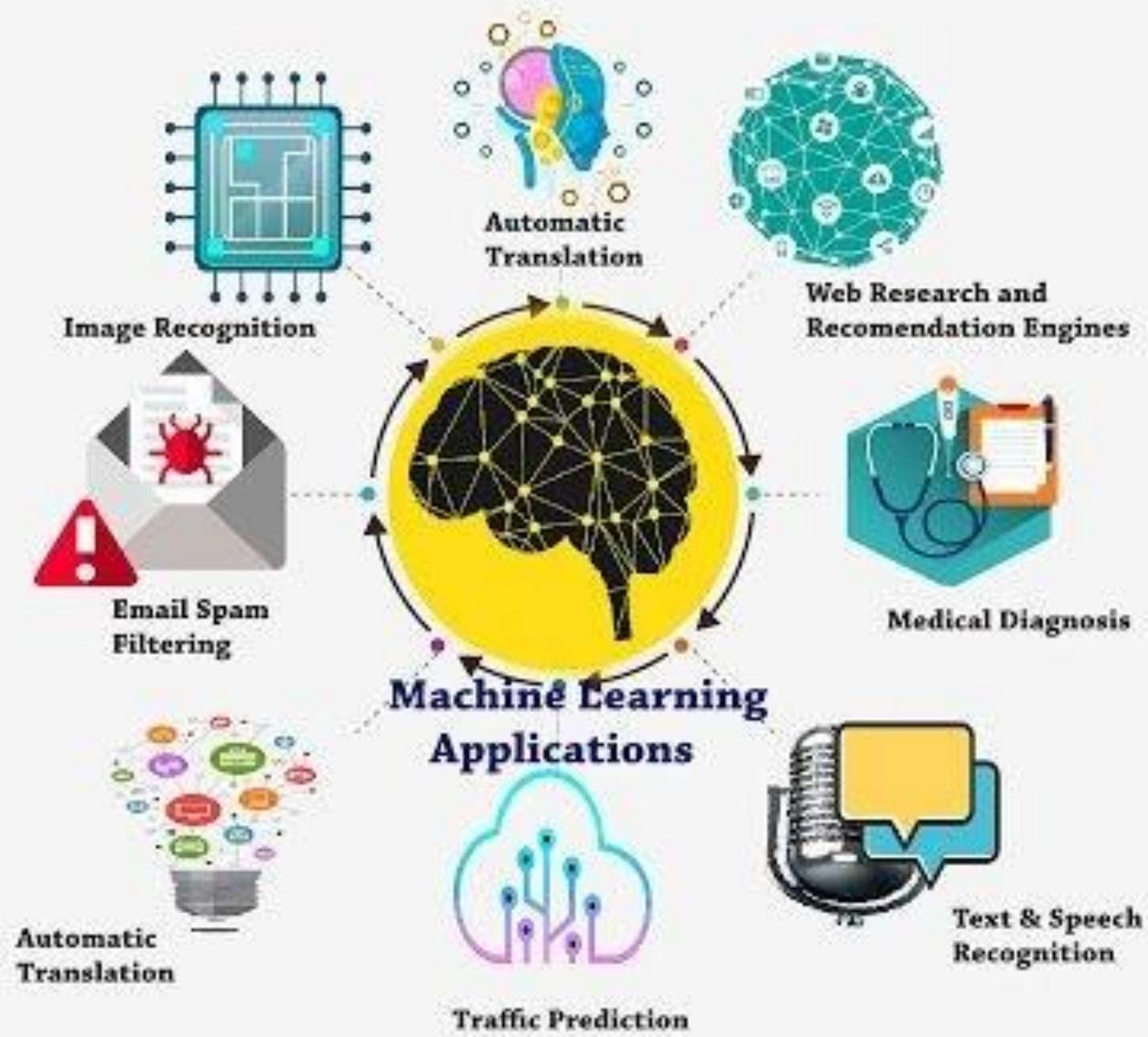
Optical Character Recognition



Image Recognition



Autonomous Driving

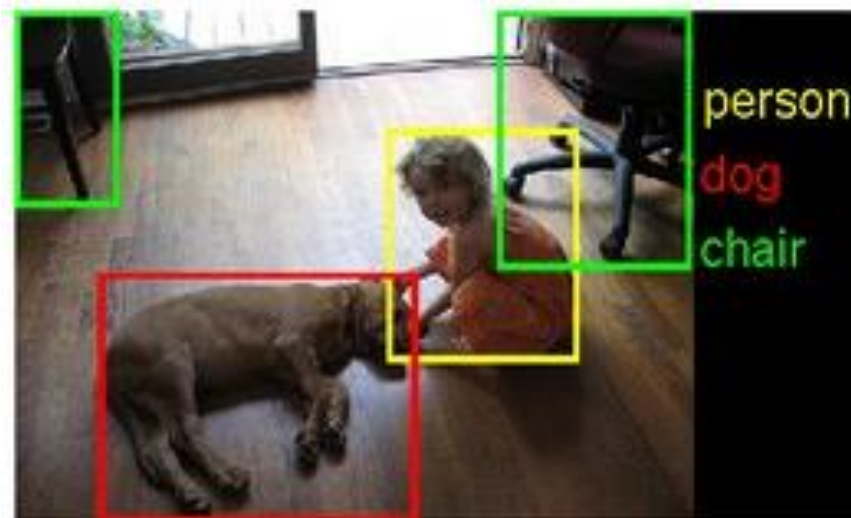
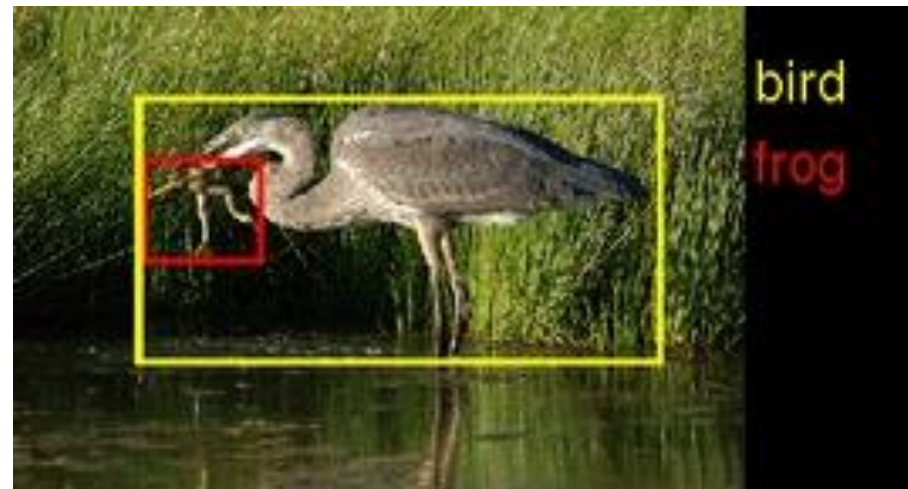


Machine Learning / Deep Learning

- Visão Computacional
- Reconhecimento de Padrões
- Processamento de Linguagem Natural
- Tradução automática
- Grandes Modelos Geradores de Linguagem – LLM
ChatGPT (OpenAI), BERT (Google) e LLaMA (Meta AI).

Visão Computacional

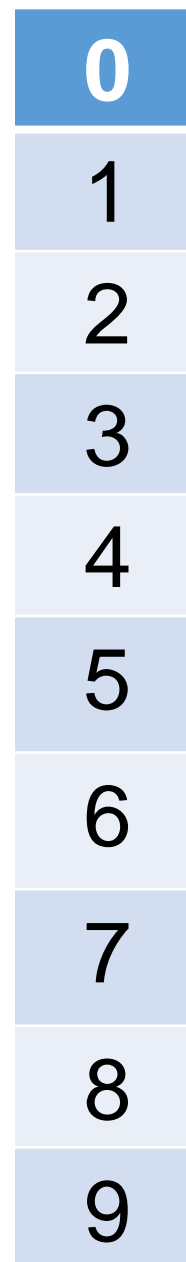
Visão Computacional

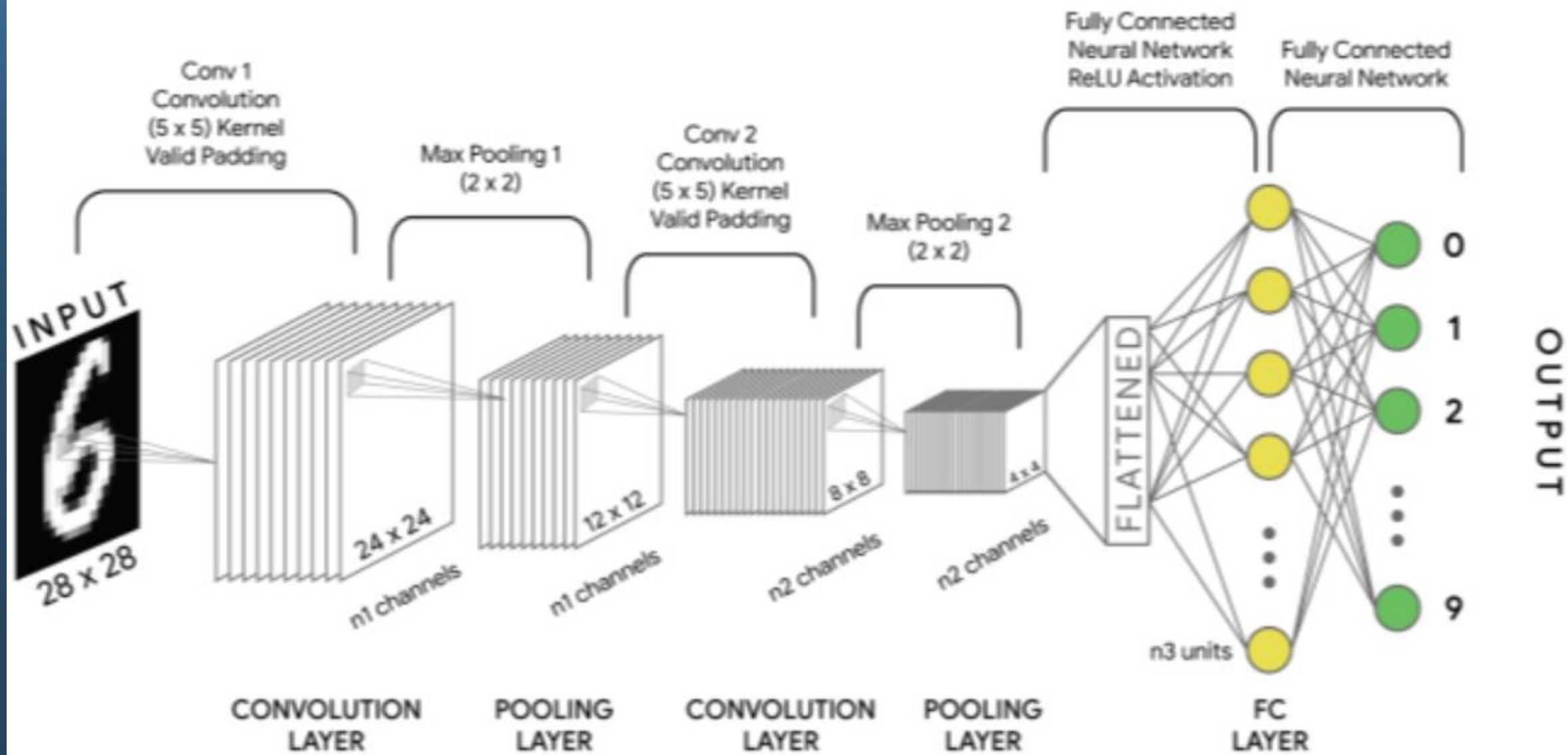


Entrada



Saída



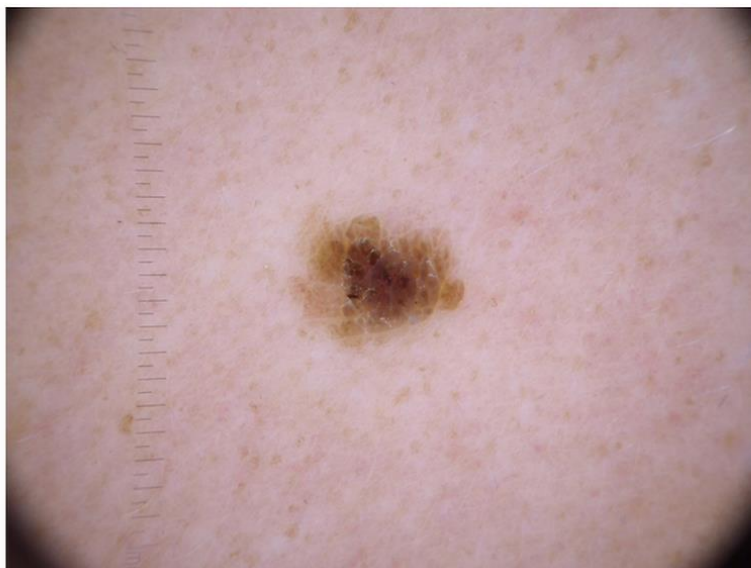


Social Distancing



Use of masks





Diagnóstico Diferencial

Melanoma - 82%

Carcinoma de Células Basais - 12%

Nevo Melanócito (Sinal) - 4%

Carregue uma imagem ou tire uma foto:

Diagnosticar Lesão

Ver outras Lesões

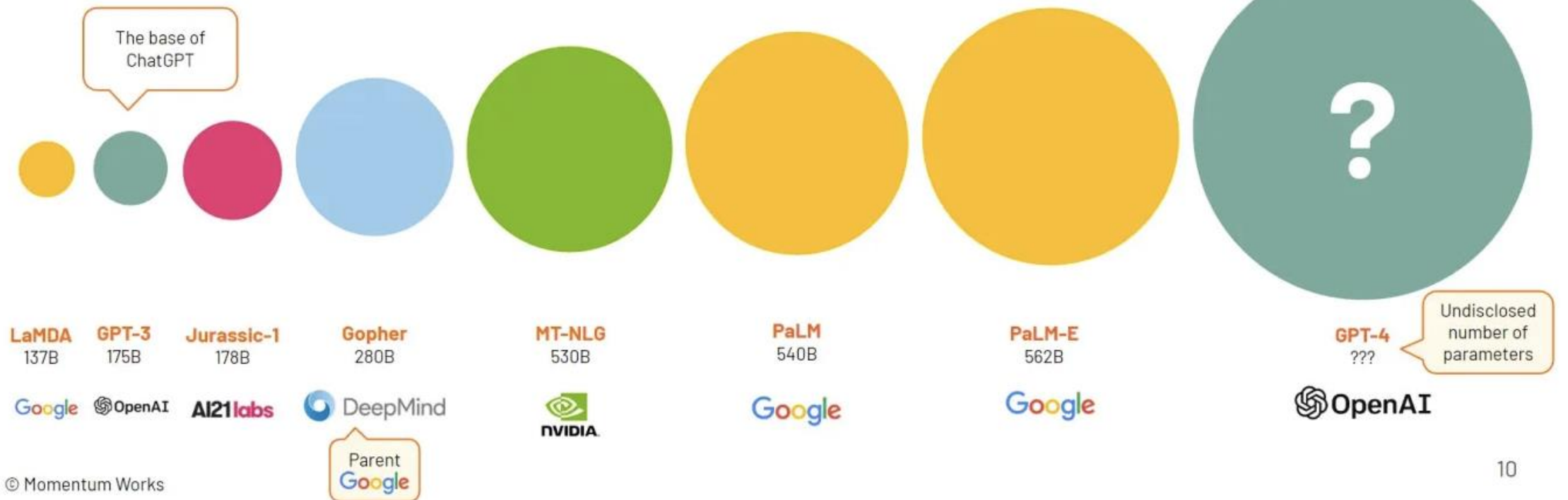
ESTE RESULTADO NÃO SUBSTITUI A AVALIAÇÃO DO MÉDICO. Este é um sistema para auxílio diagnóstico de doenças da pele usando redes neurais. Todo o processamento é feito no seu dispositivo e as imagens não são enviadas para o seu servidor. Ao continuar, você assume toda a responsabilidade com o uso.

Por



Grandes Modelos Geradores de Linguagem – LLM

Large models (>100b parameters)





ChatGPT ▾



ChatGPT

Explore GPTs

Today

Chat_2024_Setembro ...

chat_2023

Pet Nap Illustration Request

Dados Gerados Por Pessoa

Enumere os planetas do sistema solar.



Claro! Aqui estão os planetas do sistema solar, listados em ordem de distância do Sol:

1. Mercúrio
2. Vênus
3. Terra
4. Marte
5. Júpiter
6. Saturno
7. Urano
8. Netuno

Lembrando que Plutão foi reclassificado como um planeta anão pela União Astronômica Internacional em 2006.



Message ChatGPT



Upgrade plan

Get GPT-4, DALL-E, and more

ChatGPT can make mistakes. Check important info.



[+ Nova conversa](#)

Sugestões turismo Camp... ⋮

Sabiá-3 ▾



Estou passeando pela cidade de Campinas, Brasil. Contudo, não conheço esta cidade. Sugira pontos turísticos pouco conhecidos. Contudo, primeiro, me faça três perguntas que te ajudarão nesta tarefa.



Claro, para oferecer sugestões mais personalizadas, gostaria de saber:

1. Você tem interesse em atrações culturais, como museus ou galerias de arte?
2. Prefere atividades ao ar livre, como parques ou trilhas?
3. Está buscando locais históricos ou arquitetura de destaque na cidade?

museus



Entendido, se você está interessado em museus e atrações culturais, aqui estão algumas sugestões de locais em Campinas que talvez não sejam tão conhecidos pelo grande público, mas que oferecem experiências interessantes:

1. **Museu de Arte Contemporânea de Campinas (MACC)** - Embora o MACC seja conhecido entre os entusiastas da arte, ele pode não estar na lista dos grandes pontos turísticos. Possui uma coleção significativa de arte contemporânea.
2. **Casa de Vidro** - Projetada pelo arquiteto Ruy Ohtake, a Casa de Vidro é uma residência que se tornou um espaço cultural e de exposições. Além da arte exposta, a própria casa e seu jardim são obras de arte a céu aberto.



Envie uma mensagem



imagem.

Machine
Learning /
Deep
Learning



MONITORING DAMS THROUGH ARTIFICIAL INTELLIGENCE (DEEP LEARNING)

Boqueirão Dam Volume: 01/01/2009–08/01/2025



$R^2 = 0.999890$
RMSE = 0.006577
A Flood Forecast
Year: **2025**

TROPICAL DISEASE FORECAST BY TWITTER

750 Million tweets per day in the World (2019)

Disease: **Yellow Fever**

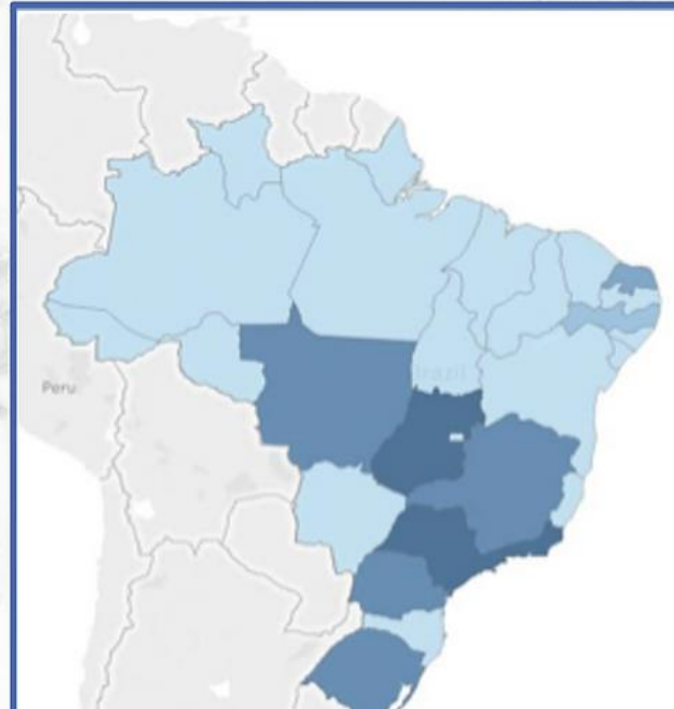
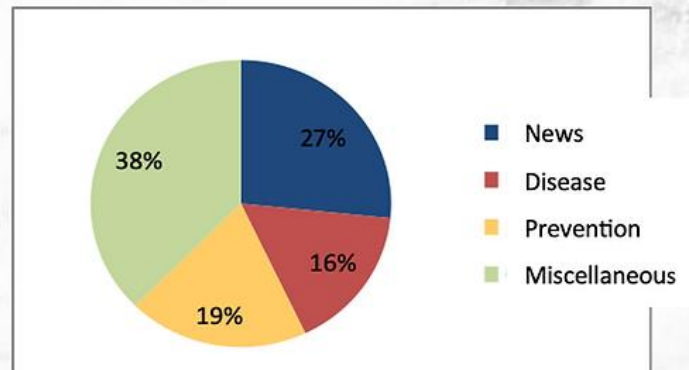
Data: 100,000 tweets

Period: 3 months collection - Year: 2018

Model: Machine Learning - Bayes

Model Accuracy: 0.89

Data compared to the Bulletins
From the Ministry of Health



Sexo
Masculino

Idade (categoria)
30-34

Etnia
Branco

Diabetes
Não

Atividade Física nos últimos 30 dias
Não

Estado Geral de Saúde
Razoável

Quantas horas você dorme por dia
7.00

Você tem asma?
Não

Doença renal (cálculo, incontinência, etc.)
Não

Câncer de pele
Não

Prever

Sistema de Auxílio à Previsão de Problemas Cardíacos



Previsão: Normal - Probabilidade: 78%

ESTE RESULTADO NÃO SUBSTITUI A AVALIAÇÃO DO MÉDICO. Este é um sistema para auxílio diagnóstico cardíaco usando o algoritmo XGBOOST. Todo o processamento é feito no seu dispositivo e as imagens não são enviadas para o seu servidor. Ao continuar, você assume toda a responsabilidade com o uso.

Sistema de Auxílio à Previsão de Problemas Cardíacos XGBOOST versão 1.0.0.

Desenvolvido por Prof. Dr. Vladimir Costa de Alencar e Equipe de Pesquisadores do LANA/UEPB. Campina Grande, Paraíba, Brasil, 2023.

valencar@gmail.com

Dashboard dados ENEM Paraíba - 2020

Numero de Inscritos

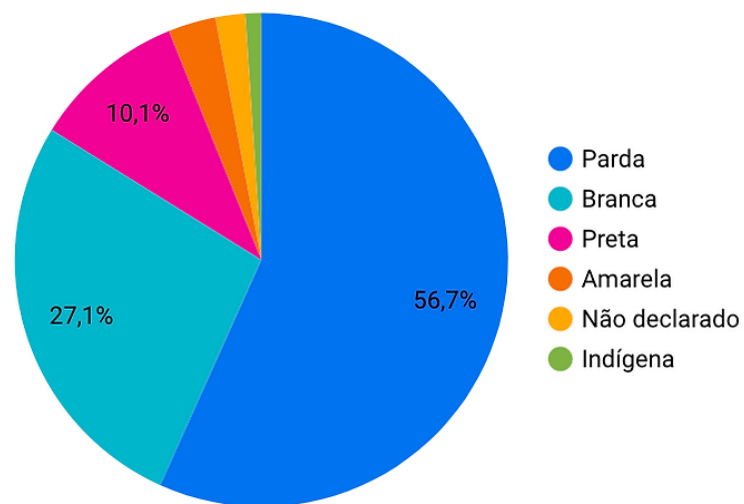
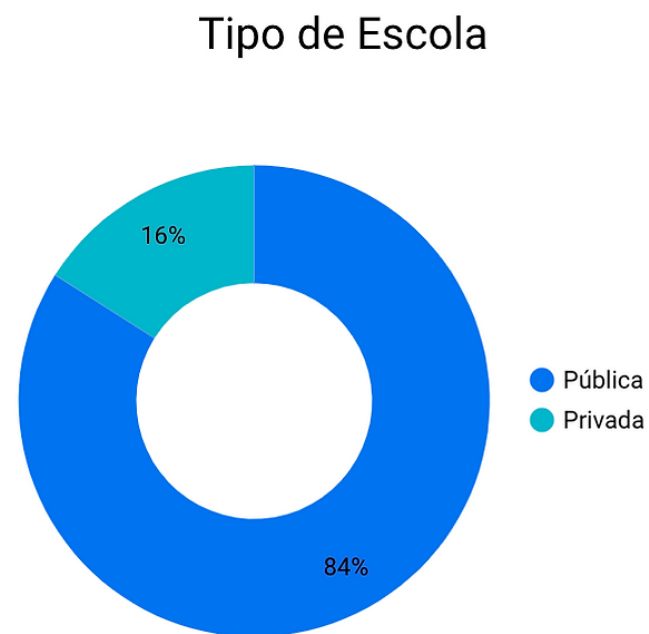
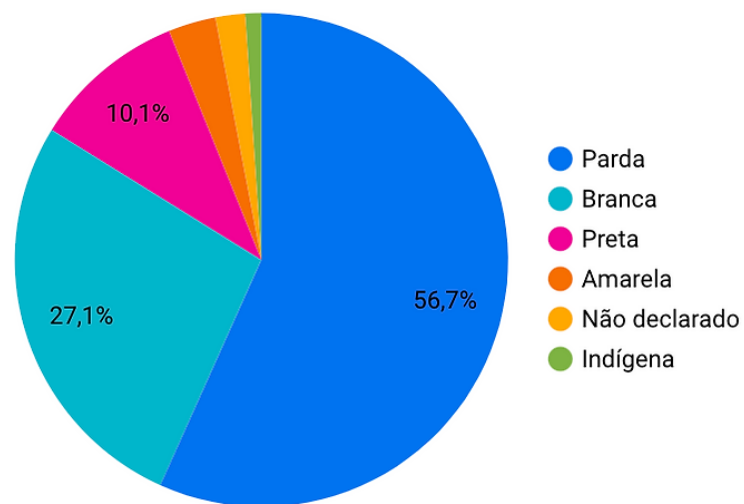
149.092

IDHM média

0,61

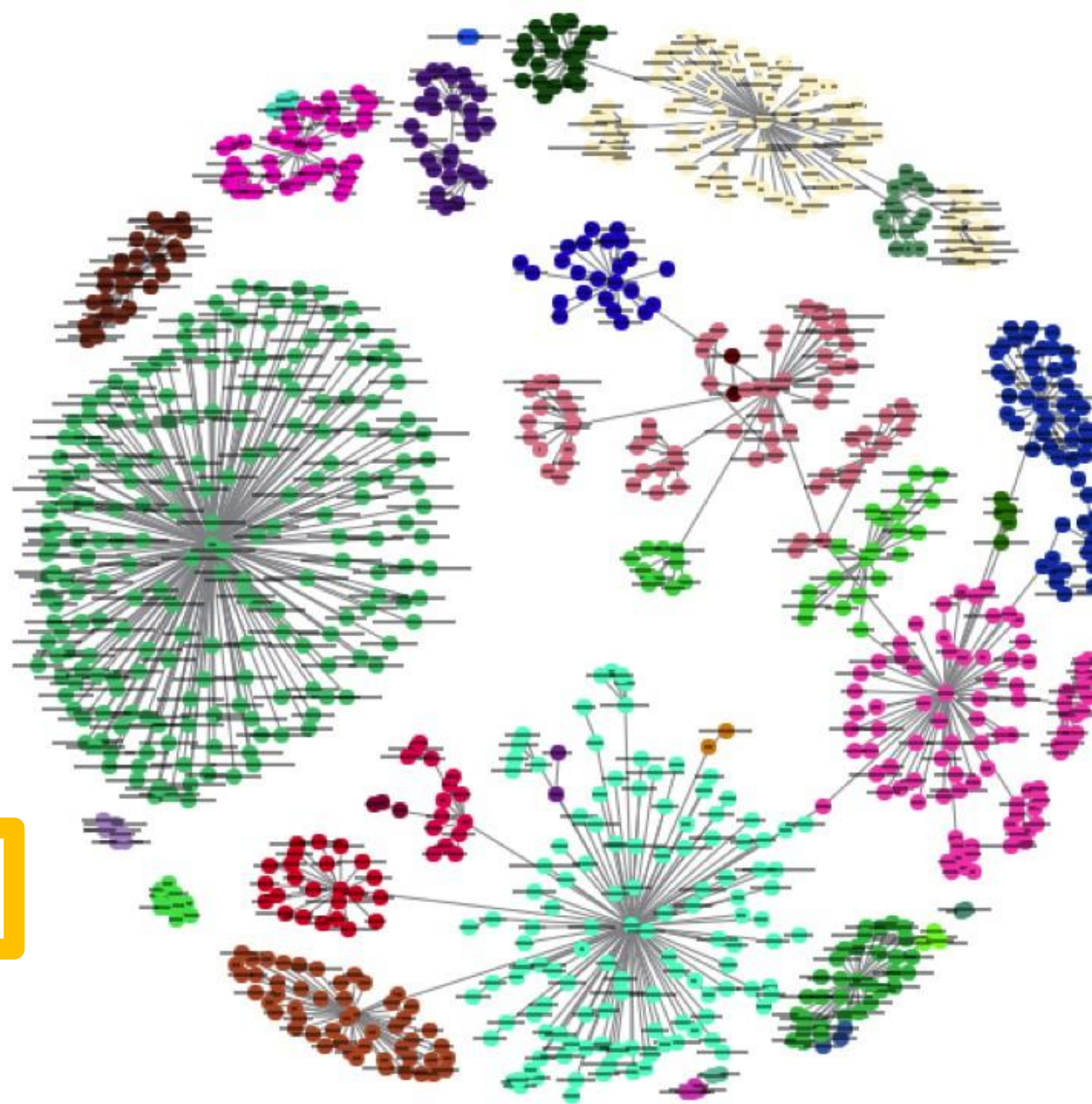
PIB per Capita

R\$ 11.987

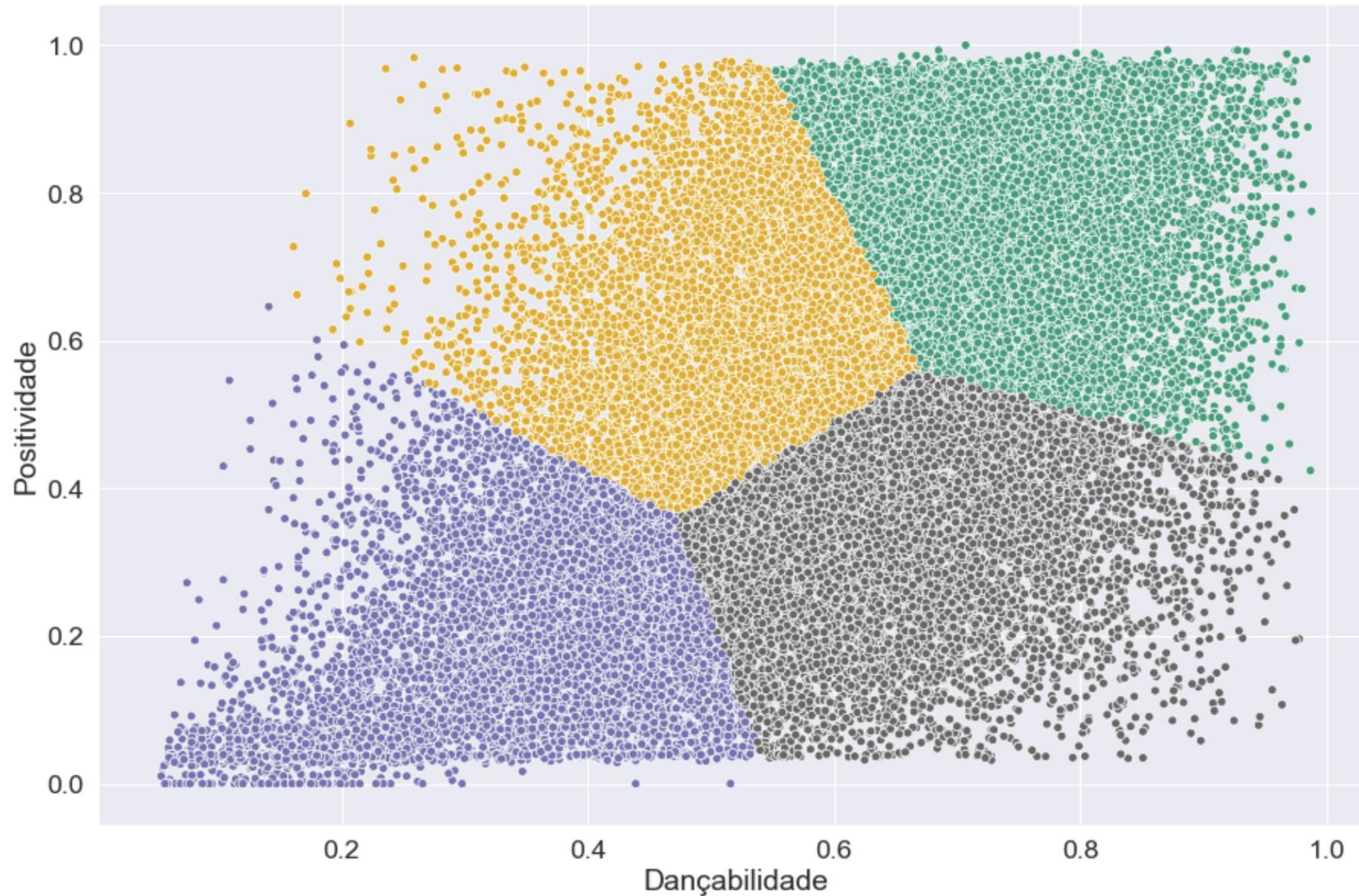


Clusterização- Determinação de agrupamentos

Influenciadores Digitais



Spotify - 30.000 músicas (USA, Brasil, França, Japão)



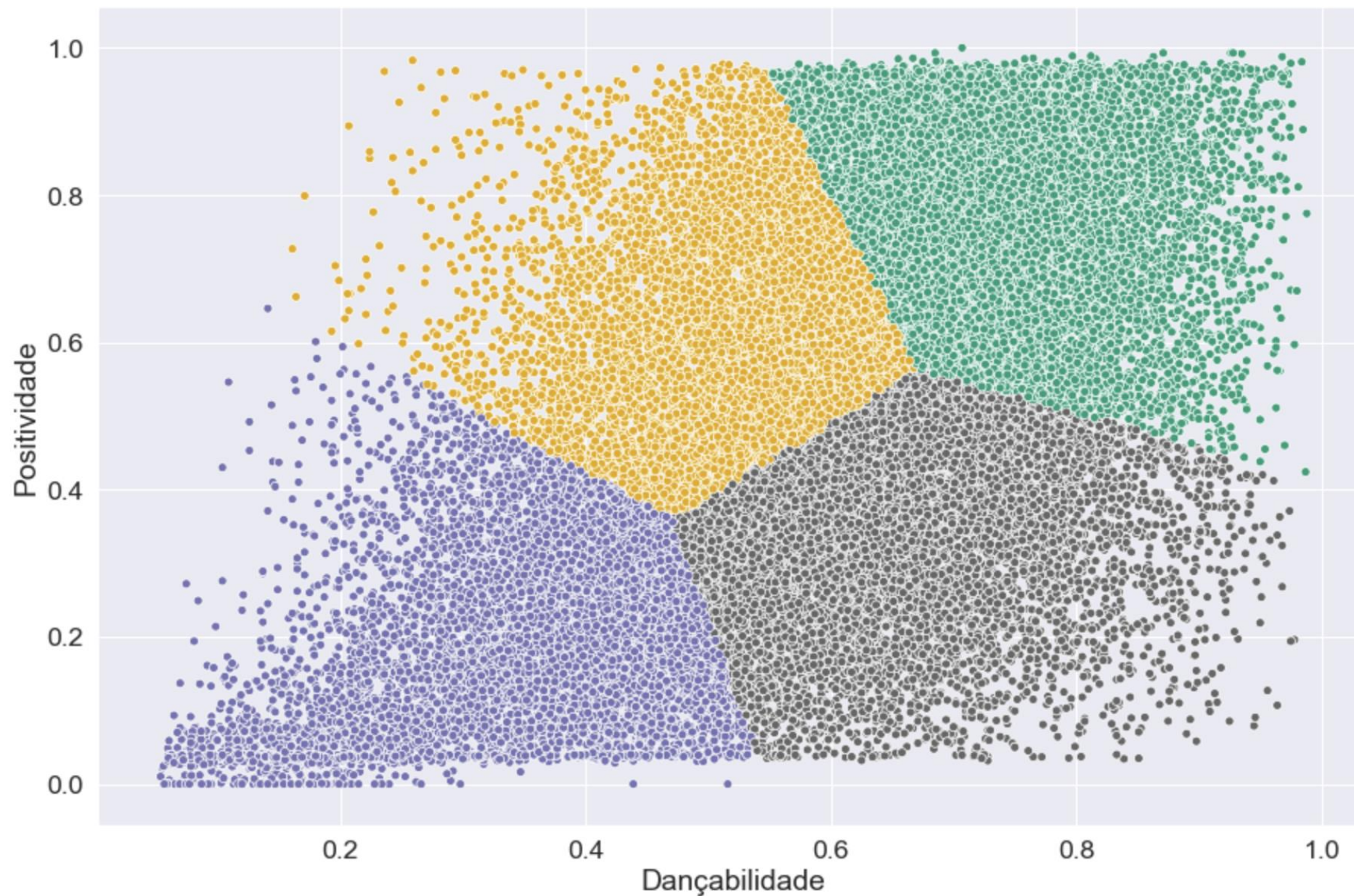
Spotify - 30.000 músicas (USA, Brasil, França, Japão)

artists	category	country	duration_ms	id	name	popularity	uri	danceability	energy	liveness
Pablo Vittar	toplists	BR	157928	29caulxU0jxucdVlyhSJTc	Amor de Que	0	spotify:track:29caulxU0jxucdVlyhSJTc	0.737	0.819	0.3600
Anitta	toplists	BR	194771	2qD7VoDYcrAPY0cVEfpsR5	Some que ele vem atrás	68	spotify:track:2qD7VoDYcrAPY0cVEfpsR5	0.648	0.795	0.3800
Dua Lipa	toplists	BR	183290	6Wri0LAC5M1Rw2MnX2ZvEg	Don't Start Now	51	spotify:track:6Wri0LAC5M1Rw2MnX2ZvEg	0.794	0.793	0.0952
Ariana Grande	toplists	BR	171840	7vJS1DPc3FzBtqBs8n3mW5	Bad To You (with Normani & Nicki Minaj)	0	spotify:track:7vJS1DPc3FzBtqBs8n3mW5	0.728	0.581	0.0983
Alok	toplists	BR	144000	6ZHW34diMINGGVwcrOfkr	I Don't Wanna Talk (feat. Amber Van Day)	21	spotify:track:6ZHW34diMINGGVwcrOfkr	0.747	0.861	0.2910

Spotify - 30.000 músicas (USA, Brasil, França, Japão)

	artists	category	country	duration_ms		id	name	popularity		uri	danceability	energy
67434	James Mason	word	US	1048466		2CUuFtUxAd95SE8GB6wS62	Andrea Del Sarto	0		spotify:track:2CUuFtUxAd95SE8GB6wS62	0.533	0.0216
67435	Richard Mitchley	word	US	185620		7zXgxqQGkg0FMI1CHXs9dA	My Last Duchess - Robert Browning	5		spotify:track:7zXgxqQGkg0FMI1CHXs9dA	0.679	0.2900
67436	Ghizela Rowe	word	US	57568		3SvAagCW5MeMmylOqHf8kh	Now - Robert Browning	0		spotify:track:3SvAagCW5MeMmylOqHf8kh	0.714	0.0624
67437	Richard Mitchley	word	US	112795		29Sh5blaEs2LyW4hTECPBV	In 3 Days - Robert Browning	0		spotify:track:29Sh5blaEs2LyW4hTECPBV	0.718	0.2670
67438	Richard Mitchley	word	US	54613		2ID3Djap997S1JuDJgL6t8	Life in a Bottle - Robert Browning	0		spotify:track:2ID3Djap997S1JuDJgL6t8	0.734	0.1910

Spotify - 30.000 músicas (USA, Brasil, França, Japão)



Machine Learning - Algoritmos

- Regressão Linear
- Árvores de Decisão
- Naive Bayes
- KNN
- K-Means
- Random Forest
- SVM
- Redes Neurais
- Deep Learning



Histórico (Ano 2.000 – Dias Atuais)

- Machine Learning – Deep Learning
- Big Data
- Data Science
- NoSQL (SGBD não Relacionais)
- Engenharia de Dados (Hadoop, Spark)