ChatGPT em Ciência de Dados

Análise ao Artigo

UC | Inteligência Artificial Simbólica para Ciência de Dados **Docente |** Isabel Machado Alexandre

Grupo 17

André Silvestre N° 104532 Maria Margarida Pereira N° 105877

CDC1 & CDC2



Evolução Histórica **ChatGPT**











A Microsoft tornou-se o seu principal investidor.

Criado por investigadores da **Google**

Em apenas 5 dias, teve 100M de utilizadores

da OpenAl



dezembro de 2015



junho de 2017





novembro de 2022





março de 2023

Evolução Histórica

Tasks

Question



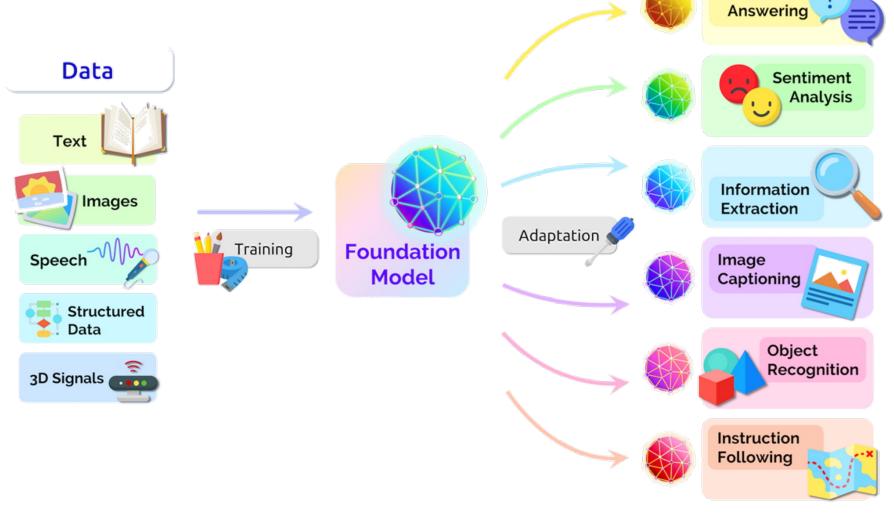


Figura 1 | Ilustração dos Modelos *Transformer*. Fonte: NVIDIA (2022)

Evolução Histórica



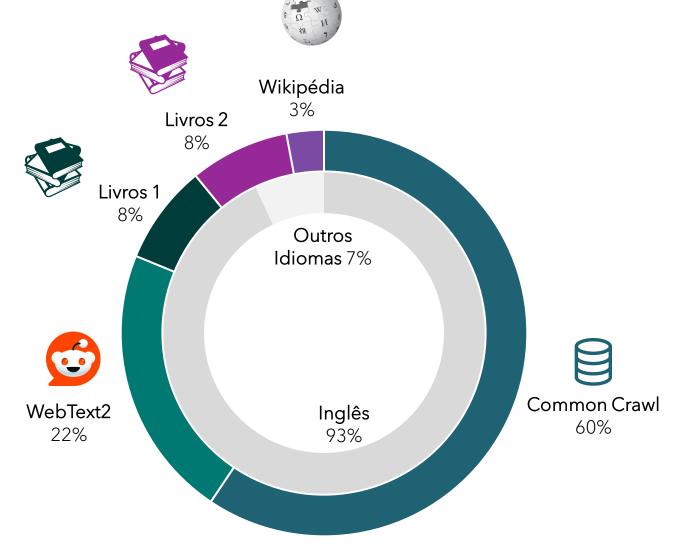


Figura 2 | Fontes do dataset de treino do ChatGPT-3.

Fonte: Representação Gráfica da <u>Tabela 2.2 do Artigo da OpenAl</u> (2020)

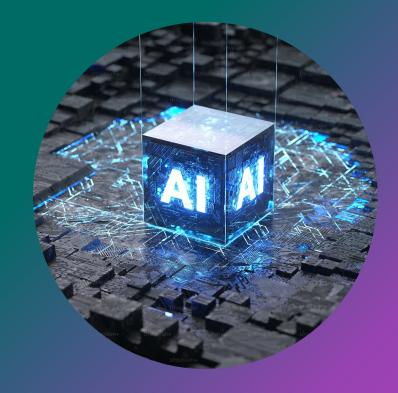




Explora o impacto do ChatGPT na Ciência de Dados, destacando como esta ferramenta influencia a programação, a solução de problemas e a ética na área

Imagem do **Unsplash**

Síntese do Artigo



- > Imprecisões
- > Privacidade
- > Parcialidade
- > Plágio

Síntese do Artigo





Automatização de Processos

Formatação ou Limpeza de Dados

Definir Estruturas de Dados

Infografias

Material de Formação

Identificação de Fontes

Criar Dados Sintéticos

Análise de Processos Analíticos

Imagem do <u>Unsplash</u>

Conformidade, Regulamentação e Medidas Práticas

Garantir que as operações de dados são legais, imparciais e éticas





Aplicabilidades do ChatGPT

Modelação da Linguagem

Gera texto preciso e coerente

Reconhecimento

Reconhece pessoas, locais e organizações no texto

Resposta a **Questões**

Responde a perguntas com base num contexto



Classificação de Texto

Identifica o tema e o sentimento do texto

Tradução Automática

Tradução de Texto

Geração de Texto

Cria texto criativo e informativo

Síntese do Artigo





Complexidade

Dificuldade em entender perguntas complexas

Ambiguidade

Diferentes interpretações, diferentes respostas

Domínios

Não é preciso em domínios específicos

Dados de Treino

A qualidade dos dados de treino influencia a precisão

Pergunta

A natureza da pergunta influencia a precisão

Viés nos Dados

Pode levar a respostas imprecisas





Crítica ao Artigo

Comparação com os **Modelos Atuais**

01



03

Recomendações Mais Específicas

02

Potenciais Impactos
Sociais e Éticos

Crítica ao Artigo





LANGUAGE MODEL SIZES TO MAR/2023

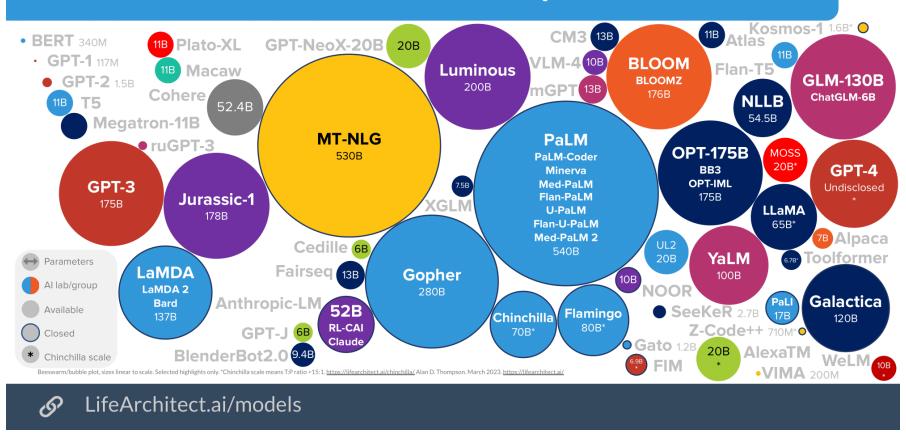


Figura 3 | Gráfico comparativo dos principais modelos linguísticos de IA 2018-2023.

Fonte: Life Architect (2023)





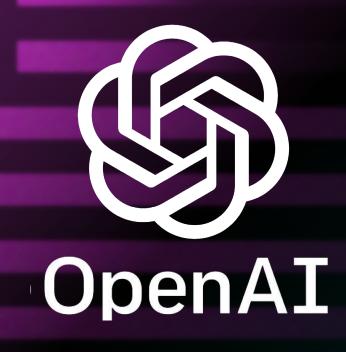


66

De acordo com o critério de originalidade da UE, as obras geradas por IA podem não ser elegíveis, uma vez que carecem das escolhas criativas e da expressão pessoal de um autor humano.

Lucchi, N. (2023). ChatGPT: A Case Study on Copyright Challenges for Generative Artificial Intelligence Systems.





GPT-4 is OpenAl's most advanced system, producing safer and more useful responses

GPT-4 can solve difficult problems with greater accuracy, thanks to its broader general knowledge and problem solving abilities.

GPT-4 surpasses ChatGPT in its advanced reasoning capabilities.

GPT-4 outperforms ChatGPT by scoring in higher approximate percentiles among test-takers.

OpenAl

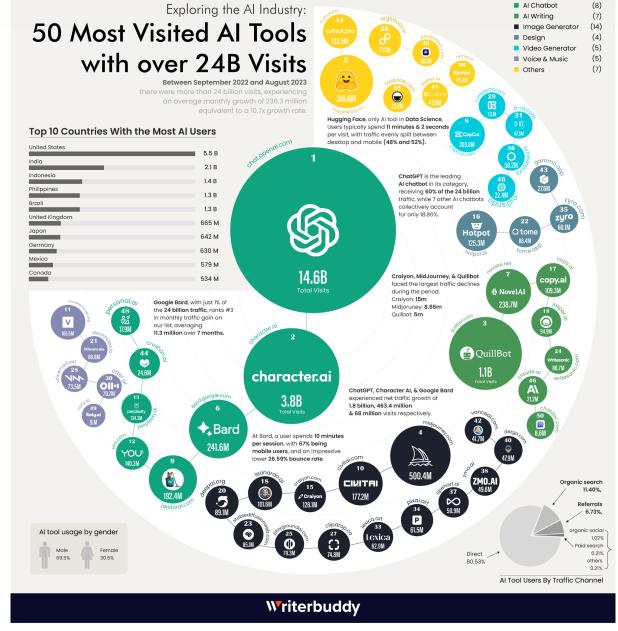


Figura 4 | Infografia das 50 plataformas de IA mais visitadas entre *Set de 2022* a *Aug de 2023*.





Figura 5 | Ilustração criada por IA alusiva à campanha de Natal da Coca-Cola. Fonte: <u>Create Real Magic © Cola-Cola</u> (2023)

Fonte: X (Coca-Cola India, 2023)

Google DeepMind About Technologies Impact Discover

Welcome to the Gemini era

The Gemini era

Capabilities

Hands-on

Safety

Bard

Build with Gemini