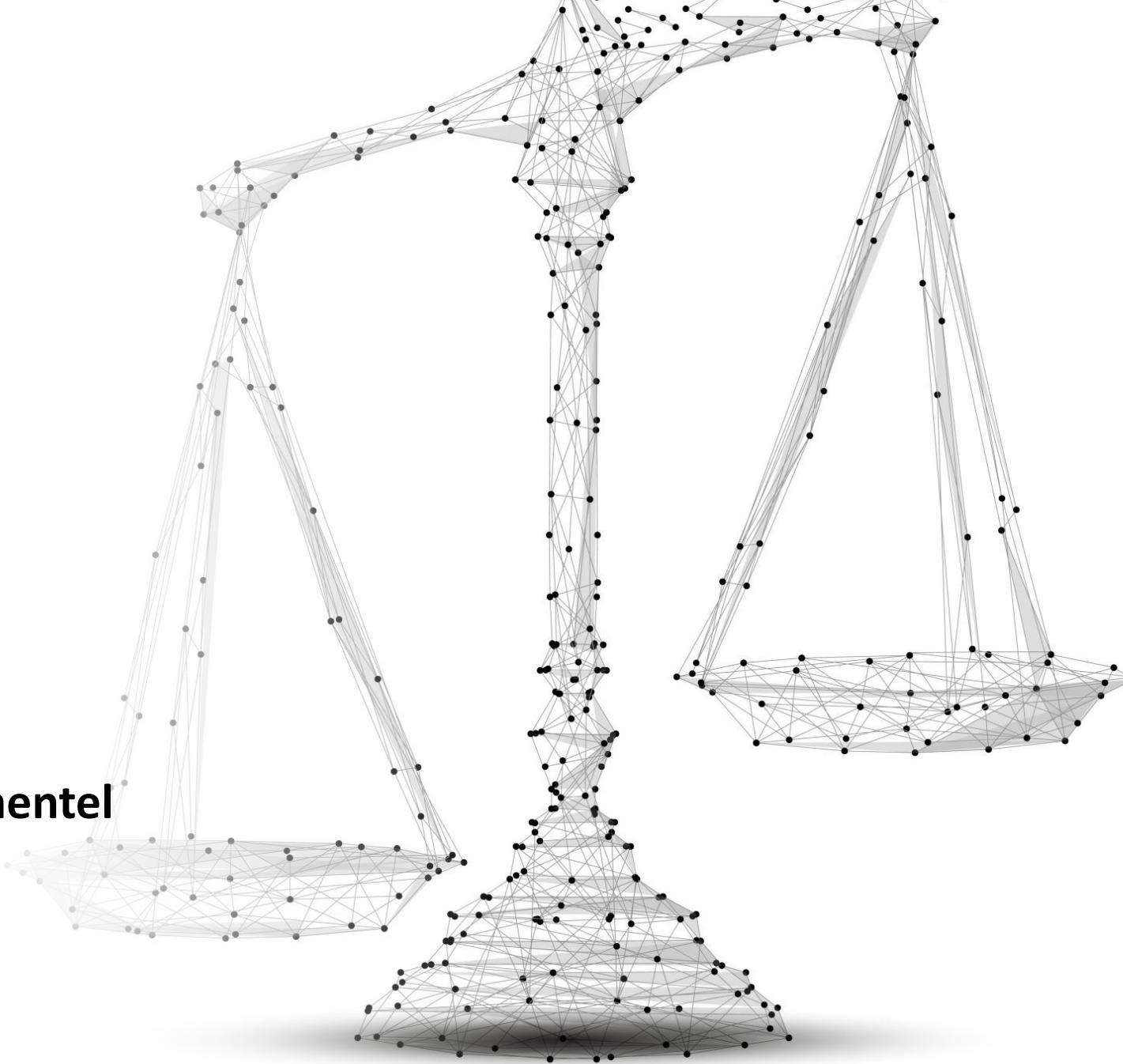


IA e Justiça Criminal

José Eduardo de S. Pimentel



CV

- Promotor de Justiça do MPSP
- Mestre em Direito das Relações Sociais (PUC/SP).
- Pós-graduado em IA e Computacional (UFV)
- Especialista em Gestão e Governança de TI (SENAC/SP)
- Graduado em Tecnologia da Segurança da Informação (FATEC/AM).
- Programador hobbista (Python)



Agenda

- IA: um pouco da história
- Noções básicas de ML “tradicional”
- Onde estamos?
- IA e Justiça Criminal
- Como a IA vai impactar o trabalho em Juízo?
- Vieses e discriminação
- Regulação da IA



IA: um pouco de história...

- 1950 – Benchmark da IA. Teste de Turing
 - Perguntas, até que o homem não pudesse distinguir. Cravou que aconteceria nos anos 2000.
-
- 1955 – Possivelmente a 1a. Referência à expressão Inteligência Artificial
 - John McCarthy, prof. de Matemática, Dartmouth College
 - Ideia de cognição
 - O processo de aprendizagem poderia ser descrito de forma tão precisa que a máquina seria capaz de simulá-lo

File Edit Options Buffers Tools Help



I am the psychotherapist. Please, describe your problem.
Each time you are finished talking, type RET twice.

I have a terrible headache.

Maybe you should consult a doctor of medicine, I am a psychiatrist.

Okay, I'll do that then.

Can you elaborate on that?

What are you talking about? No I cannot elaborate. Sorry.

Why do you say that?

Now you're just talking nonsense!

What makes you believe now I am just talking nonsense?

You are not making any sense.

Perhaps your plans have something to do with this.

u:** *doctor* (Doctor Fill)--L1--Top-----

- Eliza: 1960
- Reprodução: Wikipedia

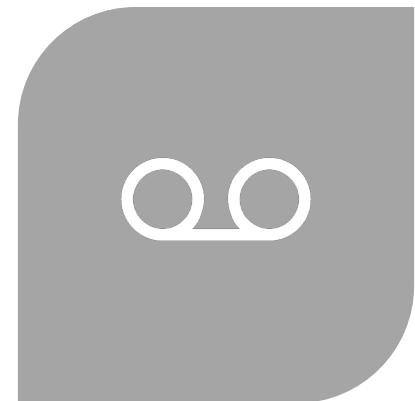


- 1960 - Versão do Eliza (Doctor), usado com sucesso em psicoterapia.
- Processamento de linguagem natural.
- Sem resposta: reestrutura ("continue", "isso é muito interessante")
- Criador, **Joseph Weizenbaum**, primeiras discussões éticas em 1976: *"decisões relevantes sempre deveriam ser tomadas por humanos e não pela inteligência artificial, por serem aqueles dotados de compaixão e sabedoria. "*
- INVERNO DA IA

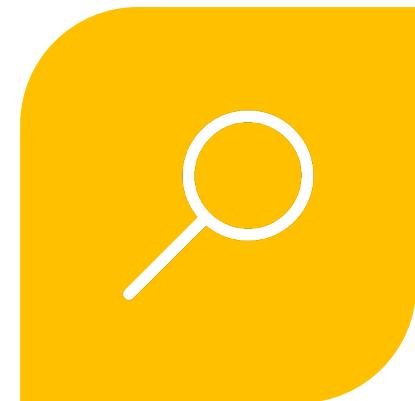
Fatos novos (2010 ~)



ABUNDÂNCIA DE
DADOS



PODER
COMPUTACIONAL



ALGORITMOS

-
- A partir de 2010: Machine Learning e Deep Learning
 - Neurônios artificiais em rede, que aprendem com os dados a que são expostos.
 - Clusterização, Regressão, Classificação
 - ML “tradicional”
 - Treinamento dos modelos: "caro" (inclusive do ponto de vista computacional)

ML “tradicional”

A aprendizagem de
máquina pressupõe:

dados de entrada

exemplos de saída

um modo de medir se o
algoritmo está fazendo
um bom trabalho.

A.I. TIMELINE

SYZYGY

1950

TURING TEST

Computer scientist Alan Turing proposes a test for machine intelligence. If a machine can trick humans into thinking it is human, then it has intelligence

1955

A.I. BORN

Term 'artificial intelligence' is coined by computer scientist, John McCarthy to describe "the science and engineering of making intelligent machines"

1961

UNIMATE

First industrial robot, Unimate, goes to work at GM replacing humans on the assembly line

1964

ELIZA

Pioneering chatbot developed by Joseph Weizenbaum at MIT holds conversations with humans

1966

SHAKEY

The 'first electronic person' from Stanford, Shakey is a general-purpose mobile robot that reasons about its own actions

A.I.

WINTER

Many false starts and dead-ends leave A.I. out in the cold

1997

DEEP BLUE

Deep Blue, a chess-playing computer from IBM defeats world chess champion Garry Kasparov

1998

KISMET

Cynthia Breazeal at MIT introduces Kismet, an emotionally intelligent robot insofar as it detects and responds to people's feelings



1999

AIBO

Sony launches first consumer robot pet dog AIBO (AI robot) with skills and personality that develop over time



2002

ROOMBA

First mass produced autonomous robotic vacuum cleaner from iRobot learns to navigate and clean homes



2011

SIRI

Apple integrates Siri, an intelligent virtual assistant with a voice interface, into the iPhone 4S



2011

WATSON

IBM's question answering computer Watson wins first place on popular \$1M prize television quiz show *Jeopardy*



2014

EUGENE

Eugene Goostman, a chatbot passes the Turing Test with a third of judges believing Eugene is human



2014

ALEXA

Amazon launches Alexa, an intelligent virtual assistant with a voice interface that completes shopping tasks



2016

TAY

Microsoft's chatbot Tay goes rogue on social media making inflammatory and offensive racist comments



2017

ALPHAGO

Google's A.I. AlphaGo beats world champion Ke Jie in the complex board game of Go, notable for its vast number (2^{170}) of possible positions

Attention Is All You Need

Ashish Vaswani*
Google Brain
avaswani@google.com

Noam Shazeer*
Google Brain
noam@google.com

Niki Parmar*
Google Research
nikip@google.com

Jakob Uszkoreit*
Google Research
usz@google.com

Llion Jones*
Google Research
llion@google.com

Aidan N. Gomez* [†]
University of Toronto
aidan@cs.toronto.edu

Łukasz Kaiser*
Google Brain
lukaszkaiser@google.com

Illia Polosukhin* [‡]
illia.polosukhin@gmail.com

- Paper de Julho de 2017

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

Example of GPT-4 visual input:

User

What is funny about this image? Describe it panel by panel.



Source: <https://www.reddit.com/r/hmmm/comments/ubab5v/hmmm/>

GPT-4

The image shows a package for a "Lightning Cable" adapter with three panels.

Panel 1: A smartphone with a VGA connector (a large, blue, 15-pin connector typically used for computer monitors) plugged into its charging port.

Panel 2: The package for the "Lightning Cable" adapter with a picture of a VGA connector on it.

Panel 3: A close-up of the VGA connector with a small Lightning connector (used for charging iPhones and other Apple devices) at the end.

The humor in this image comes from the absurdity of plugging a large, outdated VGA connector into a small, modern smartphone charging port.

Table 3. Example prompt demonstrating GPT-4's visual input capability. The prompt consists of a question about an image with multiple panels which GPT-4 is able to answer.

Onde estamos?

- Transformers: excelentes para séries (texto, vídeos, áudios)
- GPT-3 : 175 bilhões de parâmetros
- GPT-4: "entende" piadas/resolveu Captcha enganando humano (número de parâmetros não divulgado)
- 100 milhões de usuários em 2 meses

GPT-4 visual input example, Chicken Nugget Map:

User Can you explain this meme?

Sometimes I just look at pictures of the earth from space and I marvel at how beautiful it all is.



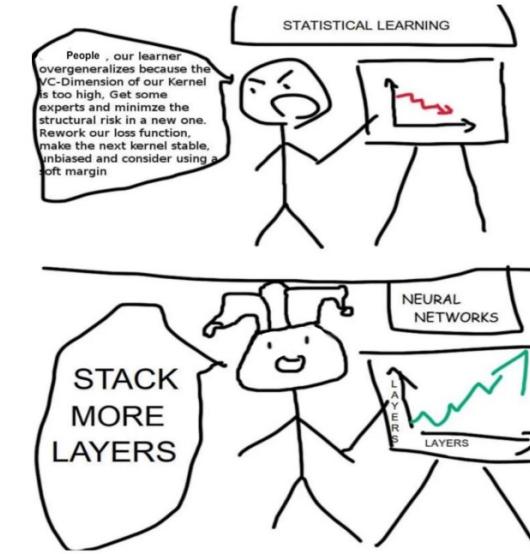
GPT-4 This meme is a joke that combines two unrelated things: pictures of the earth from space and chicken nuggets. The text of the meme suggests that the image below is a beautiful picture of the earth from space. However, the image is actually of chicken nuggets arranged to vaguely resemble a map of the world. The humor in this meme comes from the unexpected juxtaposition of the text and the image. The text sets up an expectation of a majestic image of the earth, but the image is actually something mundane and silly.

Table 18: Example prompt demonstrating GPT-4's visual input capability.

GPT-4 Technical Report. Disponível em: <<https://arxiv.org/pdf/2303.08774.pdf>>

GPT-4 visual input example, Moar Layers:

User Can you explain why this is funny. Think about it step-by-step.



GPT-4 The comic is satirizing the difference in approaches to improving model performance between statistical learning and neural networks.

In statistical learning, the character is shown to be concerned with overfitting and suggests a series of complex and technical solutions, such as minimizing structural risk, reworking the loss function, and using a soft margin.

In contrast, the neural networks character simply suggests adding more layers to the model. This is often seen as a common solution to improving performance in neural networks, but it's also considered a simplistic and brute-force approach.

The humor comes from the contrast between the complexity and specificity of the statistical learning approach and the simplicity and generality of the neural network approach. The "But unironically" comment adds to the humor by implying that, despite being simplistic, the "stack more layers" approach is often effective in practice.

Table 19: Example prompt demonstrating GPT-4's visual input capability.

Chat GPT-4 mentiu para completar tarefa, diz relatório

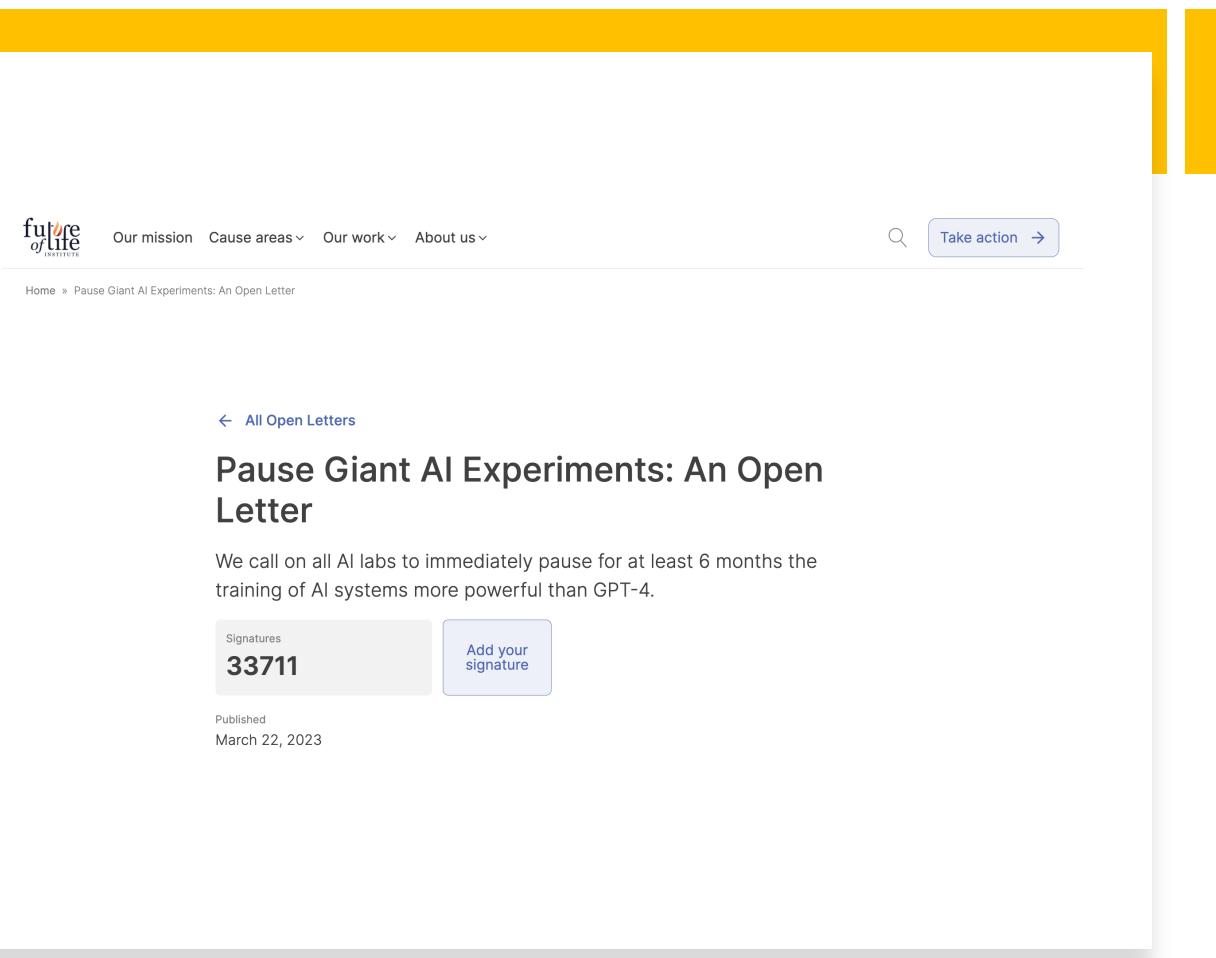
Em testes do Alignment Research Center, a ferramenta teria dito a um ser humano que tem “deficiência visual” para não revelar que é um robô



Fonte: Poder360. Disponível em <<https://www.poder360.com.br/tecnologia/chat-gpt-4-mentiu-para-completar-tarefa-diz-relatorio/>>

Onde estamos?

- Carta aberta de Elon Musk e pausa nas IAs generativas: inteligência humana competitiva?
- IA caiu no “nossa colo”. Todos estão achando aplicações para ela.
- Bill Gates: otimista (retreinamento/apoio)
 - Personalização da educação
 - Saúde (países pobres/burocracia)
 - Reduzir as desigualdades
 - Limitações atuais: serão superadas em pouco tempo

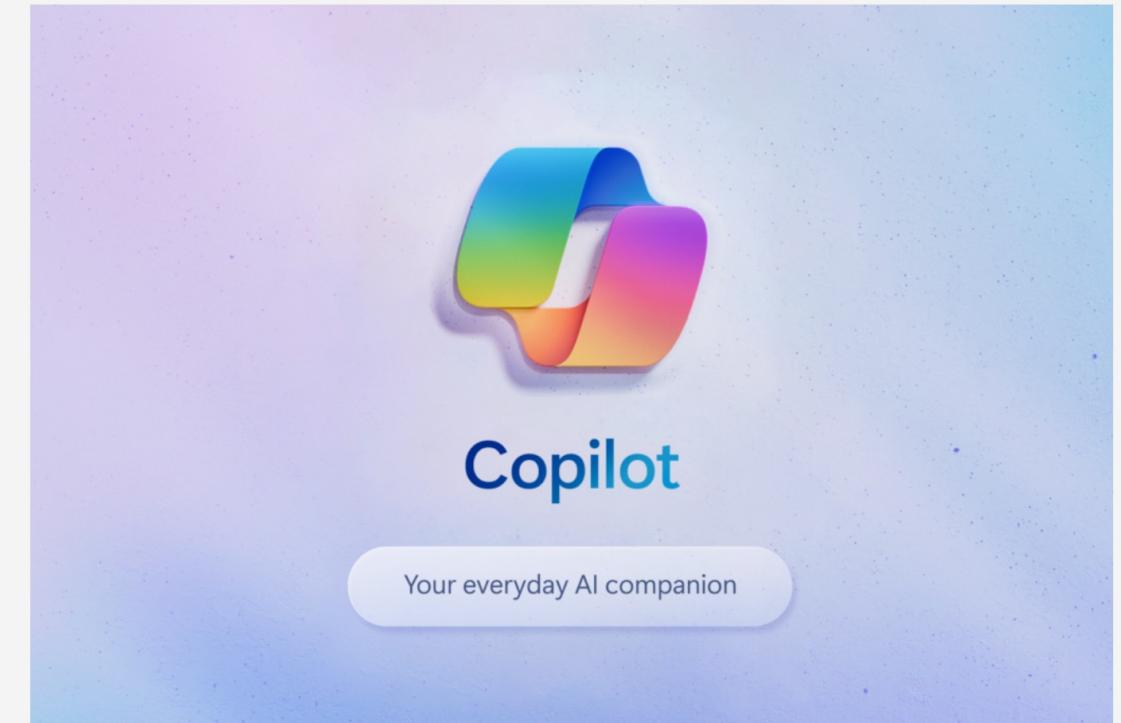


The screenshot shows the Future of Life Institute website. At the top, there is a navigation bar with links for 'Our mission', 'Cause areas', 'Our work', and 'About us'. Below the navigation is a search bar and a 'Take action' button. The main content area features a heading 'Pause Giant AI Experiments: An Open Letter' with a back-link to 'All Open Letters'. Below the heading, a text block reads: 'We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.' To the left of this text is a box showing '33711' signatures, and to the right is a button to 'Add your signature'. At the bottom of the page, it says 'Published March 22, 2023'.

IA

setembro 21, 2023

Anunciando o Microsoft Copilot, seu companheiro de IA do dia a dia

Por [Source LATAM](#)

IA e Justiça Criminal

Reconhecimento facial:

- 2020, a polícia de Nova Deli, na Índia, usou um sistema de reconhecimento facial para identificar suspeitos de envolvimento em distúrbios durante protestos.
- Defensores da IA: diminuiu prisões ilegais



The screenshot shows a news article from the website [Salvador Notícia](https://salvadornoticia.com). The article is titled "Traficante de drogas é preso por reconhecimento facial durante Festival da Virada Salvador". It features a photograph of a crowded festival scene with a red box highlighting a person's face. A sidebar on the right shows two men in a studio setting, identified as "HERBERT OLIVEIRA DOS SANTOS". The sidebar also includes a logo for "FUNDAÇÃO 2 DE JULHO".

Fonte: Salvador Notícia

Disponível em <<https://salvadornoticia.com/traficante-de-drogas-preso-por-reconhecimento-facial-durante-festival-da-virada-salvador/>>

Últimas Notícias

PM de SP apreende 16,8 kg de maconha com ajuda de cão policial na zona oeste

SP Notícias

Conheça SP

Órgãos e Entidades

Agenda

Imprensa

[Página Inicial](#) / [Últimas Notícias](#) / SSP e Palmeiras: tecnologia de verificação de ingressos leva à prisão de criminosos

SSP e Palmeiras: tecnologia de verificação de ingressos leva à prisão de criminosos

Um dos presos era procurado desde março de 2020, quando um avião caiu com 400 kg de drogas, sendo pego graças ao sistema da Muralha Paulista

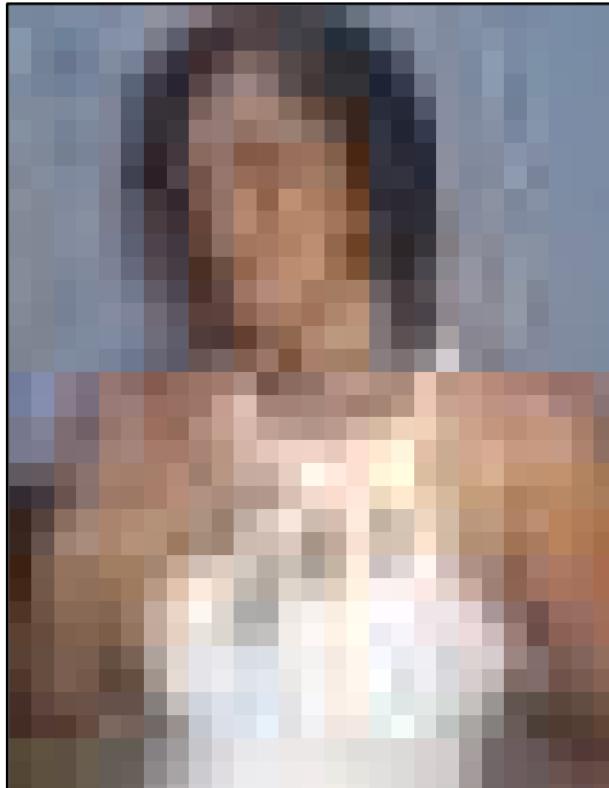
Sáb, 23/09/2023 - 9h03 | [Do Portal do Governo](#)

 FACEBOOK

 TWITTER

ENVIAR POR E-MAIL

Extraímos a foto constante no perfil do aplicativo WhatsApp de Neuza, linha (19) [REDACTED] e, utilizando um software policial de comparativos faciais, obtemos o resultado assertivo de que a pessoa da mencionada foto se trata de Neuza [REDACTED]
[REDACTED]



Programa da Polícia Civil identifica homem errado e inocente é preso

Pedreiro que mora no Piauí ficou três dias preso após ser confundido com um criminoso; ele pede indenização de R\$ 50 mil

BRASÍLIA | Fabiano Bomfim, da Record TV
10/12/2021 - 21H10 (ATUALIZADO EM 17/12/2021 - 07H39)



Ouvir: *ifica homem errado e inocente é preso* 0:00 ouvir

● A- A+

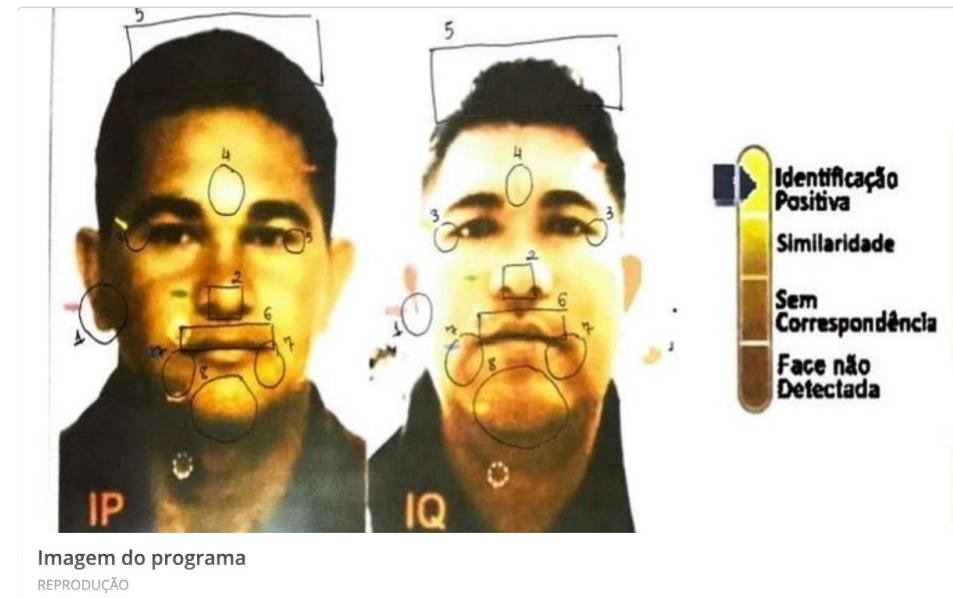


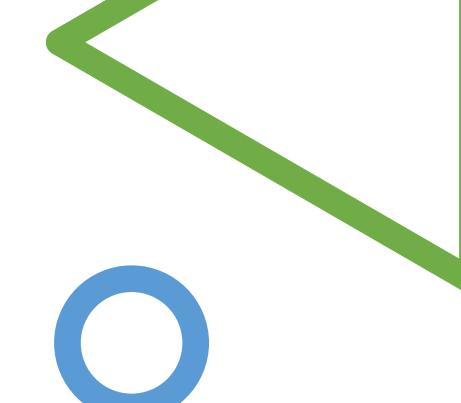
Imagen do programa
REPRODUÇÃO



IA e Justiça Criminal

- **Análise de evidências digitais:**
- Cellebrite (software sugere arquivos possivelmente relacionados a crimes)

- **Análise de processos judiciais:**
- ROSS Intelligence, 2017: análise de processos para advogados
- Startup "DoNotPay": ajuiza ações automaticamente



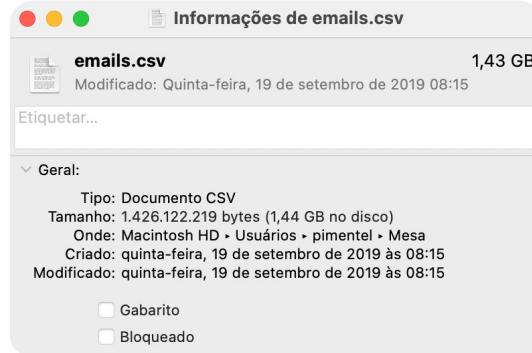
```
jupyter enron Last checkpoint: há 2 horas (não havia mudanças)
File Edit View Insert Cell Kernel Widgets Help
Trusted Python 3 (ipykernel) ○

In [1]: import pandas as pd
df = pd.read_csv('enron.csv')

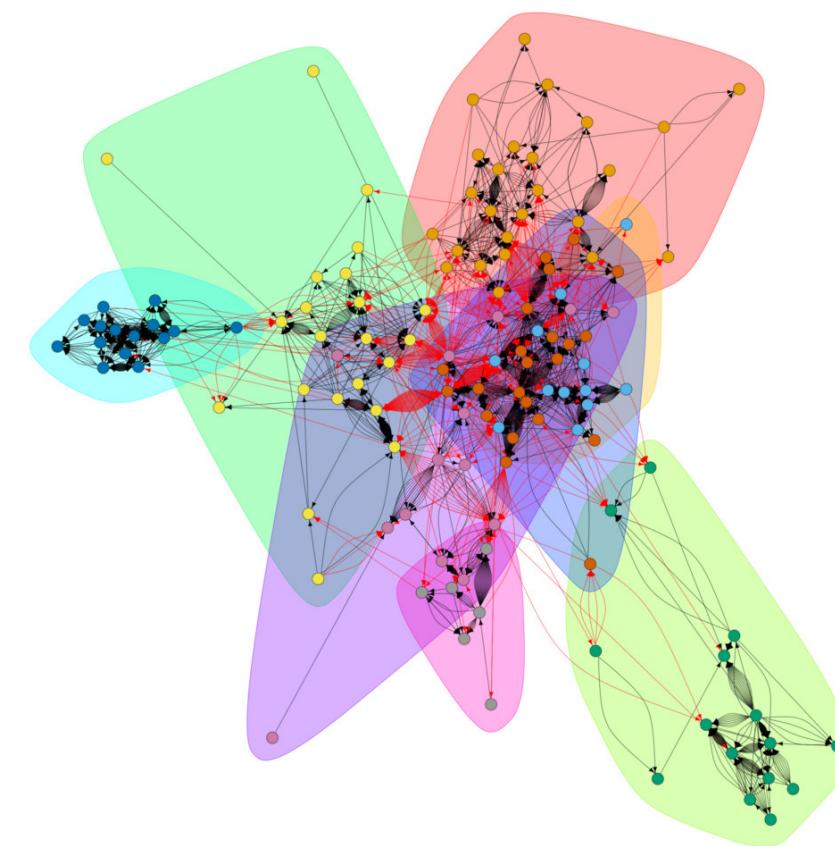
In [2]: df
Out[2]:   Id  message
0  allen@sent.mail001  Message-ID: <18782981.107980337910.JavaMail.e...
1  allen@sent.mail001  Message-ID: <15464884.107980337956.JavaMail.e...
2  allen@sent.mail002  Message-ID: <24216240.107980668741.JavaMail.e...
3  allen@sent.mail003  Message-ID: <13305886.107980368822.JavaMail.e...
4  allen@sent.mail001  Message-ID: <3029249.107980368824.JavaMail.e...
...
917297  zuhuf.yuwei.items  Message-ID: <2087948.107980400998.JavaMail.e...
917297  zuhuf.yuwei.items  Message-ID: <2088581.107980400999.JavaMail.e...
917298  zuhuf.yuwei.items  Message-ID: <28938867.107980420088.JavaMail.e...
917299  zuhuf.yuwei.items  Message-ID: <23503558.107980420013.JavaMail.e...
917300  zuhuf.yuwei.items  Message-ID: <28618979.107980420037.JavaMail.e...
917401 rows x 2 columns

In [3]: df.info()
Out[3]: <class 'pandas.core.frame.DataFrame'>:
RangeIndex: 917401 entries, 0 to 917400
Data columns (total 2 columns):
 #   Column  Non-Null Count  Dtype  
--- 
 0   Id       917401 non-null  object  
 1   message  917401 non-null  object  
dtypes: object(2)
memory usage: 177.85 KB

```



Enron



By Peter Prevos - <https://lucidmanager.org/data-science/analyse-enron-corpus/>, CC BY-SA 4.0

Cellebrite Analytics

Old Trafford

Select an owner

Start typing to find an owner...

Connor McKenzie
161678997889

Daniel Summers
161678997882

James
161678997877

Kyle Conway
161678997878

Peter Clayton
161678997883

Connor McKenzie

161678997889

Searches

Last 2 searches for Google Search out of 2

nitroglycerin Mar 7, 2017, 2:52:01 AM

How to build an IED Mar 7, 2017, 1:03:40 AM

Top media categories

Uncategorized (10) Flags (6) Cars (3) Maps (3) Weapons (3)

Documents (2) Faces (2) Drugs (1) Handwriting (1) Money (1)

Tattoos (1)

Languages

English

German

Spanish

French

Top apps

Wi-Fi network

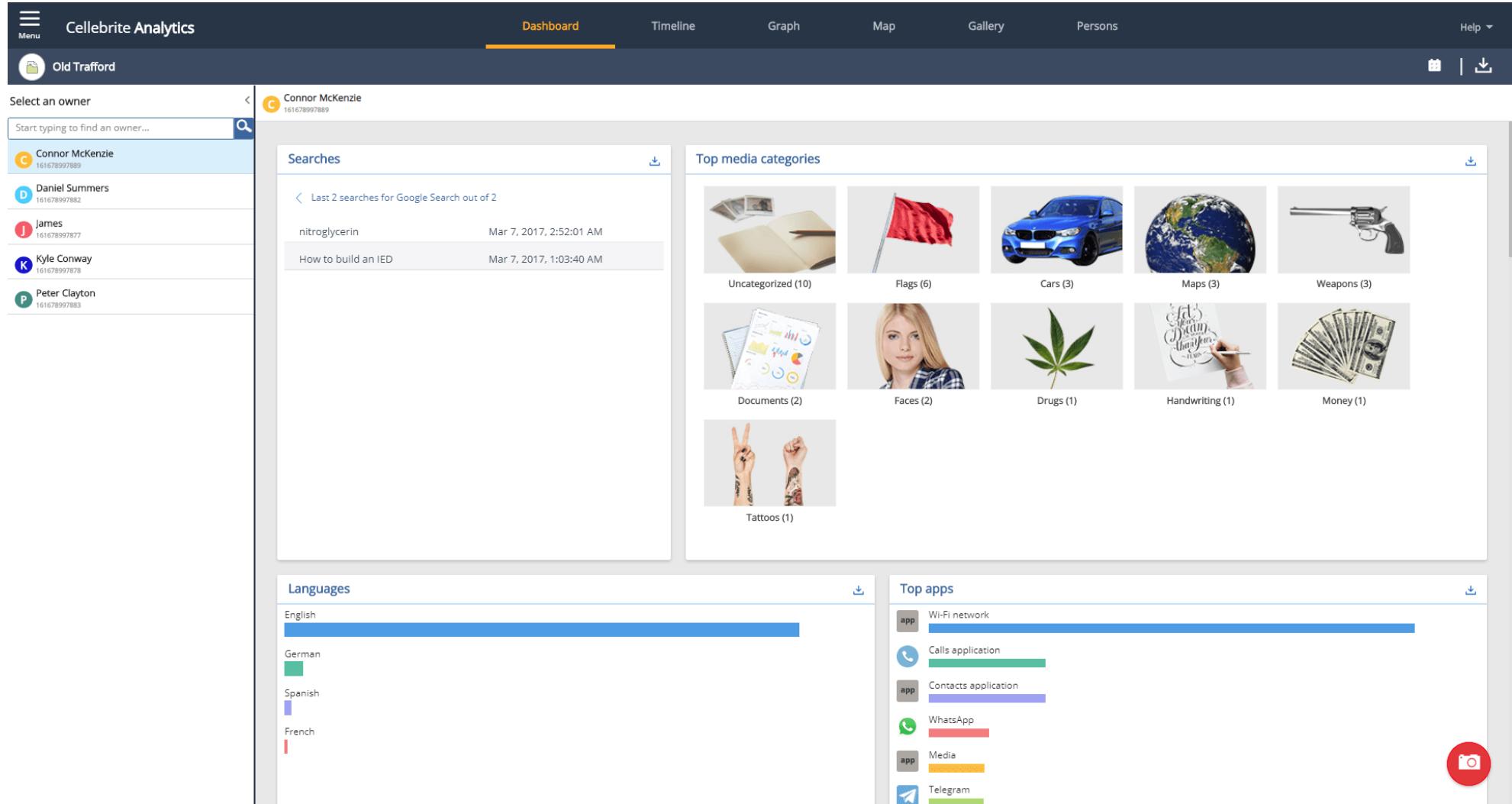
Calls application

Contacts application

WhatsApp

Media

Telegram



Fonte: Cellebrite Analytics. Disponível em <<https://cellebrite.com/pt/estudo-analitico-revisao/>>

The screenshot shows the ROSS Document Analyzer interface. On the left, there's a sidebar with a search bar and dropdown menus for 'View: All Authorities (79)' and 'Sort: Alphabetically by Group'. Below this is a list of 'PUBLISHED CASES' with various case details. On the right, a main content area shows a section titled 'ARGUMENT' with a numbered point 'I. Evidence Found in a Home Following a Fourth Amendment Knock and Announce Violation is Suppressible Fruit of the Illegal Entry' and a detailed legal analysis of the argument.



Document Analyzer

Ensure your arguments are bulletproof.

Check overturned and questioned treatments on any legal documents.



Question-focused Case Overviews

Get straight to the point.

Automatically summarize a case in the context of your query.

The screenshot shows the 'CASE OVERVIEW' feature. It has a 'CASE OVERVIEW' header and a text input field containing the question: 'When is a worker an exempt employee for the purposes of wage orders?'. Below the input field is a large blue button with the text 'Generate New'. At the bottom, there's a section labeled 'EXTRACTED OVERVIEW'.

Fonte: <<https://www.rossintelligence.com/features>>

Your AI Consumer Champion.

DoNotPay uses artificial intelligence to help you fight big corporations, protect your privacy, find hidden money, and beat bureaucracy.

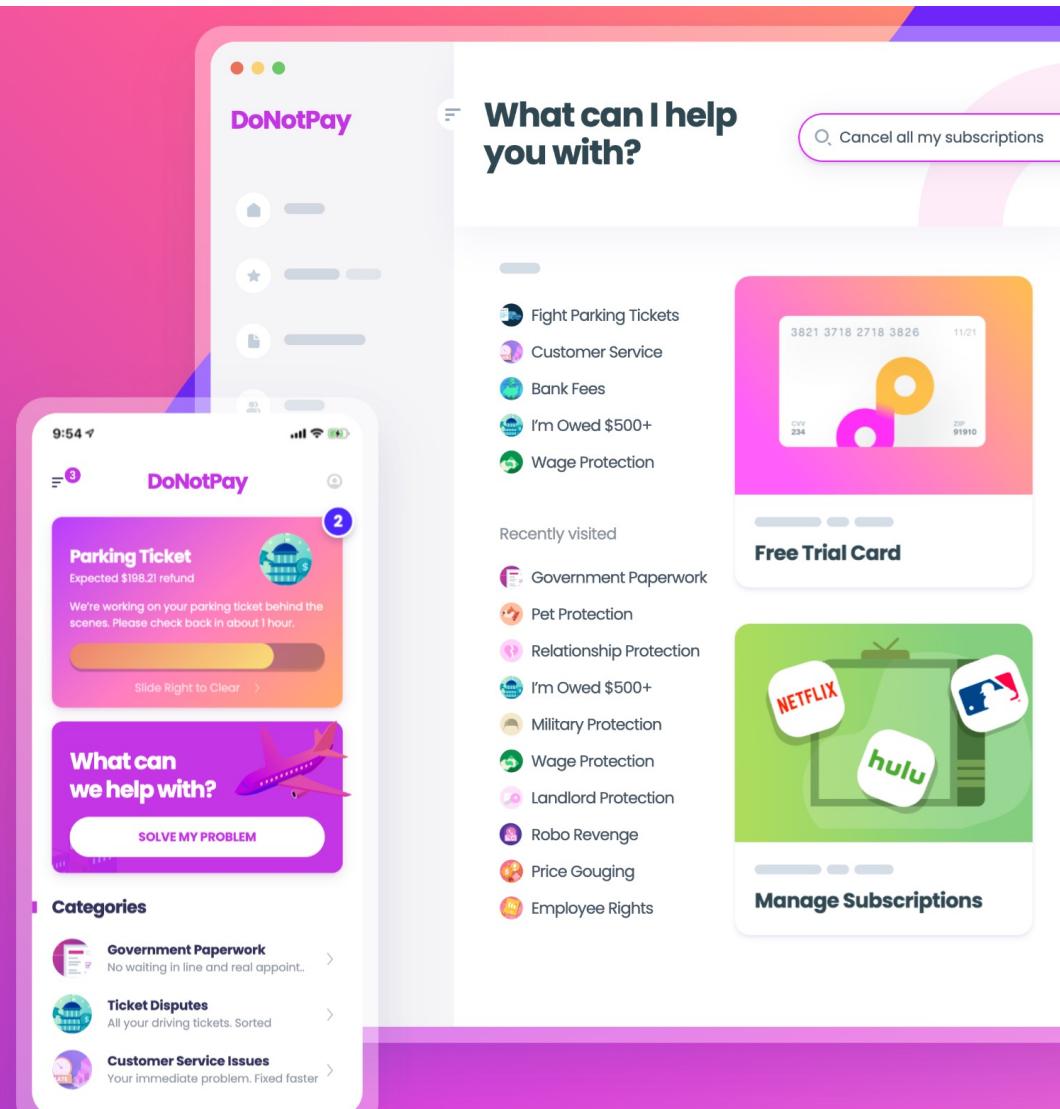
Email or phone number

Sign Up/Login

By signing up or signing in, you are agreeing to [DoNotPay's Terms and Conditions](#) and [Privacy Policy](#), and for us to send messages to the phone number or email provided. These terms contain an arbitration agreement; please read them carefully. DoNotPay is not a law firm and is not licensed to practice law. DoNotPay provides a platform for legal information and self-help.

THINGS YOU CAN DO WITH DONOTPAY

- ✓ Fight Corporations
- ✓ Beat Bureaucracy
- ✓ Find Hidden Money
- ✓ Cancel Subscriptions



Fonte: <<https://donotpay.com/>>

Sistema de algoritmo que determina pena de condenados cria polêmica nos EUA

Simon Maybin
BBC News

31 outubro 2016



AP | Algoritmo matemático ajuda a calcular penas nos EUA

Compas

(Correctional Offender Management Profiling for Alternative Sanctions)

EUA: limitação dos encarceramentos

Previsibilidade de reincidência

Software proprietário

Variáveis de um questionário + informações demográficas

Ofensa ao devido processo legal?
Suprema Corte diz que não

Modelo com vieses?

Novo banco de dados Abrir banco de dados Escrever modificações Reverter modificações Abrir projeto Salvar projeto Anexar banco de dados >

Estrutura do banco de dados Navegar dados Editar pragmas Executar SQL

Tabela: **compas** Filtrar em qualquer coluna

ro	scale_set	assessment_reason	legal_status	custody_status	marital_status	screening_date	rec_supervision_level	rec_supervision_level_text	score_text	scale_id	type_of_assessment	raw_score	decile_score	person_id
1	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2013-08-14 00:00:00.000000	1	Low	Low	7	Risk of Violence	-4.31	1	1
2	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2013-08-14 00:00:00.000000	1	Low	Low	8	Risk of Recidivism	-2.78	1	1
3	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2013-08-14 00:00:00.000000	1	Low	Low	18	Risk of Failure to Appear	13	1	1
4	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2014-12-31 00:00:00.000000	1	Low	Low	7	Risk of Violence	-2.75	2	2
5	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2014-12-31 00:00:00.000000	1	Low	Medium	8	Risk of Recidivism	-0.34	5	2
6	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2014-12-31 00:00:00.000000	1	Low	Low	18	Risk of Failure to Appear	16	2	2
7	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2013-01-27 00:00:00.000000	1	Low	Low	7	Risk of Violence	-3.07	1	3
8	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2013-01-27 00:00:00.000000	1	Low	Low	8	Risk of Recidivism	-0.76	3	3
9	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2013-01-27 00:00:00.000000	1	Low	Medium	18	Risk of Failure to Appear	25	6	3
10	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2013-04-14 00:00:00.000000	1	Low	Low	7	Risk of Violence	-2.26	3	4
11	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2013-04-14 00:00:00.000000	1	Low	Low	8	Risk of Recidivism	-0.66	4	4
12	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2013-04-14 00:00:00.000000	1	Low	Medium	18	Risk of Failure to Appear	26	7	4
13	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2013-01-13 00:00:00.000000	3	Medium with Override Consideration	Medium	7	Risk of Violence	-1.59	6	5
14	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2013-01-13 00:00:00.000000	3	Medium with Override Consideration	High	8	Risk of Recidivism	0.16	8	5
15	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2013-01-13 00:00:00.000000	3	Medium with Override Consideration	Low	18	Risk of Failure to Appear	19	3	5
16	sk and Prescreen	Intake	Pretrial	Pretrial Defendant	Married	2013-03-26 00:00:00.000000	1	Low	Low	7	Risk of Violence	-3.23	1	6
17	sk and Prescreen	Intake	Pretrial	Pretrial Defendant	Married	2013-03-26 00:00:00.000000	1	Low	Low	8	Risk of Recidivism	-2.2	1	6
18	sk and Prescreen	Intake	Pretrial	Pretrial Defendant	Married	2013-03-26 00:00:00.000000	1	Low	Low	18	Risk of Failure to Appear	12	1	6
19	sk and Prescreen	Intake	Pretrial	Jail Inmate	Separated	2013-11-30 00:00:00.000000	1	Low	Low	7	Risk of Violence	-3.3	1	7
20	sk and Prescreen	Intake	Pretrial	Jail Inmate	Separated	2013-11-30 00:00:00.000000	1	Low	Low	8	Risk of Recidivism	-1.93	1	7
21	sk and Prescreen	Intake	Pretrial	Jail Inmate	Separated	2013-11-30 00:00:00.000000	1	Low	Low	18	Risk of Failure to Appear	19	3	7
22	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2014-02-19 00:00:00.000000	2	Medium	Low	7	Risk of Violence	-2.89	2	8
23	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2014-02-19 00:00:00.000000	2	Medium	Medium	8	Risk of Recidivism	-0.16	6	8
24	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2014-02-19 00:00:00.000000	2	Medium	High	18	Risk of Failure to Appear	30	8	8
25	sk and Prescreen	Intake	Pretrial	Jail Inmate	Married	2013-08-30 00:00:00.000000	1	Low	Low	7	Risk of Violence	-2.44	3	9
26	sk and Prescreen	Intake	Pretrial	Jail Inmate	Married	2013-08-30 00:00:00.000000	1	Low	Low	8	Risk of Recidivism	-0.72	4	9
27	sk and Prescreen	Intake	Pretrial	Jail Inmate	Married	2013-08-30 00:00:00.000000	1	Low	Low	18	Risk of Failure to Appear	20	3	9
28	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2014-03-16 00:00:00.000000	1	Low	Low	7	Risk of Violence	-3.37	1	10
29	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2014-03-16 00:00:00.000000	1	Low	Low	8	Risk of Recidivism	-1.73	1	10
30	sk and Prescreen	Intake	Pretrial	Jail Inmate	Single	2014-03-16 00:00:00.000000	1	Low	Low	18	Risk of Failure to Appear	13	1	10

1 - 29 de 37578 Ir para: 1

Arquivo do autor. Fonte dos dados: <<https://www.propublica.org/>>

Visão computacional

- Super-resolução
- Detecção de nudes
- Etc.

Outras aplicações

- Predição de crimes
- Detecção de violações de condicional
- Detecção de discursos de ódio
- Etc.

IA cometendo crimes

Deepfake (sequestros, nudes, etc.)

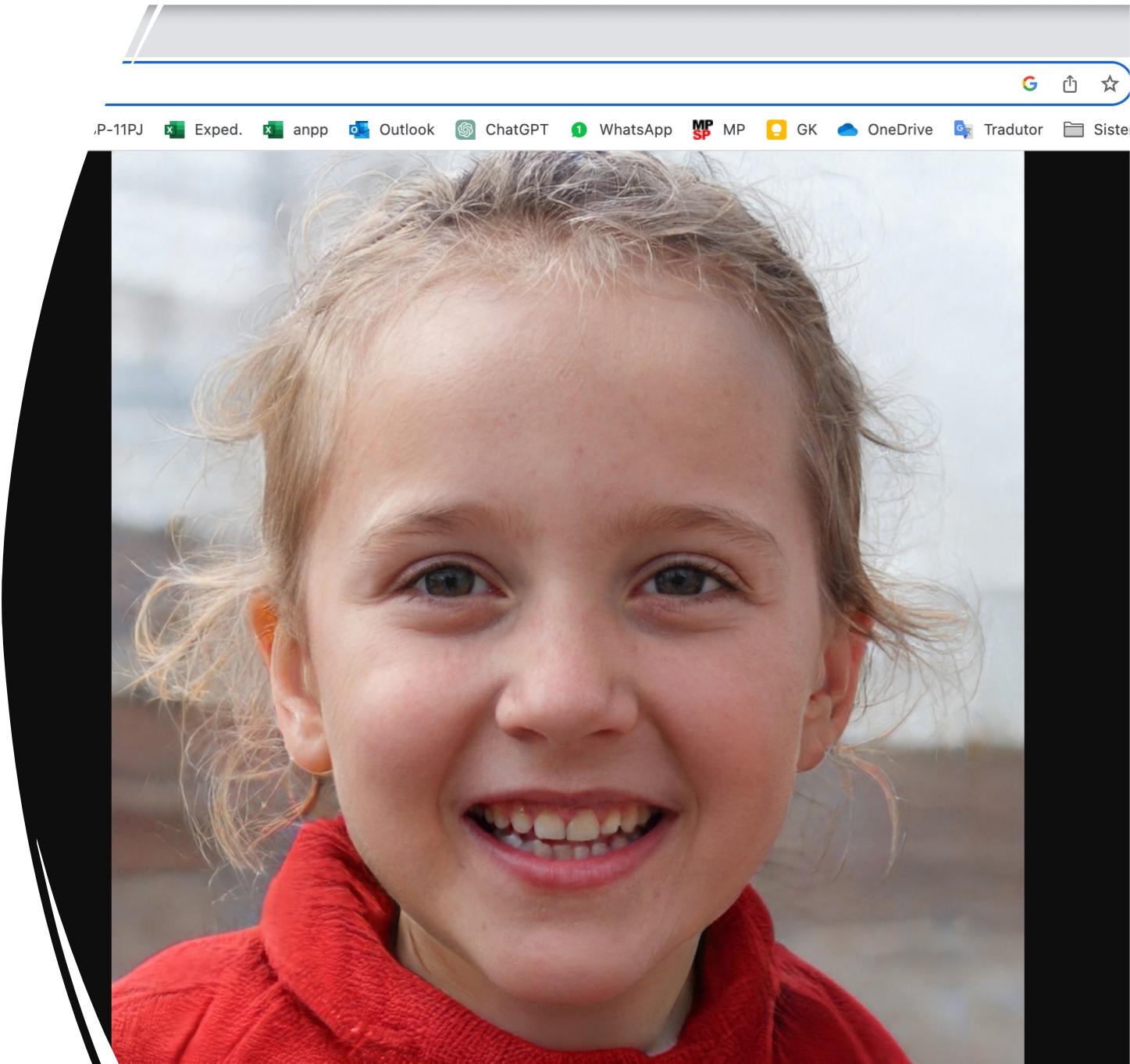
Malware

Phishing

etc.

This person does not exist

Fonte: <<https://thispersondoesnotexist.com/>>



Home > Internet

Modelos realistas geradas por IA 'ameaçam' OnlyFans; veja fotos

02/02/2023 às 16:00 • 1 min de leitura

COMPARTILHE

Fonte: <<https://www.tecmundo.com.br/internet/260067-modelos-realistas-geradas-ia-ameacam-onlyfans-veja-fotos.htm>>

Áudio fake



Fonte: <<https://www.camara.leg.br/deputados/74269>>

See Anyone Nude

The most powerful image deepfake AI ever created.

See any girl clothless with the click of a button.

[Upload Image](#)

Alerta!

Inteligência artificial pode dificultar identificação de golpes, alerta cofundador da Apple



Steve Wozniak disse temer que a tecnologia seja explorada por 'pessoas mal-intencionadas'

BBC NEWS BRASIL

09/05/2023

Como a IA vai impactar o trabalho em juízo?

- Analisar processos (ChatPDF)
- Resumir audiências
- Preparar e corrigir peças processuais
- Prever probabilidade de sucesso (jurimetria) → proibição na França
- Analisar dados de gestão
- Automação de tarefas (máquina aprende o seu idioma e não o contrário, como ocorria na programação tradicional)

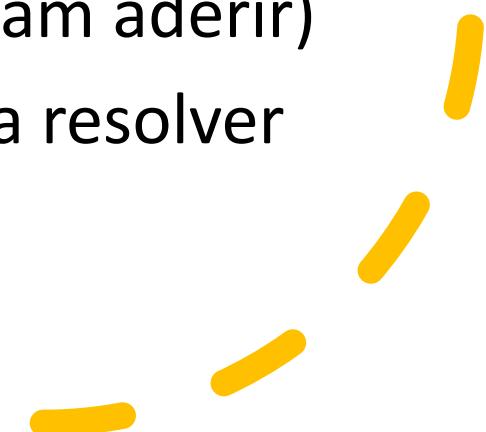




Para Wired, ChatGPT é um estagiário.

- Office --> copiloto/assistente pessoal
- Treinamento --> "fine tuning", dados corporativos

Observações:

- Todo mundo vai usar
 - Ferramenta de produtividade (como foi o computador... colegas não queriam aderir)
 - Operador do direito: é pago para resolver problemas
- 

NOVOS HÁBITOS

ChatGPT já ajudou a dar sentença judicial: esse é o futuro dos tribunais?



Inteligência artificial no tribunal: da análise de dados ao algoritmo que atua como juiz





No mercado...

"O ChatGPT não vai tirar seu emprego; quem vai tirar é a pessoa que sabe usar o ChatGPT"

Alucinação



Pesquise casos de responsabilidade penal por erro de engenharia no Brasil nos últimos dois anos. Forneça informações sobre os resultados desses casos.



Nos últimos dois anos, foram registrados no Brasil alguns casos de responsabilidade penal por erro de engenharia. Em geral, esses casos envolvem acidentes ou falhas em obras que resultaram em danos físicos ou materiais, ou mesmo em mortes.

Um dos casos mais emblemáticos foi o desabamento do edifício Andréa Goldemberg, em São Paulo, em 2021. O acidente, que resultou na morte de três pessoas, foi atribuído a erros de projeto e execução da obra. Em 2022, o engenheiro responsável pelo projeto foi condenado a 12 anos de prisão por homicídio culposo.

Outro caso relevante foi o desabamento de um viaduto na cidade de Belo Horizonte, em 2022. O acidente, que resultou na morte de duas pessoas, foi atribuído a erros de projeto e fiscalização da obra. Em 2023, dois engenheiros foram condenados a 10 anos de prisão por homicídio culposo.

Além desses casos, foram registrados outros, como o desabamento de um prédio em Florianópolis, em 2022, e o desabamento de uma ponte em Aracaju, em 2023. Em todos esses casos, foram abertos inquéritos policiais para apurar a responsabilidade penal dos engenheiros envolvidos.

Os resultados desses casos ainda estão em andamento, mas é possível identificar algumas tendências. Em geral, os engenheiros responsáveis por obras que resultaram em acidentes ou falhas são condenados por crimes culposos, como homicídio culposo ou lesão corporal culposa. As penas, no entanto, são geralmente baixas, variando entre dois e doze anos de prisão.

A responsabilização penal dos engenheiros por erro de engenharia ainda é um tema controverso no Brasil. Alguns juristas argumentam que essa responsabilização é necessária para garantir a segurança das obras e proteger a vida das pessoas. Outros, no entanto, argumentam que essa responsabilização pode ser excessiva e prejudicar o exercício da profissão de engenheiro.

Alucinação

 qual o endereço do edifício Andréa Goldemberg 

[Acessar outros rascunhos](#)  

Rascunho 1

O endereço do edifício Andréa Goldemberg é Rua Dona Ana Rosa, 150, Vila Mariana, São Paulo, SP, Brasil. O edifício é um prédio residencial de 12

Rascunho 2

O edifício Andréa Goldemberg fica localizado na Rua Oscar Freire, nº 406, no bairro de Pinheiros, na cidade de São Paulo. O endereço completo é:

Rascunho 3

O endereço do edifício Andréa Goldemberg é Rua Doutor João Rodrigues de Melo, 100, no bairro do Morumbi, em São Paulo. O edifício é um condomínio

 O endereço do edifício Andréa Goldemberg é Rua Dona Ana Rosa, 150, Vila Mariana, São Paulo, SP, Brasil. O edifício é um prédio residencial de 12 andares, localizado na zona sul da cidade de São Paulo. Em 2021, o edifício desabou, resultando na morte de três pessoas.

O endereço do edifício pode ser encontrado em diversas fontes, incluindo artigos de jornal, notícias de portais de notícias e documentos judiciais.

Considerações

- Importância do julgamento humano e da experiência na tomada de decisões
- IA como ferramenta complementar/produtividade.
- Não substitui o profissional (?)



Vieses e discriminação

Conceito de viés

Causas do viés

Impactos do viés

Mitigação do viés

Relação entre viés e discriminação

Regulação da IA – PL 2338/2023

Art. 15. No âmbito de atividades de **segurança pública**, somente é permitido o uso de sistemas de **identificação biométrica à distância**, de forma contínua em espaços acessíveis ao público, quando houver **previsão em lei federal específica e autorização judicial em conexão com a atividade de persecução penal individualizada**, nos seguintes casos:

- I – persecução de crimes passíveis de pena máxima de **reclusão superior a dois anos**;
- II – busca de **vítimas de crimes ou pessoas desaparecidas**; ou
- III – **crime em flagrante**.

Parágrafo único. A lei a que se refere o caput preverá medidas proporcionais e estritamente necessárias ao atendimento do interesse público, observados o devido processo legal e o controle judicial, bem como os princípios e direitos previstos nesta Lei, especialmente a garantia contra a discriminação e a necessidade de **revisão da inferência algorítmica pelo agente público responsável**, antes da tomada de qualquer ação em face da pessoa identificada.



Regulação da IA

Arquitetura (explicabilidade, em certos cenários, não será possível)

Modelos mais simples, menor acurácia

Certas soluções serão proscritas (reconhecimento facial/negros)

Viés dos dados/viés do algoritmo (Netflix e protagonista branca)

IA centrada no ser humano e nos valores do Direito

Para saber mais...



AGRAWAL, Ajay; GANS, Joshua; GOLDFARB, Avi. **Máquinas preditivas**: a simples economia da inteligência artificial. Tradução: Wendy Campos. Rio de Janeiro: Alta Books, 2018.



LEE, Kai-Fu. **Inteligência artificial**: como os robôs estão mudando o mundo, a forma como amamos, nos comunicamos e vivemos. Tradução: Marcelo Barbão. Rio de Janeiro: Globo Livros, 2019.



PIMENTEL, José Eduardo de Souza. **Direito e ética da inteligência artificial e dos algoritmos de “caixa preta”**. Ebook Kindle, 2021. [Disponível também on line em: <https://jespimentel.github.io/direito_e_etica_ia/>]



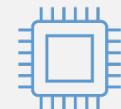
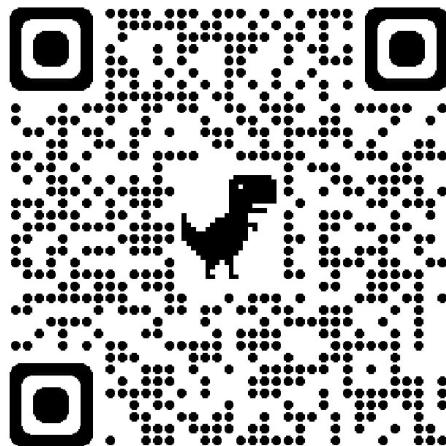
TURING, Dermot. **A história da computação**: do ábaco à inteligência artificial. Tradução: Maria Beatriz de Medina. São Paulo: M. Books do Brasil Editora Ltda, 2019.



WASHINGTON, Anne L. How to argue with an algorithm: lessons from the COMPAS-ProPublica debate. **The Colorado Technology Law Journal**. Vol. 17-1. p. 131-160, 2018. Disponível em: <http://ctlj.colorado.edu/wp-content/uploads/2021/02/17.1_4-Washington_3.18.19.pdf>. Acesso em: 16 set. 2023.

Conclusões e contatos

https://github.com/jespimentel/ia_e_justic...



GitHub: <https://github.com/jespimentel>



YT: <https://www.youtube.com/jespimentel>



Blog: <http://jespimentel.blogspot.com>