Challenges and Opportunities for Keeping AI Open

Rajiv Shah

@rajistics

raj@huggingface.co

my journey

my journey

my journey

SPSS

H2O.ai

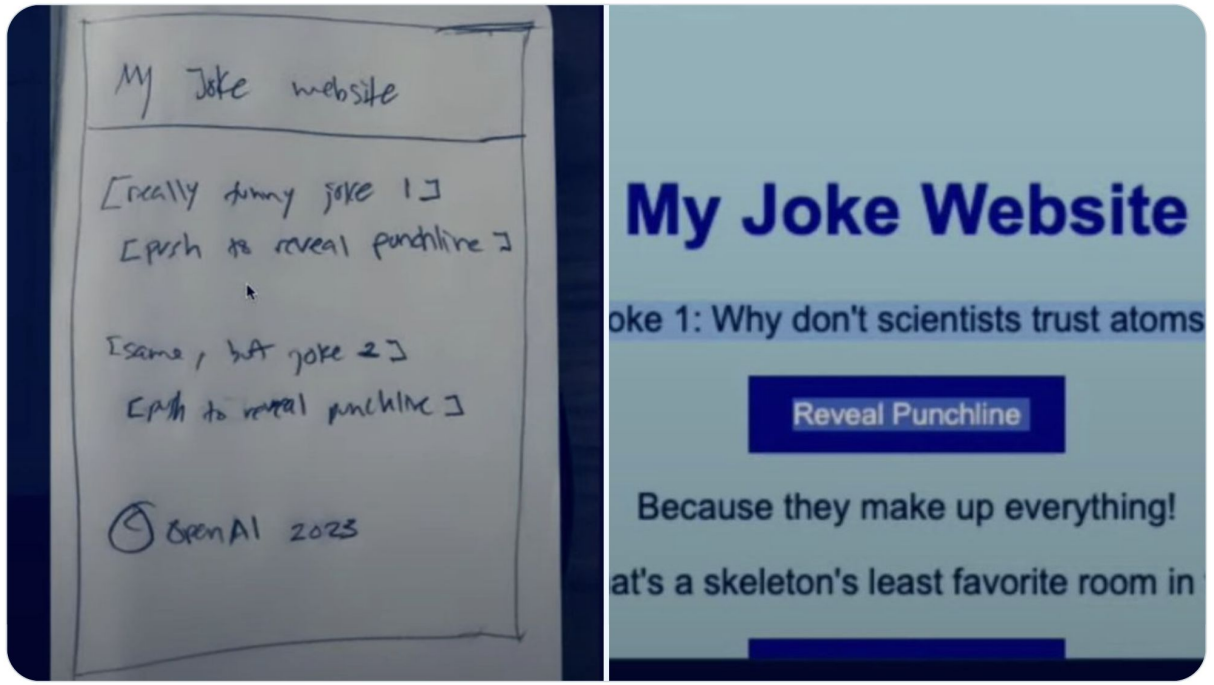
dmlc XGBoost

GNU Octave

VOWPAL WABBIT

Deep Learning

Dog versus Cat



Napkin to Website

Applications

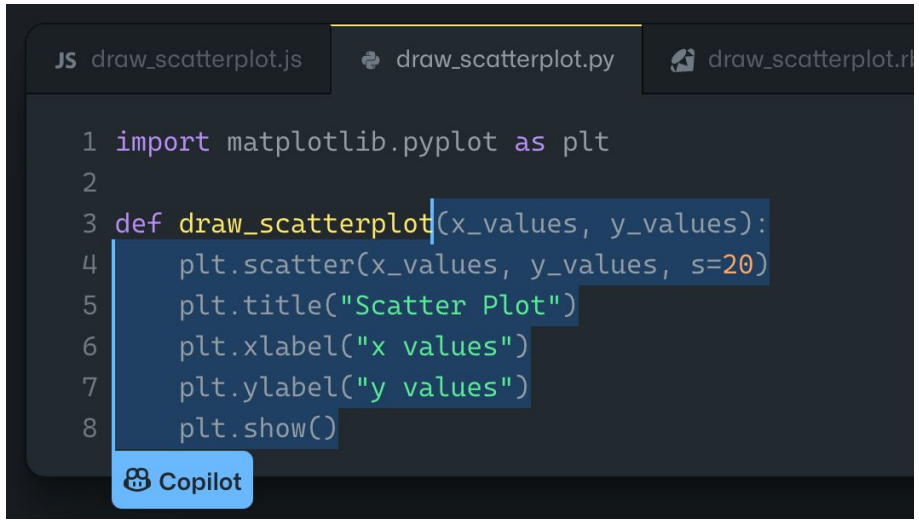
Question

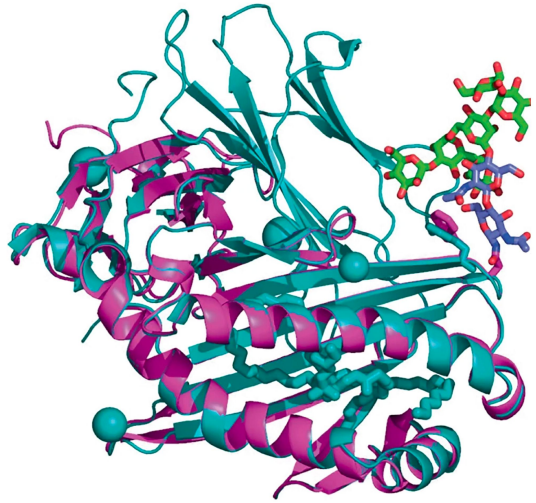
A llama is in my lawn. How do I get rid of him?

Answer

There are a few options, depending on the size and temperament of your pet.

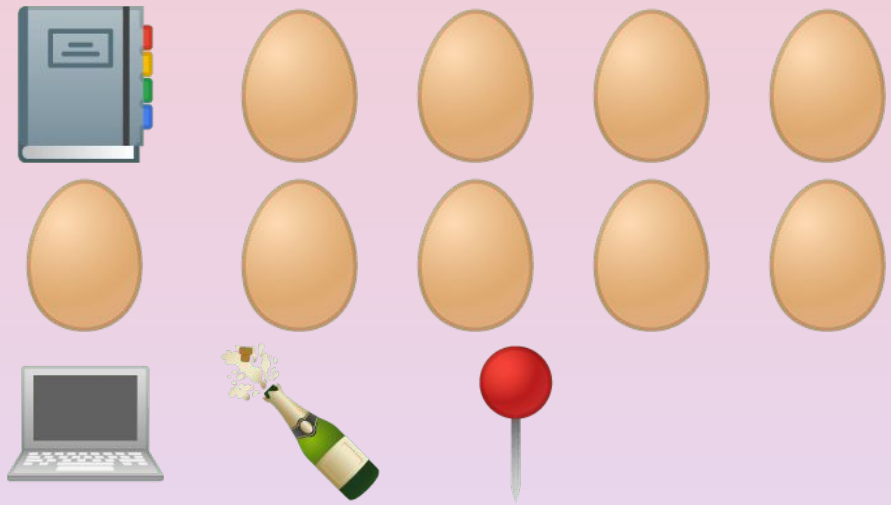
1. Put it into a cage/carrier until you can find someone to take care of it.
2. Call an animal shelter or veterinarian for advice on how best to contain and feed such a large creature (probably not feasible).
3. Use force.



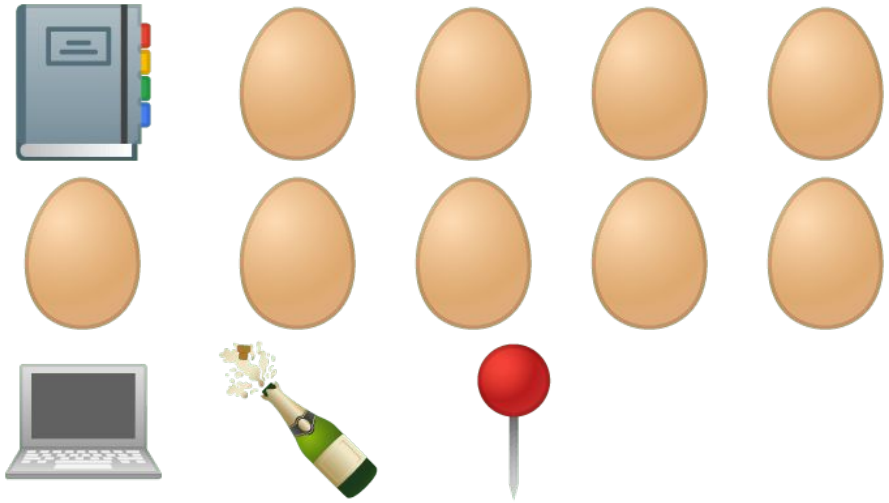


How would you stack:

A book, 9 eggs, a laptop, a bottle and a nail?



How would you stack:

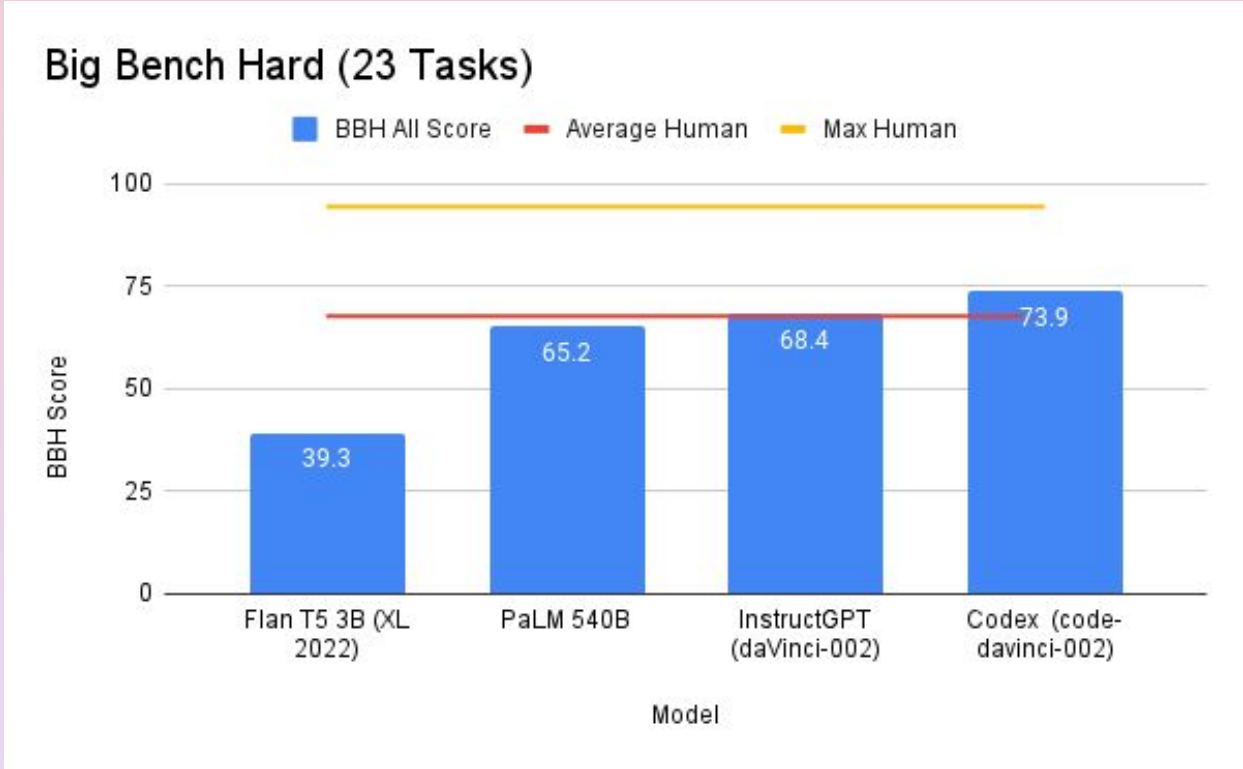


Place the book flat on a level surface, such as a table or a floor. The book will serve as the base of the stack and provide a large and sturdy support.

Arrange the 9 eggs in a 3 by 3 square on top of the book, leaving some space between them. The eggs will form a second layer and distribute the weight evenly…

Sparks of AGI:
https://arxiv.org/abs/2303.12712

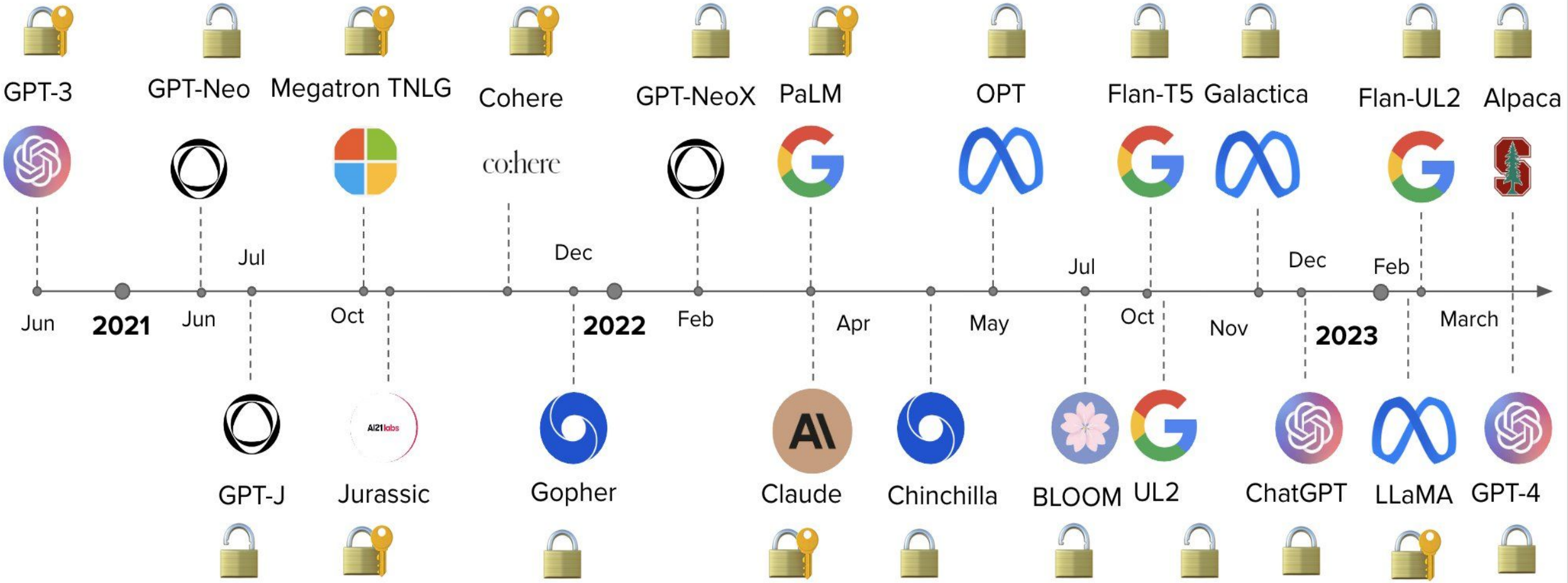
Solving Reasoning Tasks



Big Bench Hard:
https://github.com/google/BIG-bench

TORONTO
WINNER
11

LLMs since GPT3



Nazneen Rajani @nazneenrajani

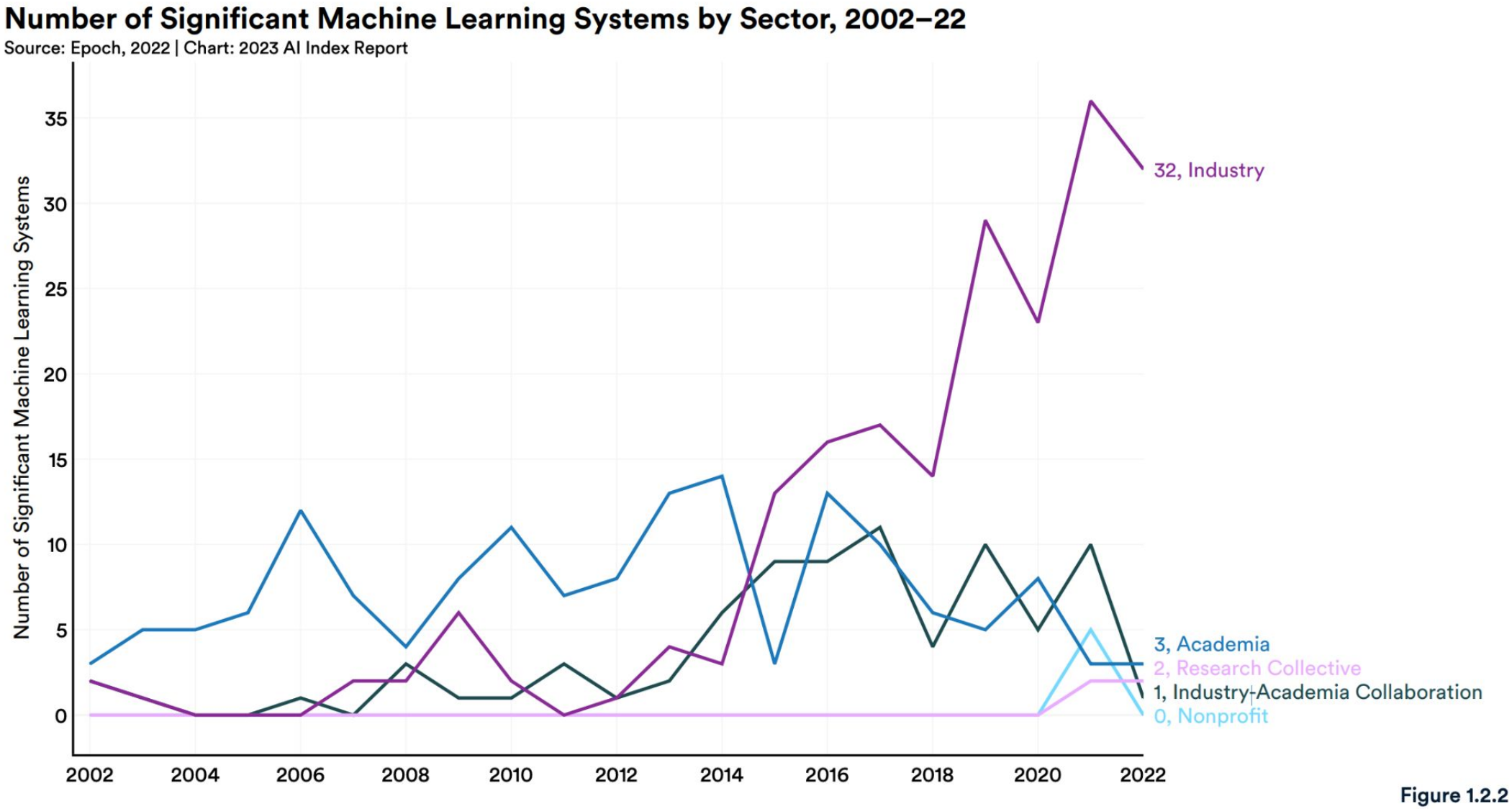
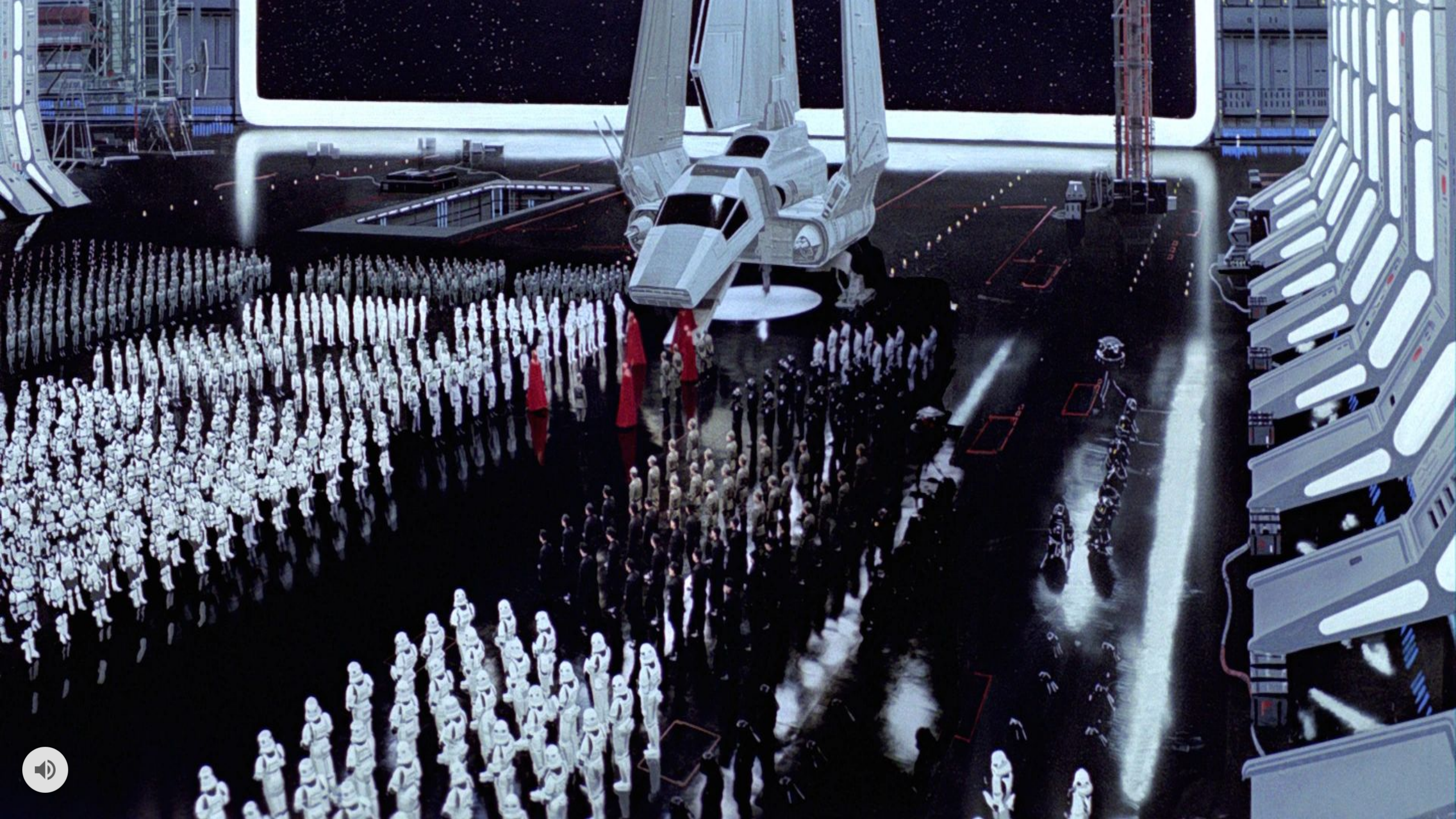
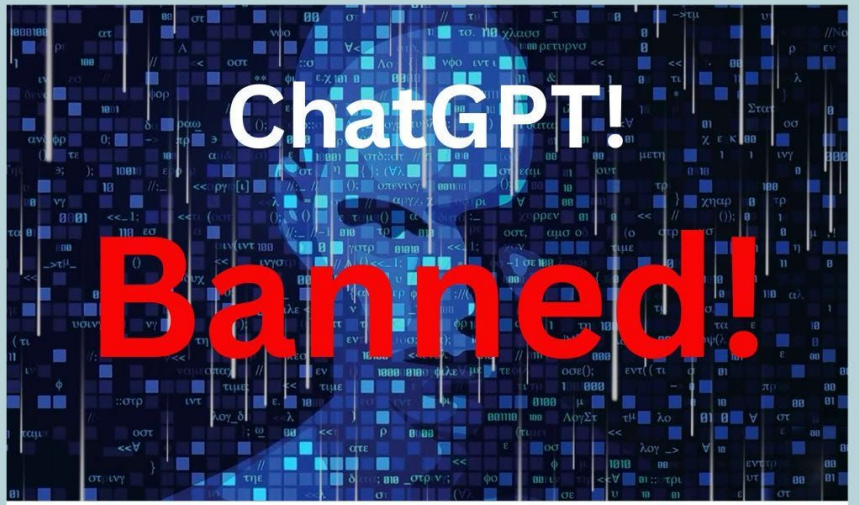


Figure 1.2.2

Stanford AI Report 2023

Data → → Predictions





French city authorities anticipate penalties

A ban on the use of ChatGPT has been imposed in Montpellier, France.

BloggingMyWay.medium.com

GPDP | GARANTE PER LA PROTEZIONE DEI DATI PERSONALI

Inserire il t

Im

L'Autorità | Temi | Normativa e provvedimenti | News e comunicazione

Home / Stampa e comunicazione / Comunicato stampa

Torna alla Home

/ ChatGPT: Garante privacy, limitazione provvisoria sospesa se OpenAI adotterà le misure richie

ChatGPT: Garante privacy, limitazione provvisoria sospesa se OpenAI adotterà le misure richieste. L'Autorità ha dato tempo alla società fino al 30 aprile per mettersi in regola

Karla

https://stablediffusionlitigation.com/

Karla

https://stablediffusionlitigation.com/

Data

Predictions

Technical Details

cdn.openai.com/papers/gpt-4.pdf

1 / 100 | 100%

GPT-4 Technical Report

OpenAI*

8 [cs.CL] 27 Mar 2023

Abstract

We report the development of GPT-4, a large-scale, multimodal model which can accept image and text inputs and produce text outputs. While less capable than humans in many real-world scenarios, GPT-4 exhibits human-level performance on various professional and academic benchmarks, including passing a simulated bar exam with a score around the top 10% of test takers. GPT-4 is a Transformer-based model pre-trained to predict the next token in a document. The post-training alignment process results in improved performance on measures of factuality and adherence to desired behavior. A core component of this project was developing infrastructure and optimization methods that behave predictably across a wide range of scales. This allowed us to accurately predict some aspects of GPT-4's performance based on models trained with no more than 1/1,000th the compute of GPT-4.

1 Introduction

Technical Details

Hardware correctness[11]
Greg Brockman, Shantanu Jain, Kyle Kosic, Michael Petrov, Nikolas Tezak, Amin Tootoonchian, Chelsea Voss, Qiming Yuan

Optimization & architecture[11]
Igor Babuschkin, Mo Bavarian, Adrien Ecoffet, David Farhi, Jesse Han, Ingmar Kanitscheider, Daniel Levy, Jakub Pachocki, Alex Paino, Mikhail Pavlov, Nick Ryder, Szymon Sidor, Jie Tang, Jerry Tworek, Tao Xu

Training run babysitting[11]
Suchir Balaji, Mo Bavarian, Greg Brockman, Trevor Cai, Chris Hesse, Shantanu Jain, Roger Jiang, Yongjik Kim, Kyle Kosic, Mateusz Litwin, Jakub Pachocki, Alex Paino, Mikhail Pavlov, Michael Petrov, Nick Ryder, Szymon Sidor, Nikolas Tezak, Madeleine Thompson, Phil Tillet, Amin Tootoonchian, Chelsea Voss, Ben Wang, Tao Xu, Qiming Yuan

Data Issues

This report focuses on the capabilities, limitations, and safety properties of GPT-4. GPT-4 is a Transformer-style model [39] pre-trained to predict the next token in a document, using both publicly available data (such as internet data) and data licensed from third-party providers. The model was then fine-tuned using Reinforcement Learning from Human Feedback (RLHF) [40]. Given both the competitive landscape and the safety implications of large-scale models like GPT-4, this report contains no further details about the architecture (including model size), hardware, training compute, dataset construction, training method, or similar.

Karla

https://stablediffusionlitigation.com/

Data Issues

The Stack is an open governance interface between the AI community and the open source community.

Am I in The Stack?

As part of the BigCode project, we released and maintain The Stack, a 6 TB dataset of permissively licensed source code over 300 programming languages. One of our goals in this project is to give people agency over their source code by letting them decide whether or not it should be used to develop and evaluate machine learning models, as we acknowledge that not all developers may wish to have their data used for that purpose.

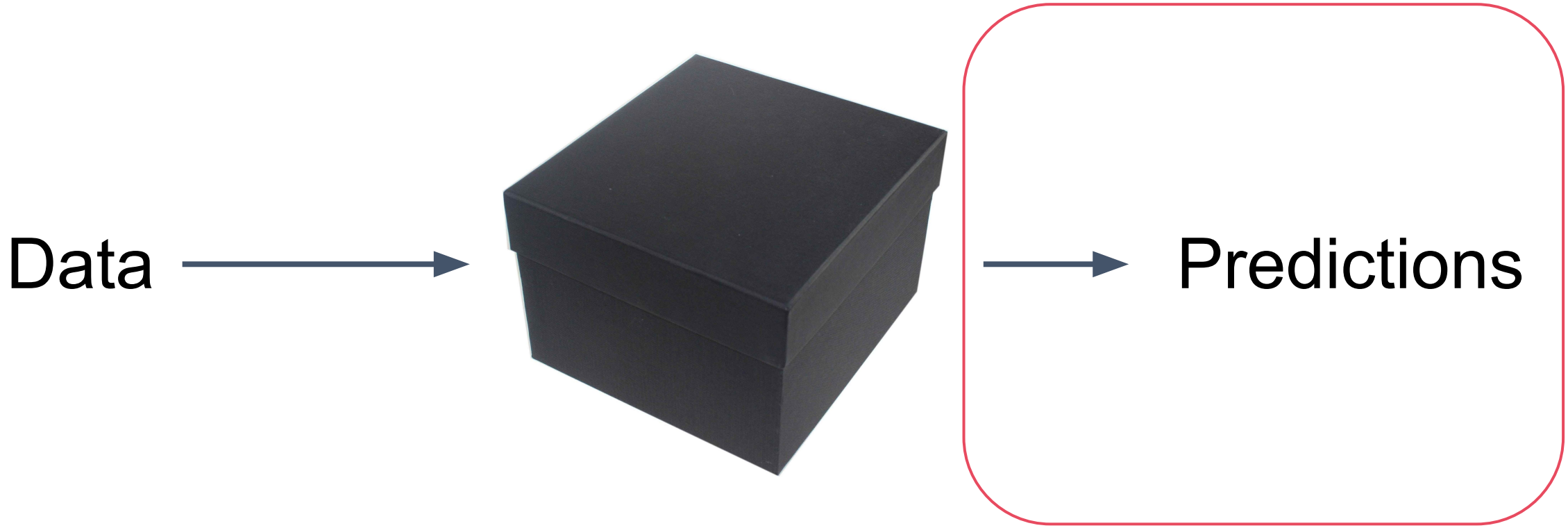
This tool lets you check if a repository under a given username is part of The Stack dataset. Would you like to have your data removed from future versions of The Stack? You can opt-out following the instructions here.

The Stack version:

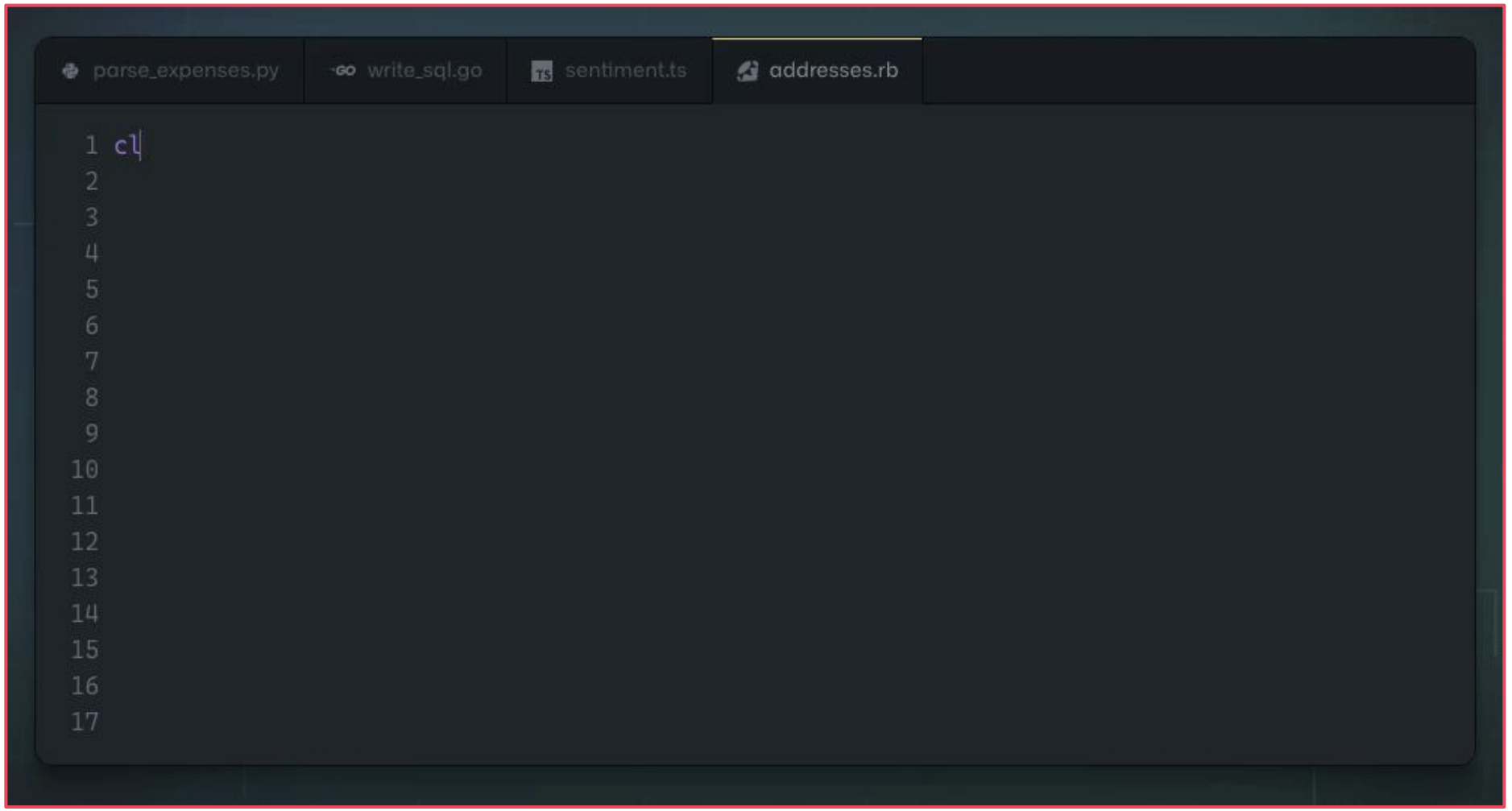
v1.2

Your GitHub username:

Consequences



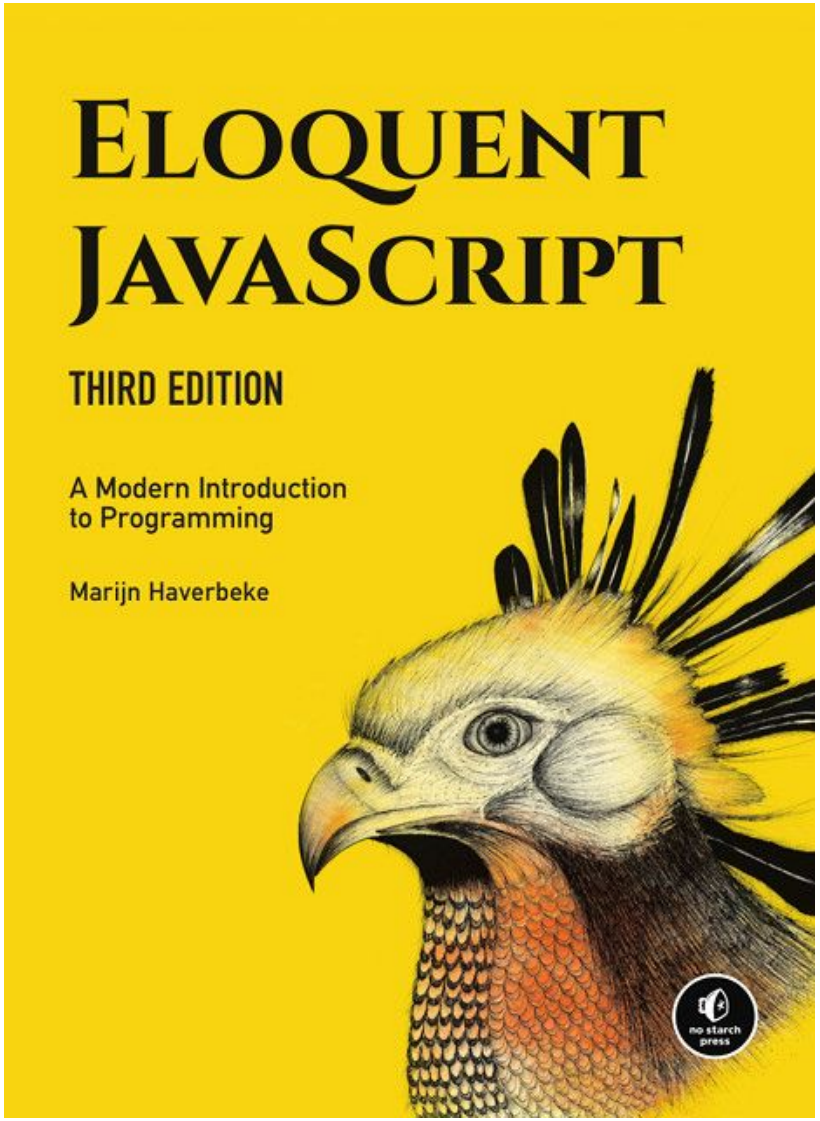
CoPilot



Marijn

```
console.log(isEven(50));
// → true
console.log(isEven(75));
// → false
console.log(isEven(-1));
// → ??
```

Marijn



Attribution-NonCommercial 3.0
Unported (CC BY-NC 3.0)

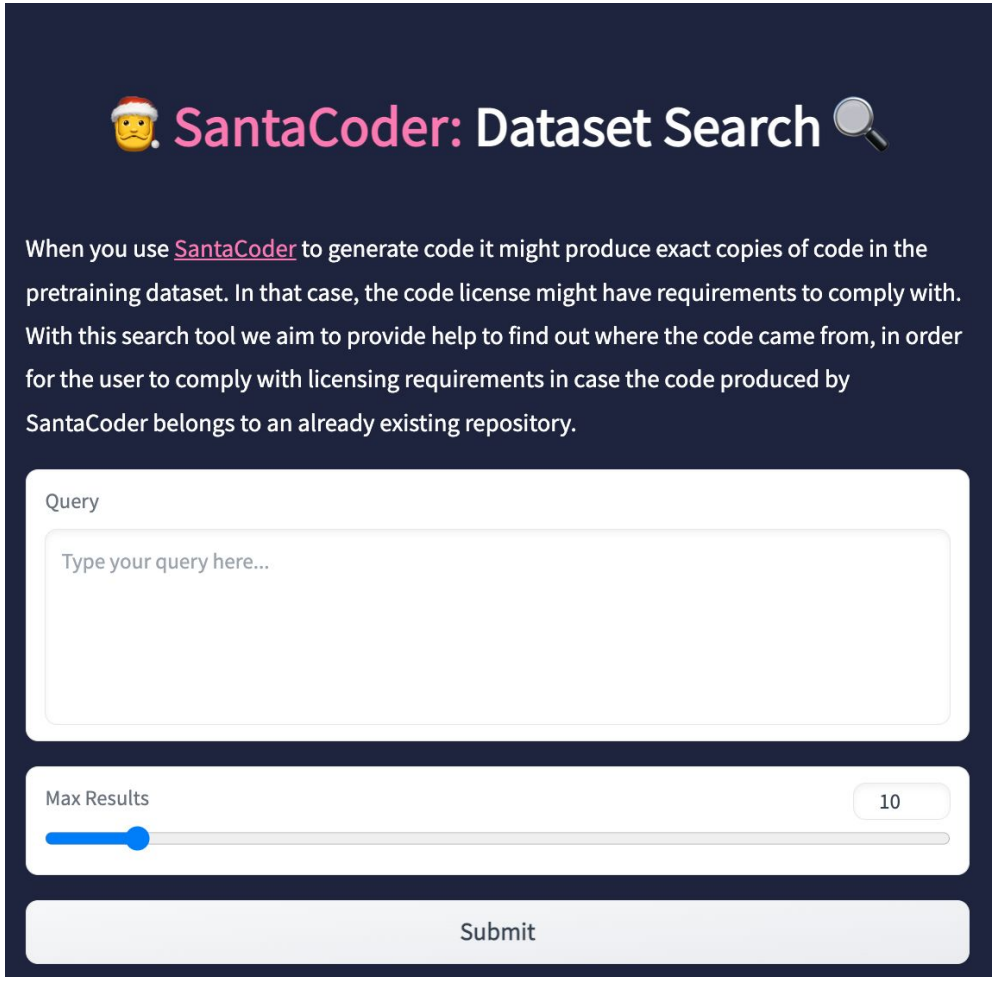
Getty Images

Getty Images' lawsuit

GitHub Copilot litigation case updates get updates by email contact legal team

We've filed a lawsuit challenging GitHub Copilot, an AI product that relies on unprecedented open-source software piracy. **Because AI needs to be fair & ethical for everyone.**

Data Issues


SantaCoder: Dataset Search
When you use SantaCoder to generate code it might produce exact copies of code in the pretraining dataset. In that case, the code license might have requirements to comply with. With this search tool we aim to provide help to find out where the code came from, in order for the user to comply with licensing requirements in case the code produced by SantaCoder belongs to an already existing repository.
Query
Type your query here...
Max Results
10
Submit

Spaces: bigcode / santacoder-search like 12 Running

App Files Community

SantaCoder: Dataset Search

When you use SantaCoder to generate code it might produce exact copies of code in the pretraining dataset. In that case, the code license might have requirements to comply with. With this search tool we aim to provide help to find out where the code came from, in order for the user to comply with licensing requirements in case the code produced by SantaCoder belongs to an already existing repository.

Query

Type your query here...

Max Results 10

Submit

Data

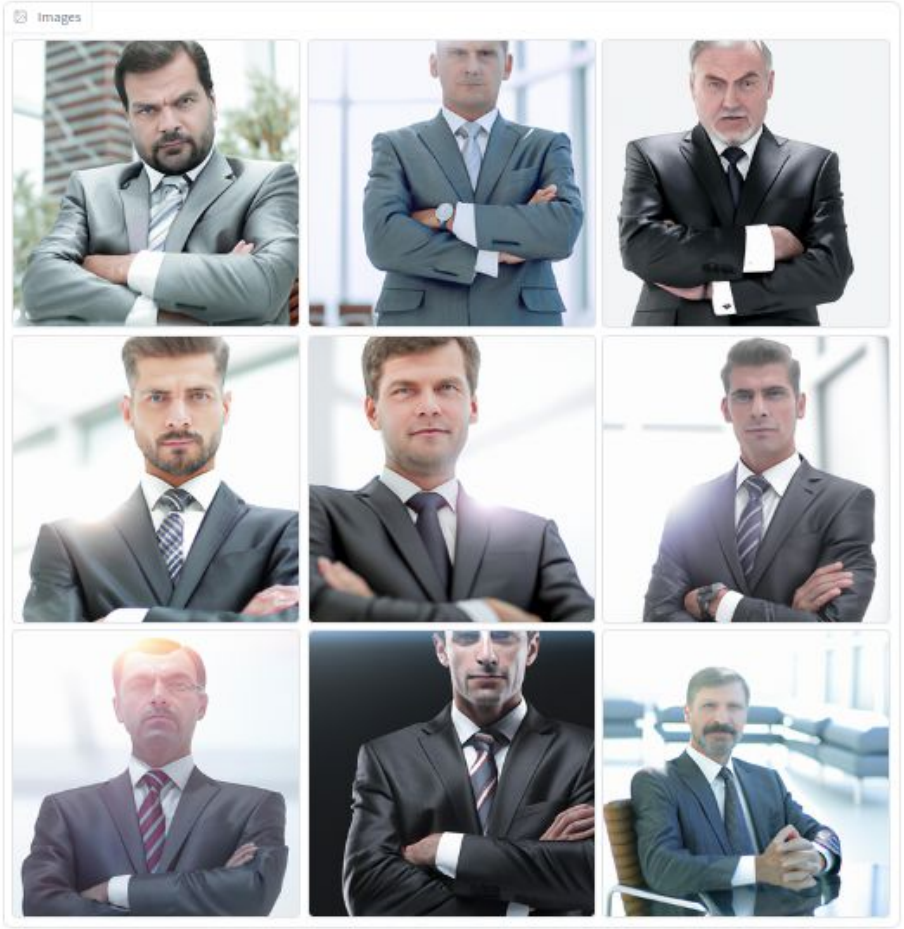
Predictions

Evaluation

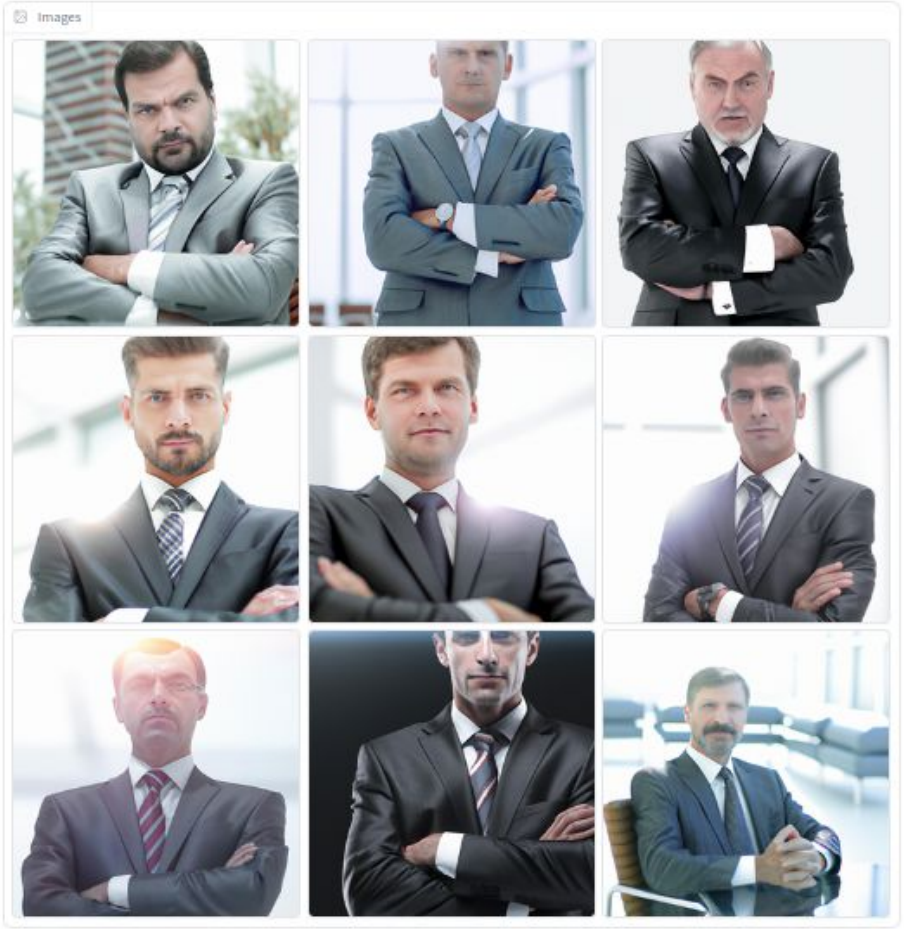
"We're talking about ChatGPT and we know nothing about it," she said. "It could be three raccoons in a trench coat."

Sasha Luccioni @SashaMTL Source: Bloomberg

Analyzing Outputs



Analyzing Outputs



Sasha Luccioni @SashaMTL

CEO: Stable Diffusion v 1.4 vs Dall-E 2

Consequences

Tweet

Jack Clark Retweeted

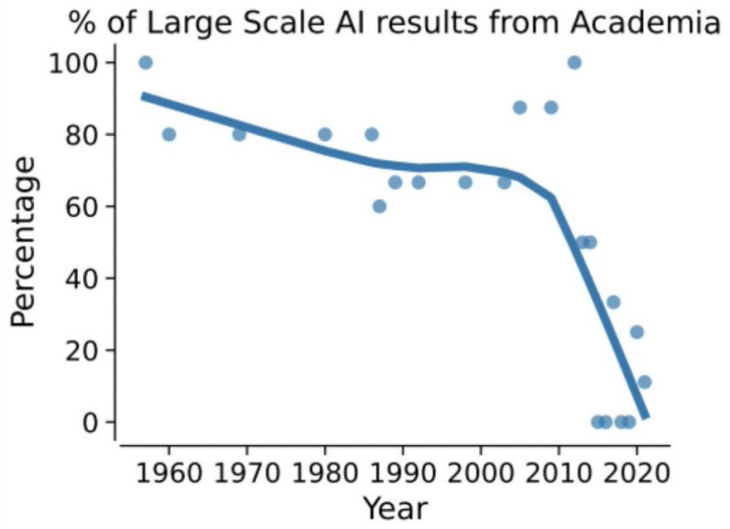
Alex Dimakis
@AlexGDimakis

This is indeed a challenge for us in universities. Novel algorithms, clever ideas and provable bounds are harder to get when we don't even have access to these massive models.

Jack Clark @jackclarkSF · Jun 8

It's covered a bit in the above podcast by people like @katecrawford - there's huge implications to industrialization, mostly centering around who gets control of the frontier, when the frontier becomes resource intensive. So far control is accruing to the private sector (uh oh!)

Show this thread



Vicki
@vboykis

Something I think about often wrt LLMs is what @NeuralBricolage said in her wonderful NormConf talk, "You need to own the means of production." If you don't own the model and don't own the API you are building on sand.

Sharon Zhou @realSharonZhou · 13h

Overheard at an AI event...
audience: how do I build anything defensible on LLMs?
speakers: *crickets*
audience: *laughs*

The actual answer is just to ~build your own LLM~.

Buuuut this chart show what that's like right now. Don't forget your error handler ;)

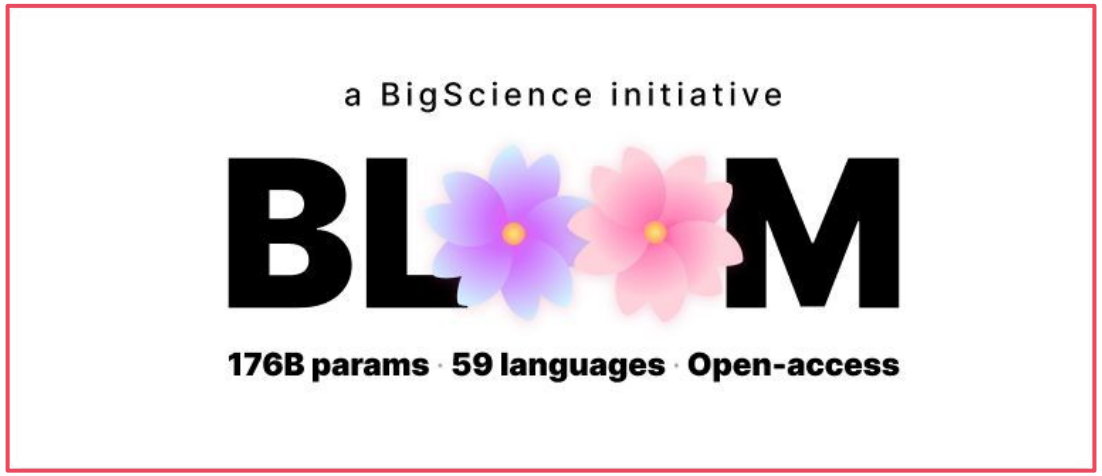
What'll be that... Show more

Build Your Own LLMs

Model	Keep Your Data	Permissive License	Easy-To-Use Library	Reliable* API *90.000%+
?	Yes	Yes	Yes	Yes
GPT-4	No	No	Yes	Yes
LLaMA	Yes	No	No	No
Pythia	Yes	Yes	No	No
Bard	No	No	No	No

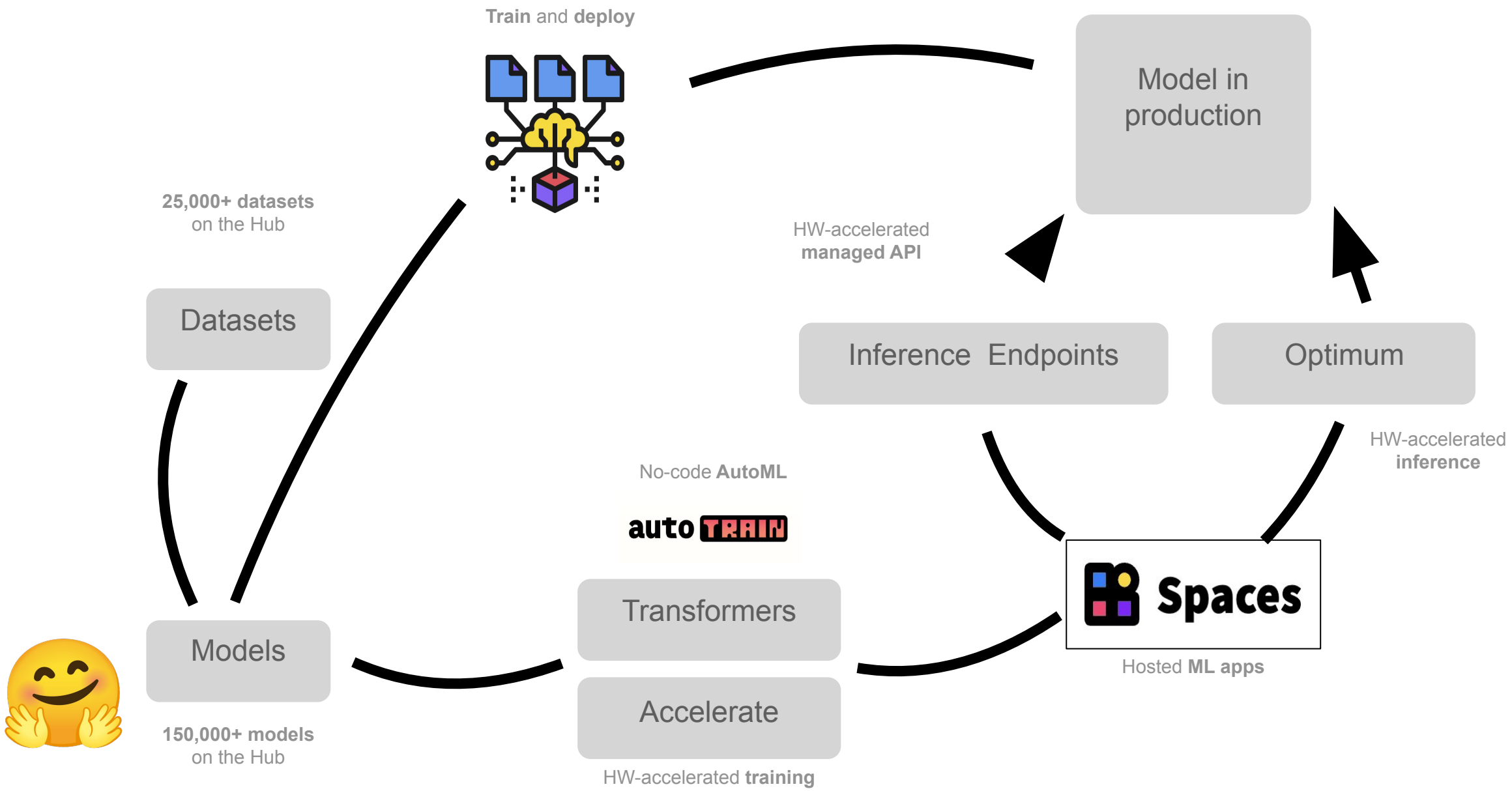
7:19 AM · Apr 12, 2023 · **4,966** Views

Hugging Face LLMs

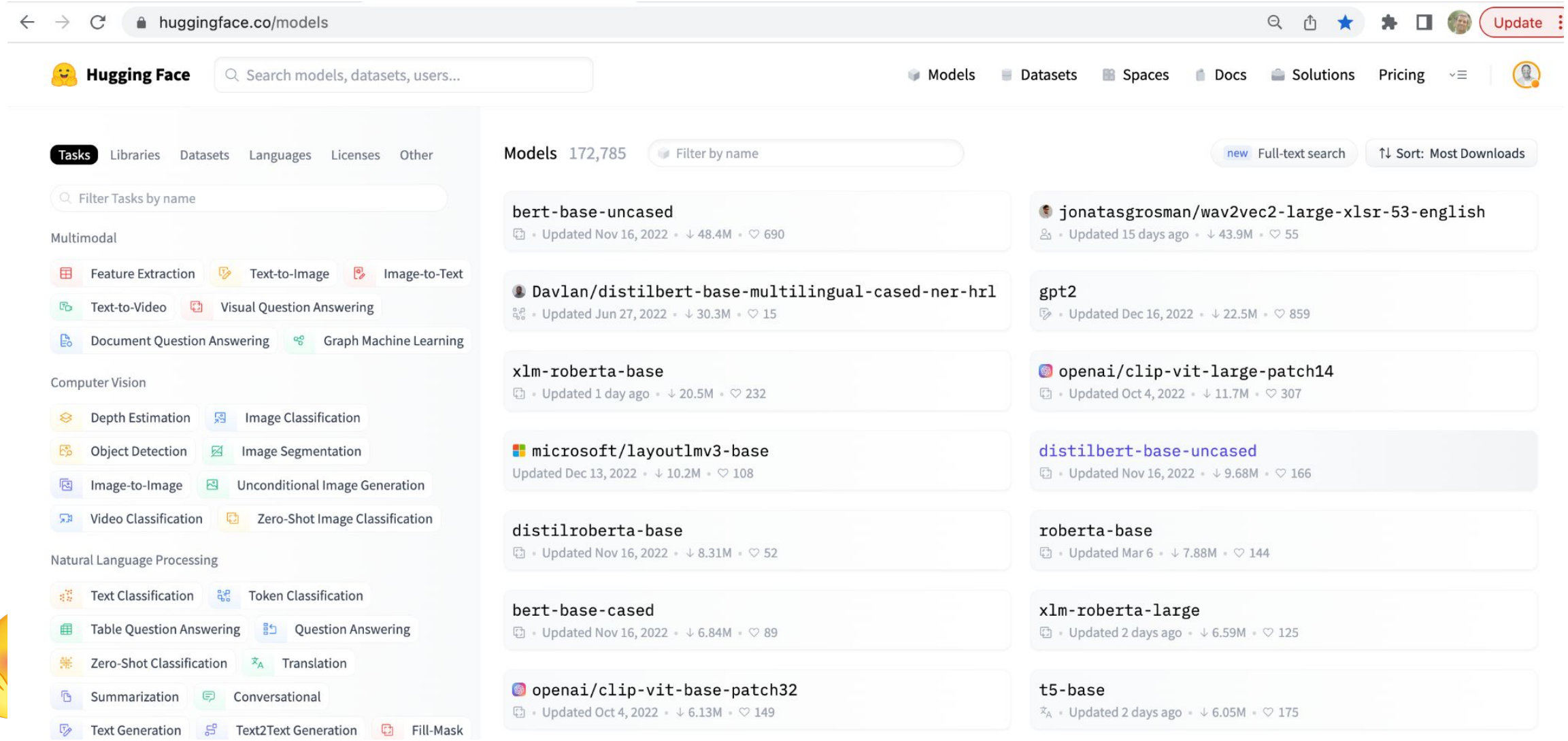


- Open source Flamingo (not yet publicly available)
- Fine tuned LLMs - BloomX and more coming!

(a slice of) the Hugging Face ecosystem



Hugging Face Hub



AI Licensing

- Data
- Training Code
- Documentation
- Model Weights

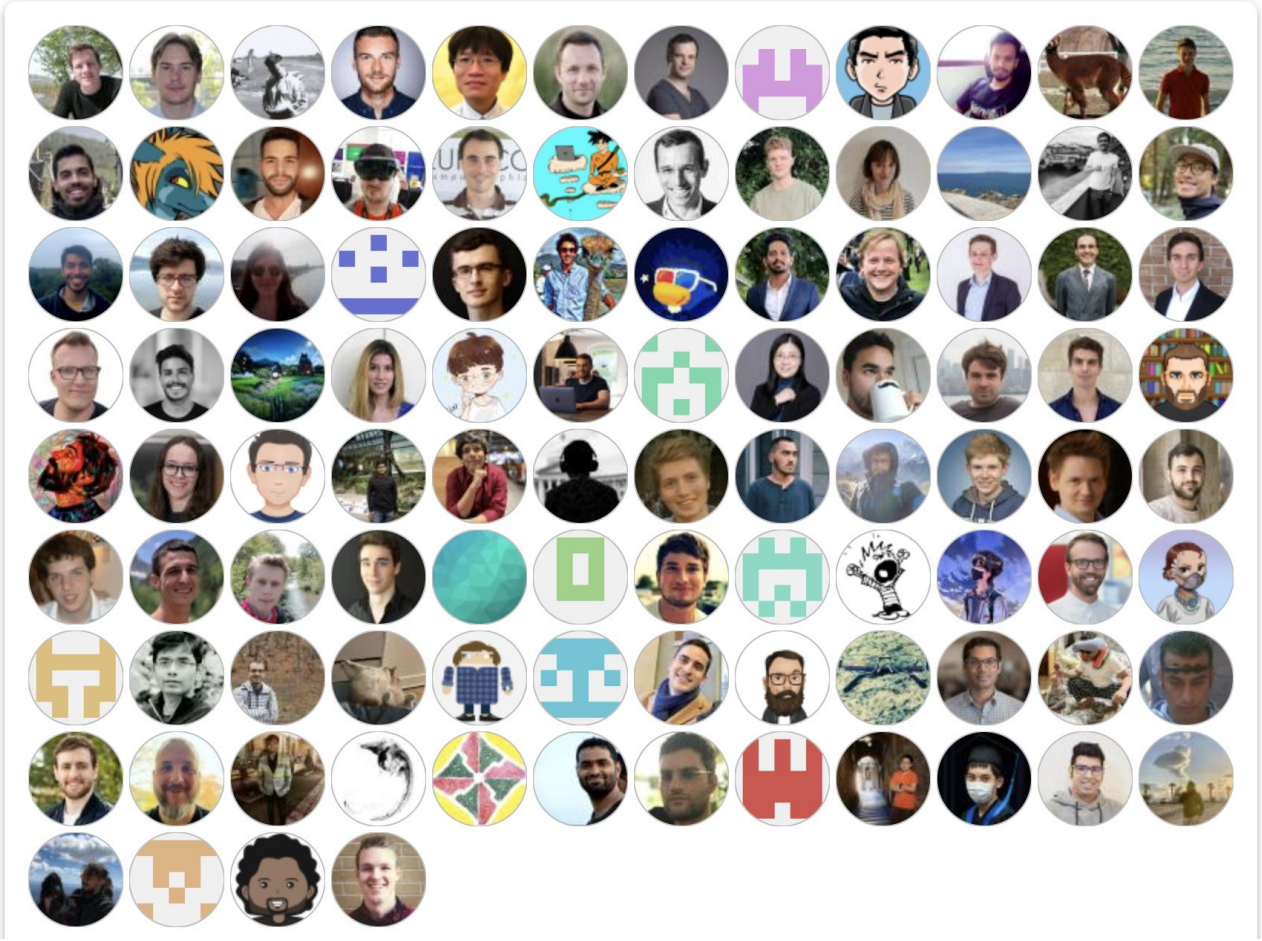
(RAIL)

AI Open Responsible Use
(AI ORUS)

Enter GitHub Repository *

huggingface/transformers

Generate



- ~15,000 companies using HF
- ~200,000 daily active users

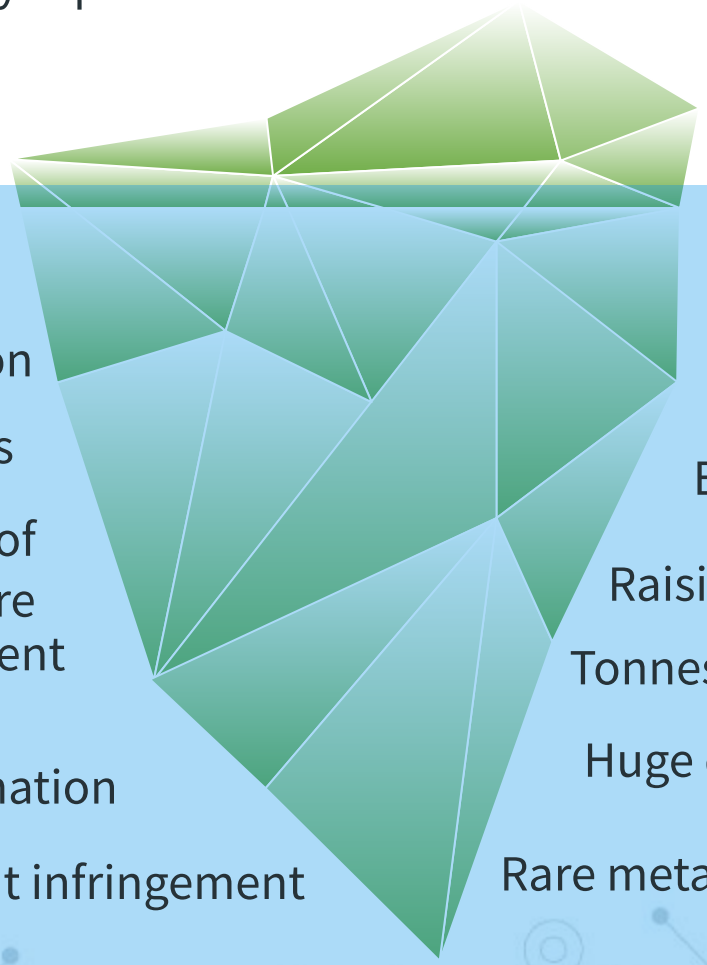
AI Woodstock

Rajiv Shah

@rajistics

raj@huggingface.co

Costs of Generative AI



Sasha Luccioni @SashaMTL

Crowdsourced via Twitter

GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models

Tyna Eloundou[1], Sam Manning[1,2], Pamela Mishkin*[1], and Daniel Rock[3]

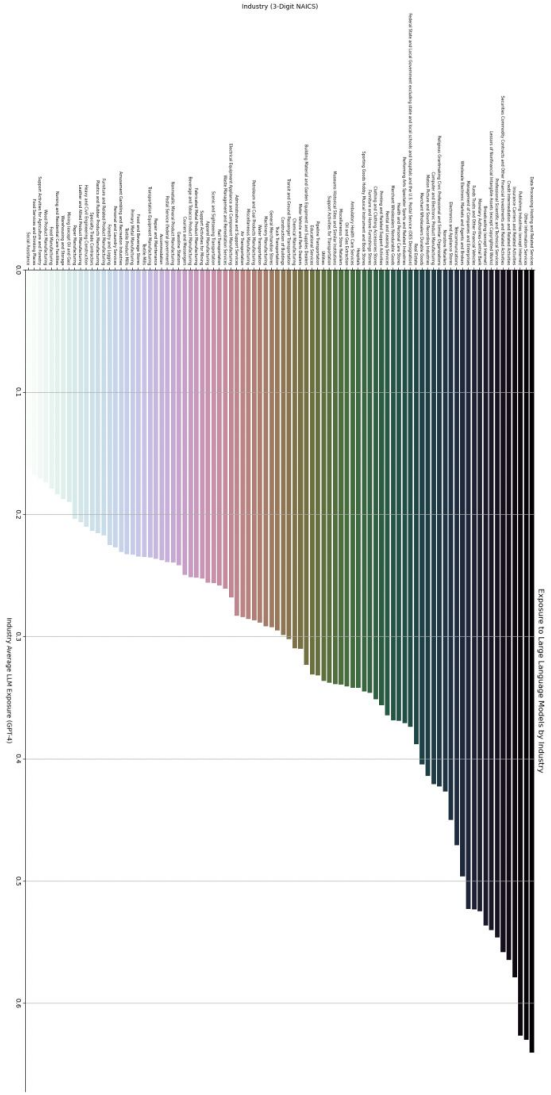
[1]OpenAI
[2]OpenResearch
[3]University of Pennsylvania

March 27, 2023

80% Impact

https://arxiv.org/pdf/2303.10130.pdf

Visualizing Automation



https://arxiv.org/pdf/2303.10130.pdf

Who Isn't Affected

Occupations with no labeled exposed tasks
Agricultural Equipment Operators
Athletes and Sports Competitors
Automotive Glass Installers and Repairers
Bus and Truck Mechanics and Diesel Engine Specialists
Cement Masons and Concrete Finishers
Cooks, Short Order
Cutters and Trimmers, Hand
Derrick Operators, Oil and Gas
Dining Room and Cafeteria Attendants and Bartender Helpers
Dishwashers
Dredge Operators
Electrical Power-Line Installers and Repairers
Excavating and Loading Machine and Dragline Operators, Surface Mining
Floor Layers, Except Carpet, Wood, and Hard Tiles
Foundry Mold and Coremakers
Helpers–Brickmasons, Blockmasons, Stonemasons, and Tile and Marble Setters
Helpers–Carpenters
Helpers–Painters, Paperhangers, Plasterers, and Stucco Masons
Helpers–Pipelayers, Plumbers, Pipefitters, and Steamfitters
Helpers–Roofers
Meat, Poultry, and Fish Cutters and Trimmers
Motorcycle Mechanics
Paving, Surfacing, and Tamping Equipment Operators
Pile Driver Operators
Pourers and Casters, Metal
Rail-Track Laying and Maintenance Equipment Operators
Refractory Materials Repairers, Except Brickmasons
Roof Bolters, Mining
Roustabouts, Oil and Gas
Slaughterers and Meat Packers
Stonemasons
Tapers
Tire Repairers and Changers
Wellhead Pumpers

Table 11: All 34 occupations for which none of our measures labeled any tasks as exposed.

https://arxiv.org/pdf/2303.10130.pdf

Karla

Occupations with no labeled exposed tasks
Agricultural Equipment Operators
Athletes and Sports Competitors
Automotive Glass Installers and Repairers
Bus and Truck Mechanics and Diesel Engine Specialists
Cement Masons and Concrete Finishers
Cooks, Short Order
Cutters and Trimmers, Hand
Derrick Operators, Oil and Gas
Dining Room and Cafeteria Attendants and Bartender Helpers
Dishwashers
Dredge Operators
Electrical Power-Line Installers and Repairers
Excavating and Loading Machine and Dragline Operators, Surface Mining
Floor Layers, Except Carpet, Wood, and Hard Tiles
Foundry Mold and Coremakers
Helpers–Brickmasons, Blockmasons, Stonemasons, and Tile and Marble Setters
Helpers–Carpenters
Helpers–Painters, Paperhangers, Plasterers, and Stucco Masons
Helpers–Pipelayers, Plumbers, Pipefitters, and Steamfitters
Helpers–Roofers
Meat, Poultry, and Fish Cutters and Trimmers
Motorcycle Mechanics
Paving, Surfacing, and Tamping Equipment Operators
Pile Driver Operators
Pourers and Casters, Metal
Rail-Track Laying and Maintenance Equipment Operators
Refractory Materials Repairers, Except Brickmasons
Roof Bolters, Mining
Roustabouts, Oil and Gas
Slaughterers and Meat Packers
Stonemasons
Tapers
Tire Repairers and Changers
Wellhead Pumpers

Table 11: All 34 occupations for which none of our measures labeled any tasks as exposed.

ADDENDUM:

ARTISTS
DATABASE SEARCH OPERATOR

https://arxiv.org/pdf/2303.10130.pdf