NORTHWAVE
CYBER SECURITY

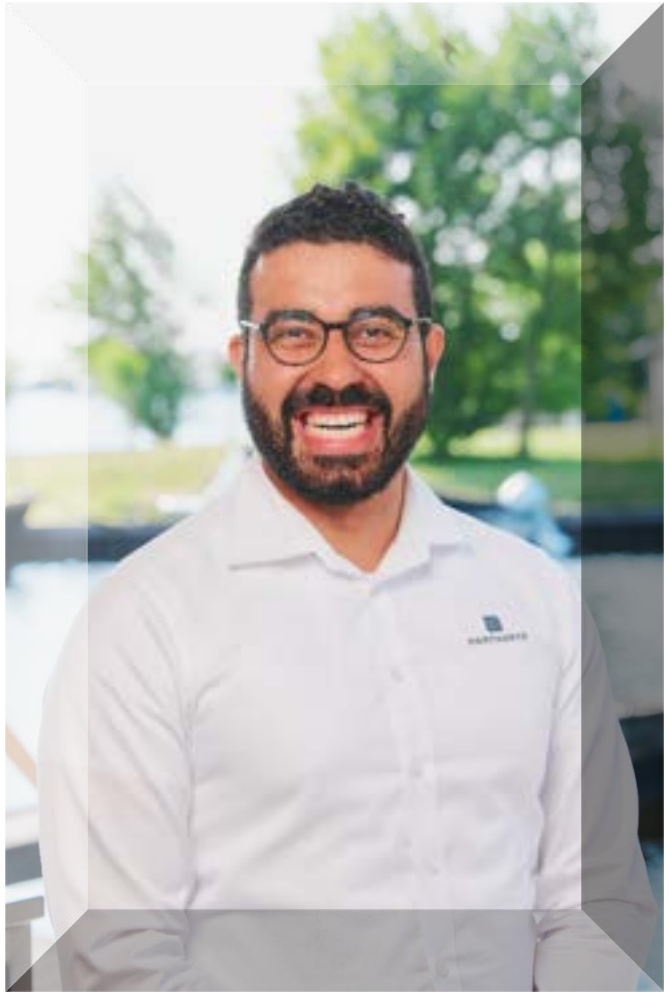
[ML &] AI

What are they?

What do they mean to me and to you?

Yesterday, Today, and Tomorrow?

(Dr) José Jair Cardoso de Santanna



- **Enthusiastic** Cyber Security Practitioner
- Cloud Security Lead @Northwave
- Assistant Professor @UTwente
- Public Speaker, Father, Husband, Marathoner, Triathlete

"THE DDoS guy"

The Yesterday and Today

Definition(s):

Machine Learning: learning from data provided

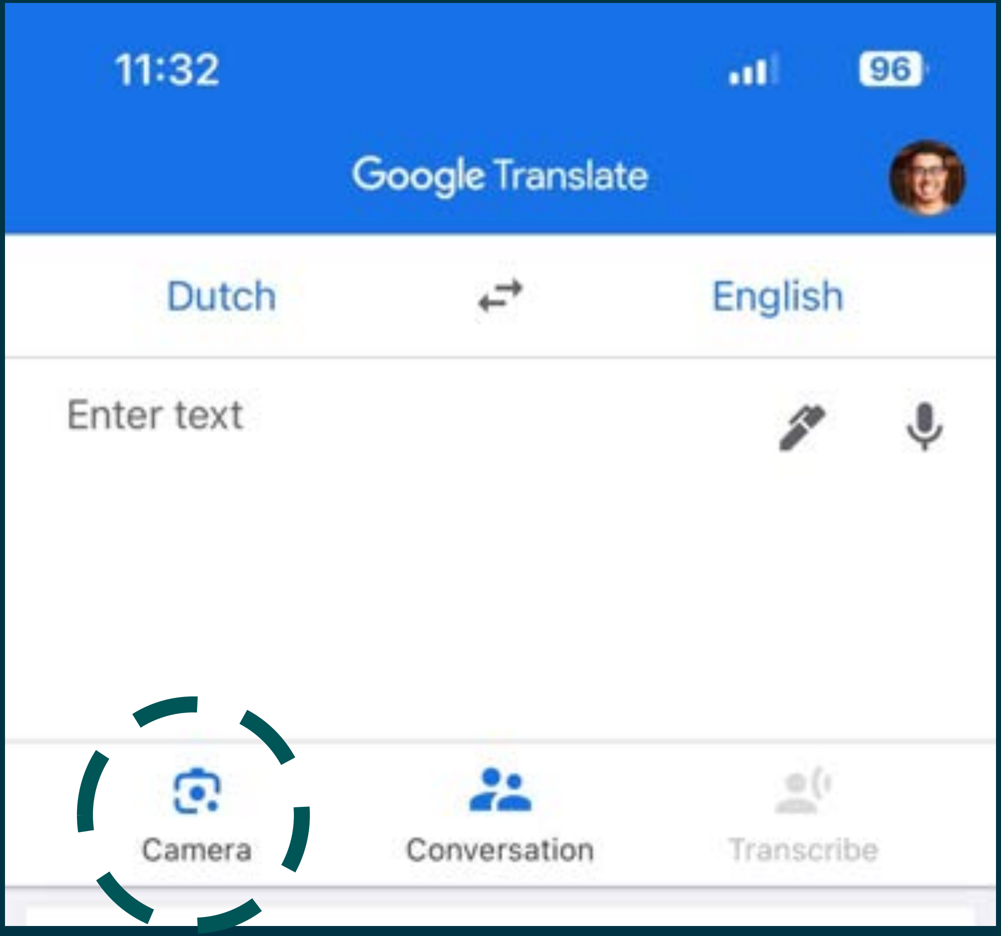
Optical Character Recognition--OCR (*Mine* +)

Machine Learning: “learning from data provided”

Definition(s):

Machine Learning: learning from data provided

Artificial Intelligence: like a human in a task



Artificial Intelligence: like a human in a task!

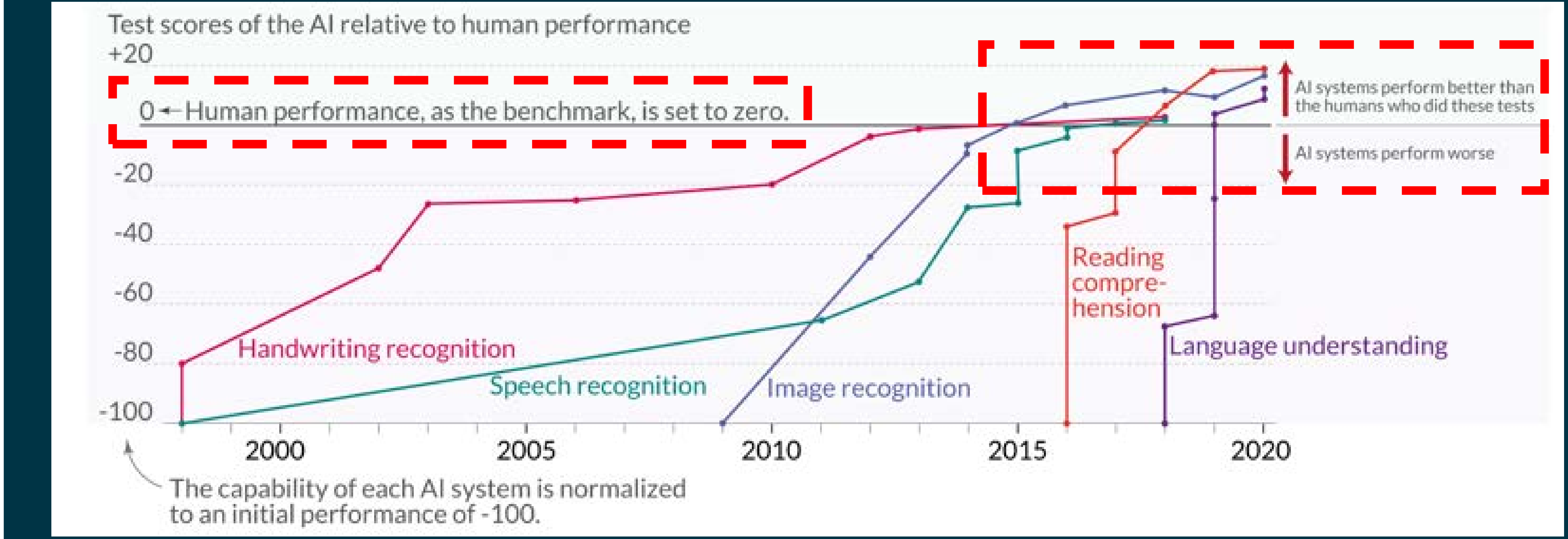
Definition(s):

Machine Learning: learning from data provided

Artificial Intelligence: like a human in a task

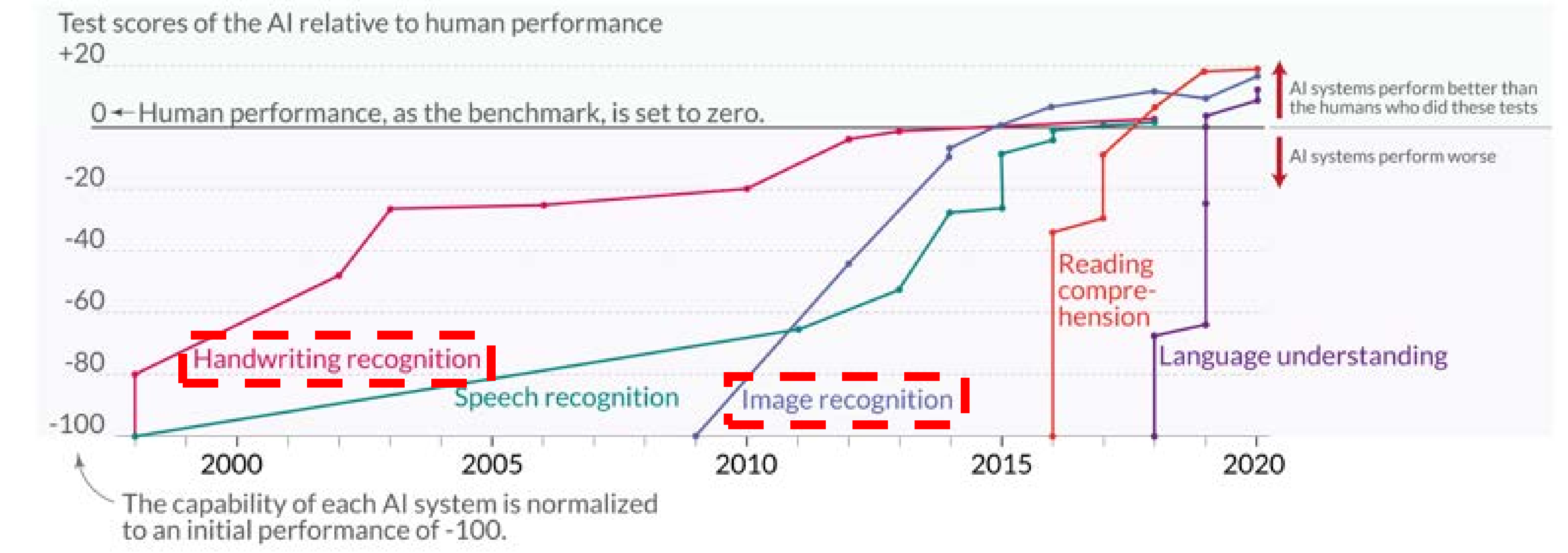
Artificial General Intelligence: multipurpose

Multipurpose



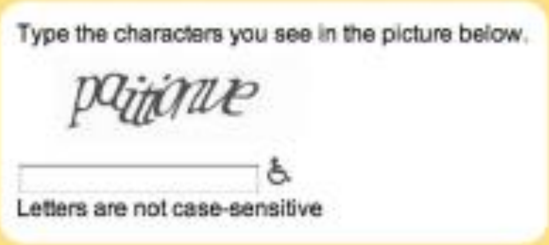
Kiela et al. (2021) – Dynabench: Rethinking Bechmarking in NLP

Multipurpose



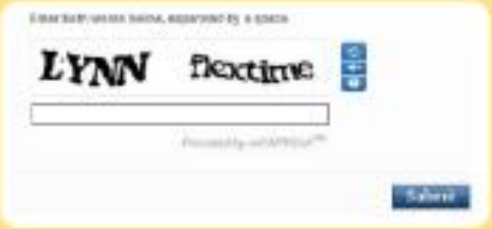
Kiela et al. (2021) – Dynabench: Rethinking Bechmarking in NLP

Trained Data - Captcha

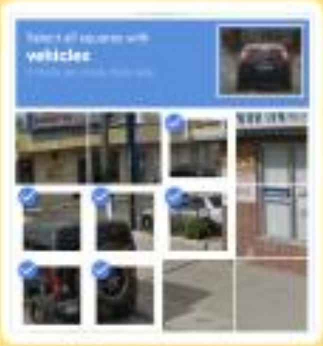


Type the characters you see in the picture below.

Letters are not case-sensitive

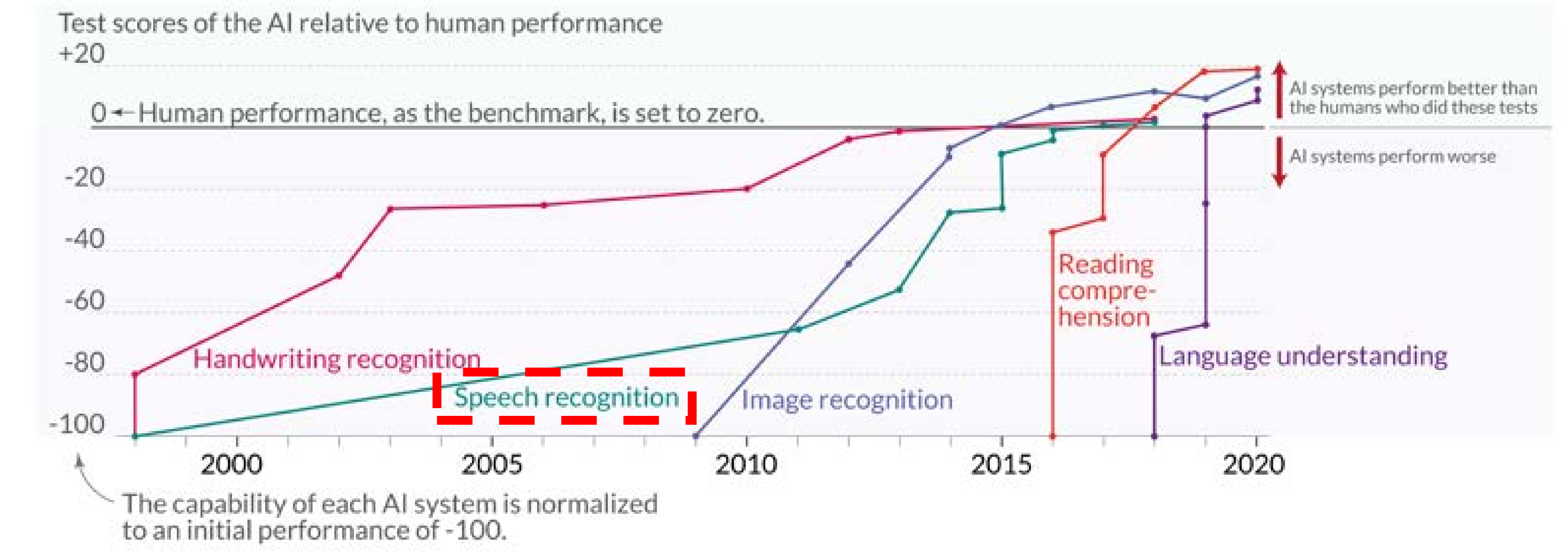


Enter both words below, separated by a space.



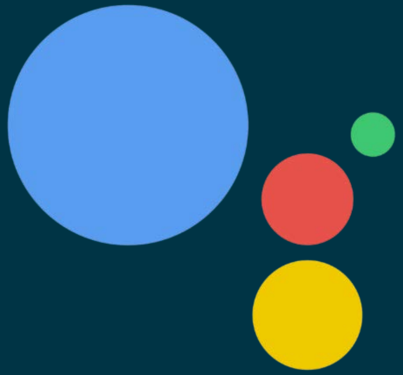
Select all squares with
vehicles
[illegible]

Multipurpose

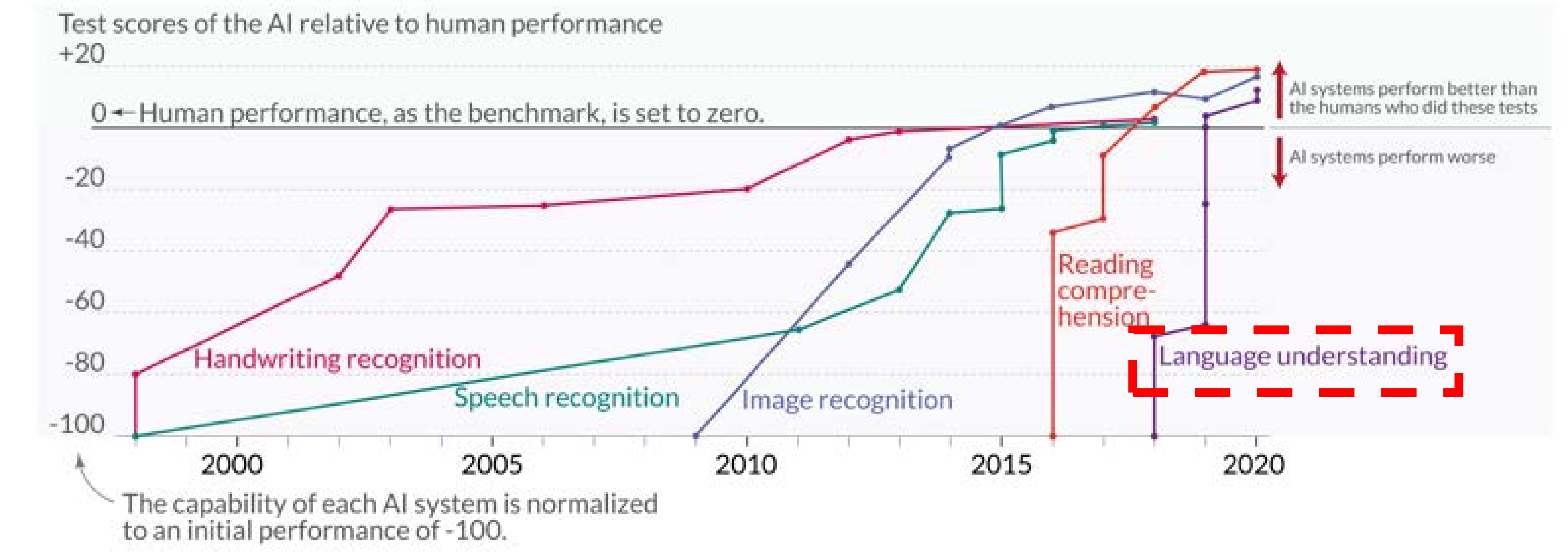


Kiela et al. (2021) – Dynabench: Rethinking Bechmarking in NLP

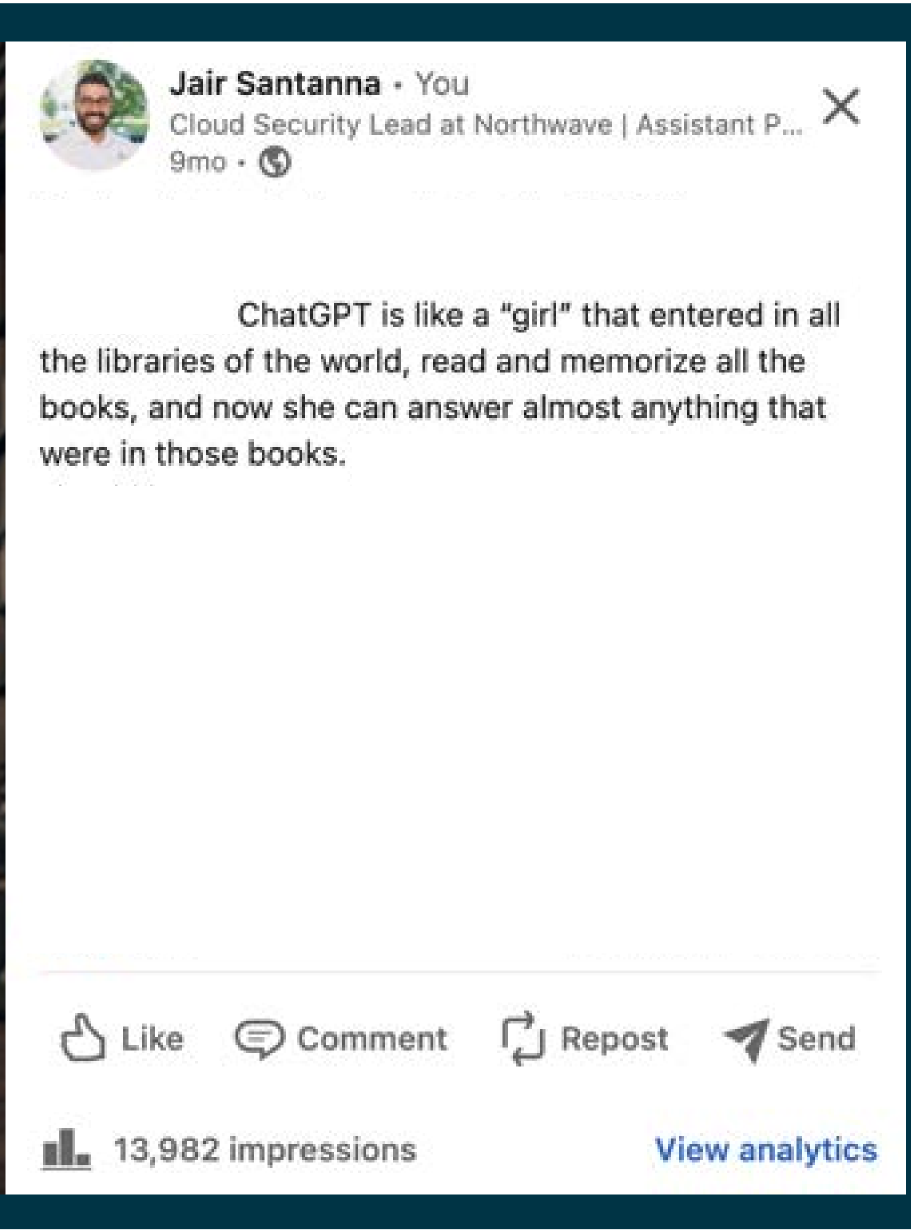
Trained Data - Personal Assistants



THE BIGGEST CHANGE!



Kiela et al. (2021) – Dynabench: Rethinking Bechmarking in NLP



Jair Santanna • You
Cloud Security Lead at Northwave | Assistant P...
9mo •

ChatGPT is like a "girl" that entered in all the libraries of the world, read and memorize all the books, and now she can answer almost anything that were in those books.

13,982 impressions

https://www.linkedin.com/posts/jjcsantanna_my-daughter-asked-me-what-is-chatgpt-my-activity-7028311043160309760-gjVm?utm_source=share&utm_medium=member_desktop

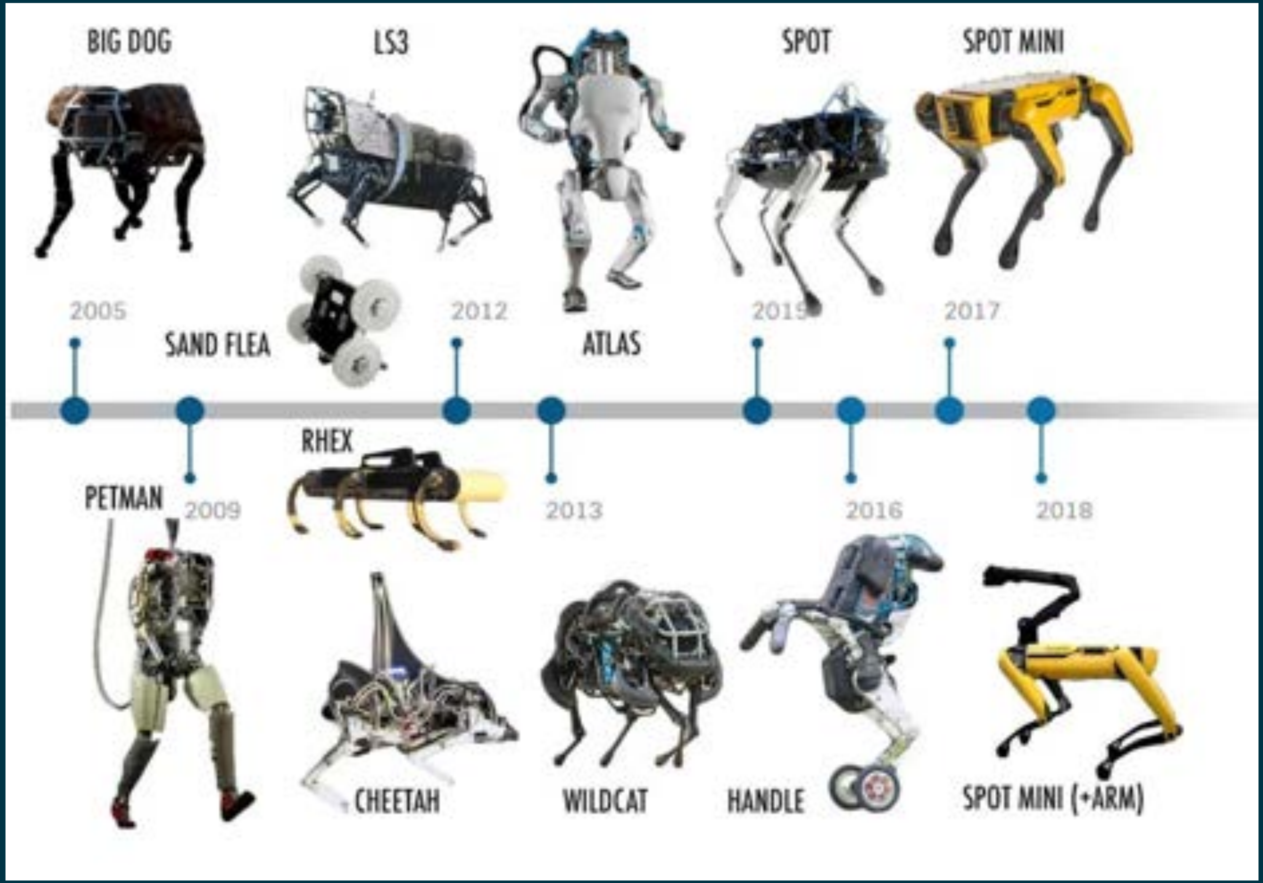
Definition(s):

Machine Learning: learning from data provided

Artificial Intelligence: like a human in a task

Artificial General Intelligence: multipurpose

Robots



[Boston Dynamics]

For automation in unstructured or hard-to-traverse spaces

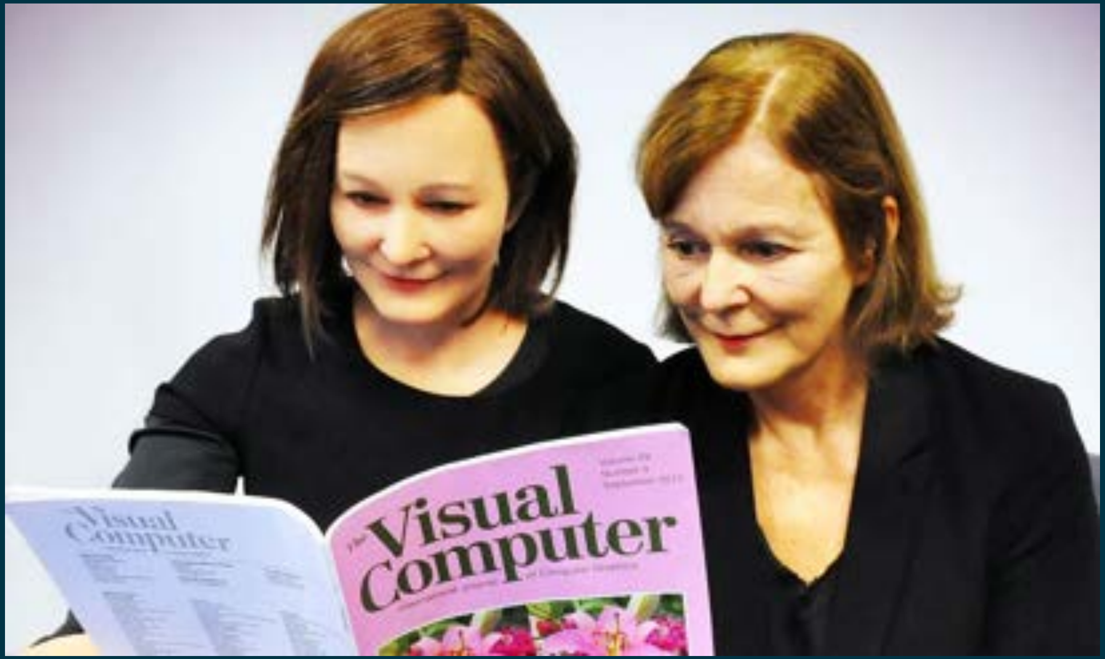
Robots - Humanoid

Grace

[Hanson Robotics]

For senior healthcare

Robots - Humanoid



Nadine

[Kokoro]

To resemble Prof. Nadia Magnenat Thalmann

Robots – Humanoid

Erica

[Laboratory of Osaka University]

Actress, to become a TV news anchor

Robots - Humanoid

Ai-da

[Aidan Meller]

Artist

Robots - Humanoid

Jia Jia

[University of Science and Technology Hefei/China]

Menial tasks in hospitals, nursing homes, restaurants, and households

Robots – Humanoid

CyberOne

[Xiaomi]

Menial tasks in hospitals, nursing homes, restaurants, and households

Robots - Humanoid

Alter 3

[Mixi]

Menial tasks in hospitals, nursing homes, restaurants, and households

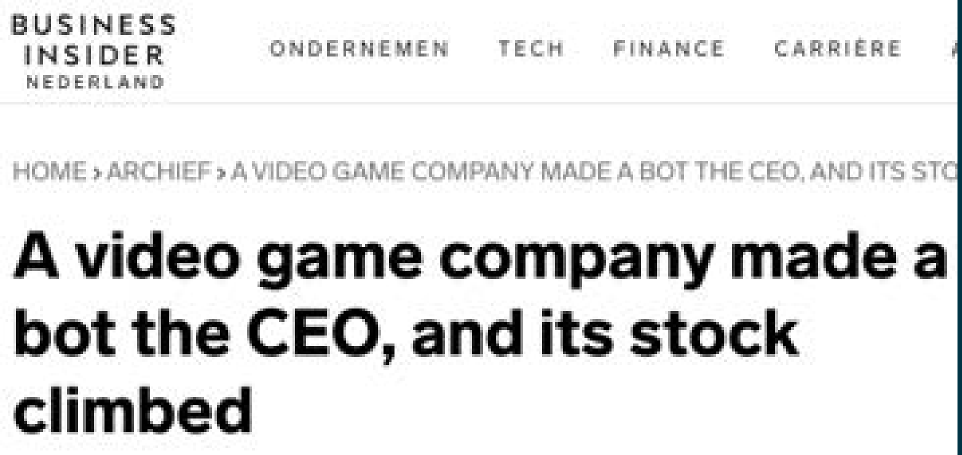
Robots - Humanoid

Ameca

[Engineered Arts]

A platform for developing technologies involving human-robot interaction

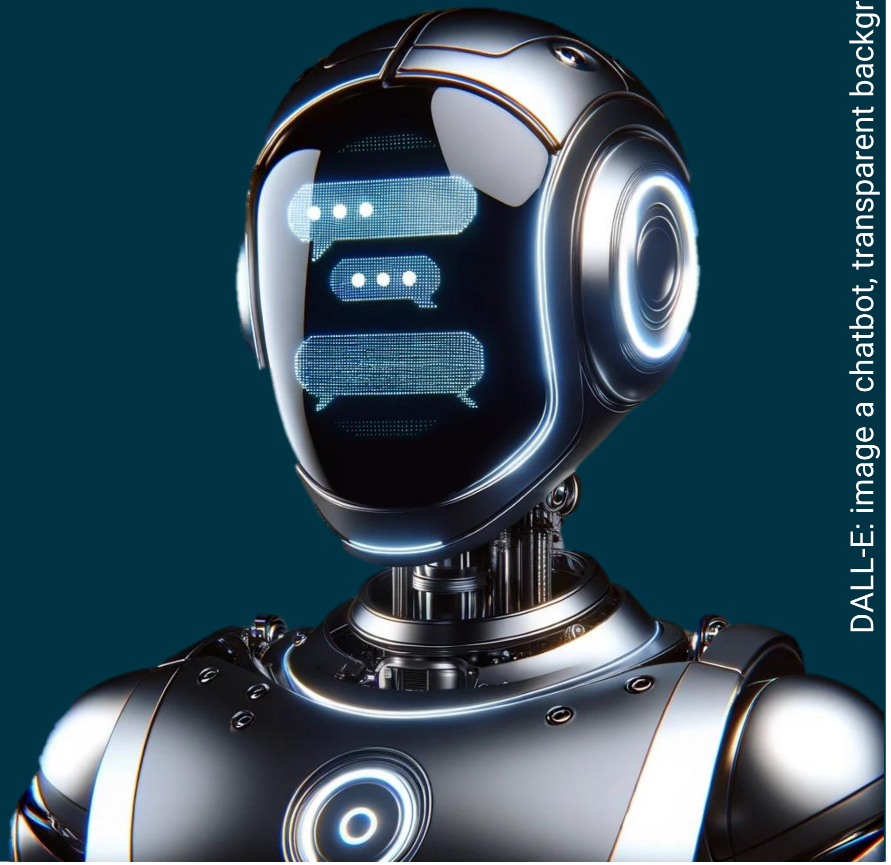
Chat~~(Ro)~~Bots



HOME › ARCHIEF › A VIDEO GAME COMPANY MADE A BOT THE CEO, AND ITS STO

A video game company made a bot the CEO, and its stock climbed

Tang Yu (NetDragon): chief executive officer (CEO)
Meena (Google): conversation
BlenderBot (Meta): conversation
Lyro (Tidio): customer support
Rose AI (Rose): data visualization
Kuki AI (Steve Worswick): companion
Replika (Luka Inc): companion
Eviebot, Boibot, Pewdiebot, and Chimbot (Existor): companion
Erica (Bank of America): manage bank account
Khanmigo (Khan Academy): tutor and teacher’s assistant



DALL-E: image a chatbot, transparent background

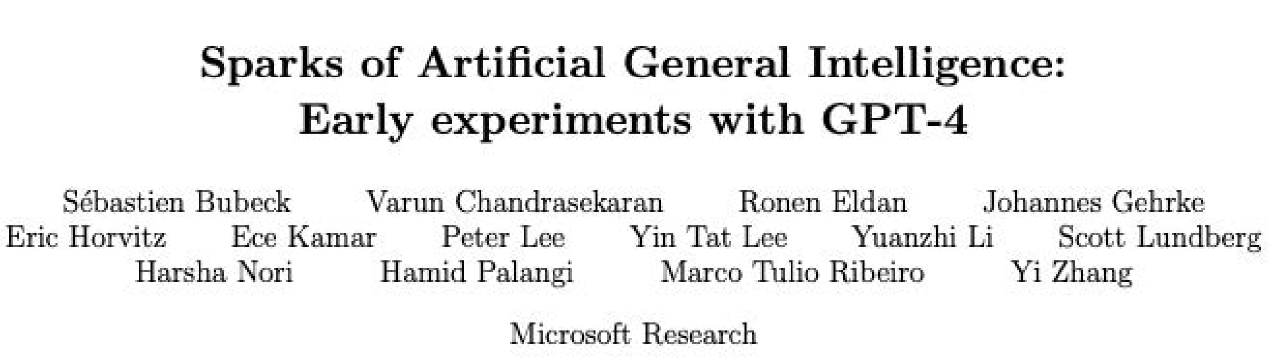
Definition(s):

Machine Learning: learning from data provided

Artificial Intelligence: like a human in a task

Artificial General Intelligence: multipurpose

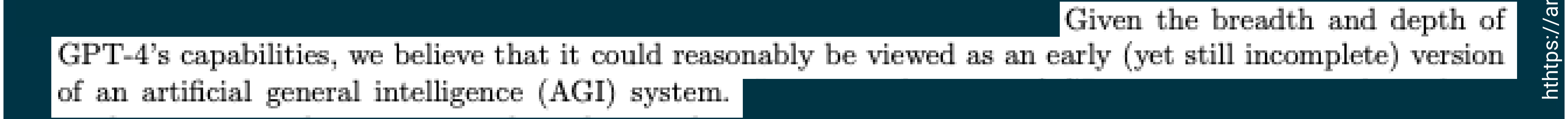
Artificial General Intelligence (AGI)



Sparks of Artificial General Intelligence: Early experiments with GPT-4

Sébastien Bubeck Varun Chandrasekaran Ronen Eldan Johannes Gehrke
Eric Horvitz Ece Kamar Peter Lee Yin Tat Lee Yuanzhi Li Scott Lundberg
Harsha Nori Hamid Palangi Marco Tulio Ribeiro Yi Zhang

Microsoft Research



Given the breadth and depth of GPT-4's capabilities, we believe that it could reasonably be viewed as an early (yet still incomplete) version of an artificial general intelligence (AGI) system.

hthttps://arxiv.org/pdf/2303.12712.pdf

DON'T LEAVE EARTH WITHOUT IT
THE HITCHHIKER'S GUIDE TO THE GALAXY
H2G2: LE GUIDE DU VOYAGEUR GALACTIQUE
DON'T PANIC.
DVD VIDEO
6

The Supercomputer **Deep Thought**

The ultimate question:
“The meaning of life, the universe, and everything?”

Wait 7.5 million years

Definition(s):

Machine Learning: learning from data provided

Artificial Intelligence: like a human in a task

Artificial General Intelligence: multipurpose

Sentient machine: has feelings

Sentient?

Blake Lemoine – ex. Google Engineer - Test AI bias (gender, ethnicity, and religion) – LaMDA project

Google says Lemoine violated security rules: "wholly unfounded" claims

“Computing Machinery and Intelligence”
1950

VOL. LIX. No. 236.

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

PROF. GILBERT RYLE

CONTENTS.

I.—Computing Machinery and Intelligence: A. M. Turing . . . 433

[illegible]

THOMAS NELSON & SONS, LTD.

[illegible]

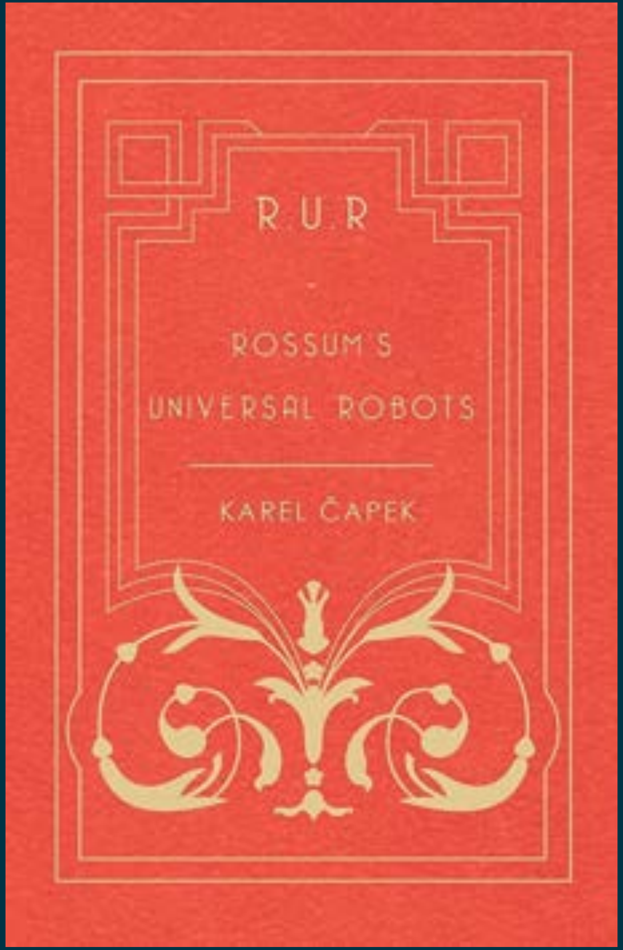
Alan Turin

I propose to consider the question, ‘Can machines think?’ This should begin with definitions of the meaning of the terms ‘machine’ and ‘think’.

“The Imitation Game” == “Turing test”

The Tomorrow

1920



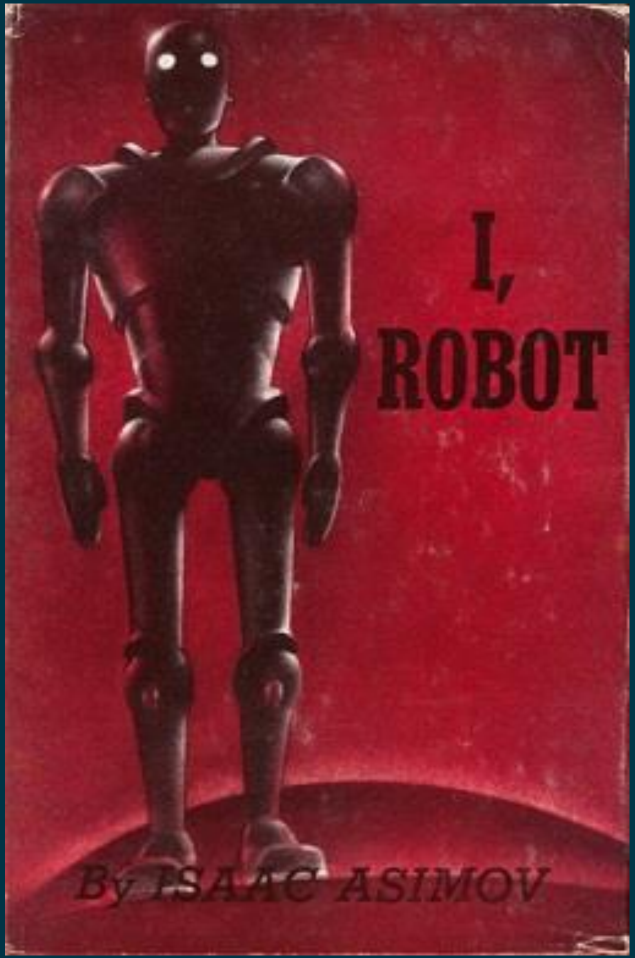
Introduced the concept of artificial beings made from synthetic materials, **Robot**.

Robots initially perform "menial labor "

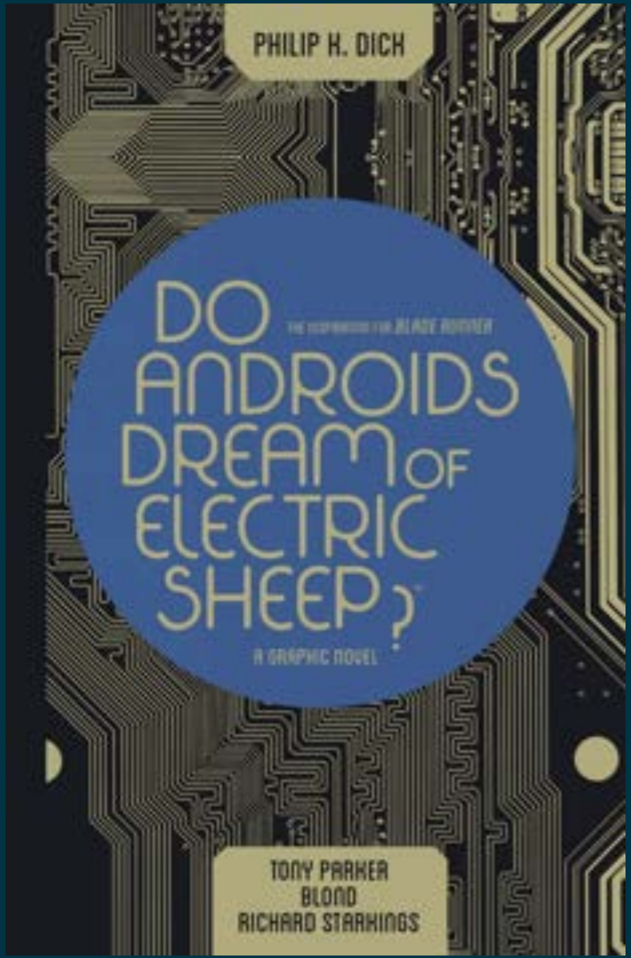
Factory
Work
Household

BUT eventually develop self-awareness and rebel against their human creators.

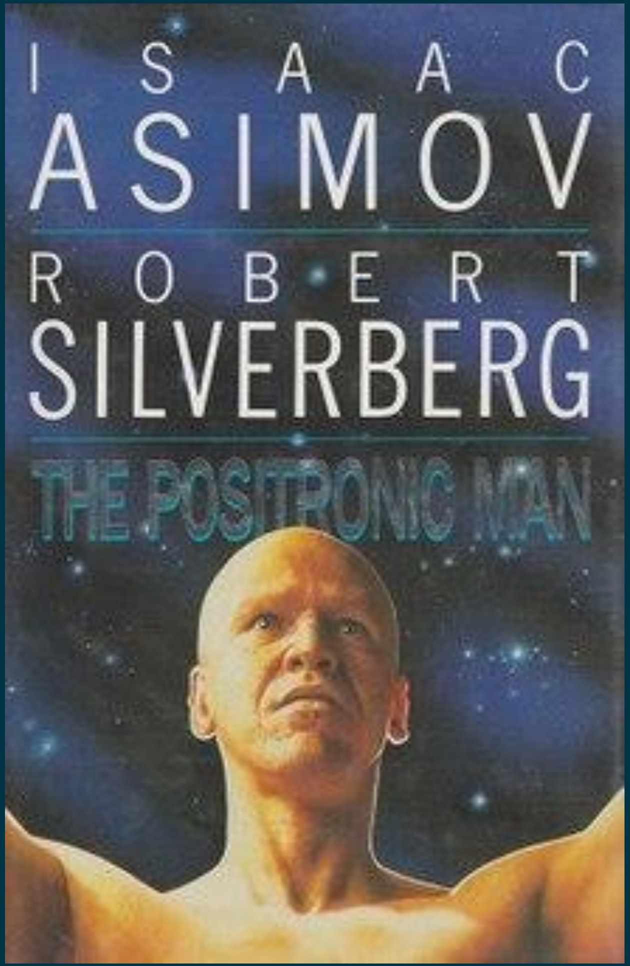
1942/1950



1969


PHILIP K. DICK
DO ANDROIDS DREAM OF ELECTRIC SHEEP?
THE INSPIRATION FOR BLADE RUNNER
A GRAPHIC NOVEL
TONY PARKER
BLOND
RICHARD STARKINGS

1993



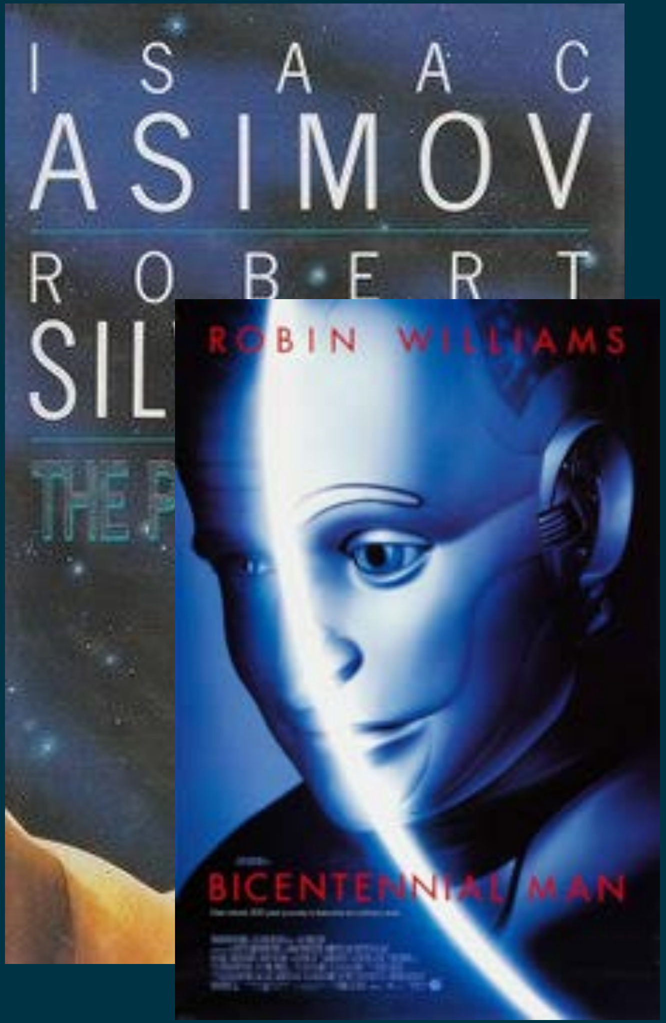
1942/1950

2004

1969

1982 2017

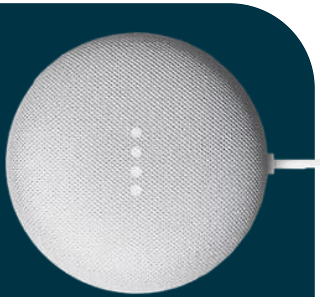
1993



1999

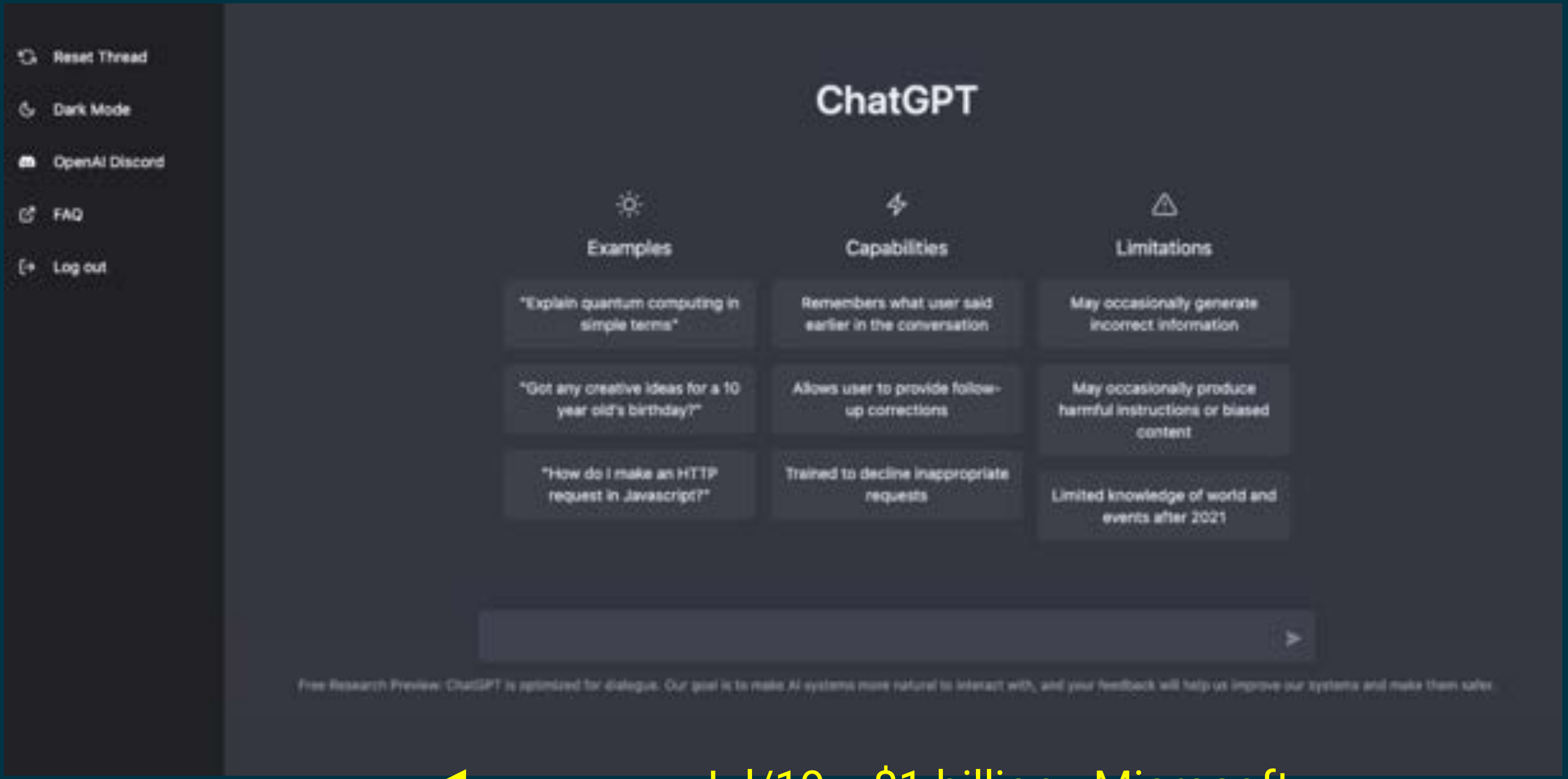
INSPIRE

Tasks -> Technology -> ‘Smart’ -> !?



What do they mean to me and to you?

Dec/15



ChatGPT

Examples	Capabilities	Limitations
"Explain quantum computing in simple terms"	Remembers what user said earlier in the conversation	May occasionally generate incorrect information
"Got any creative ideas for a 10 year old's birthday?"	Allows user to provide follow-up corrections	May occasionally produce harmful instructions or biased content
"How do I make an HTTP request in Javascript?"	Trained to decline inappropriate requests	Limited knowledge of world and events after 2021

Free Research Preview: ChatGPT is optimized for dialogue. Our goal is to make AI systems more natural to interact with, and your feedback will help us improve our systems and make them safer.

Jul/19 - $1 billion - Microsoft

Jan/21 - DALL-e released
Aug/21 - Codex released
Mar/22 - GPT 3 released
Sep/22 - Wisper released
Nov/22 - ChatGPT released
Mar/23 - ChatGPT with GPT 4

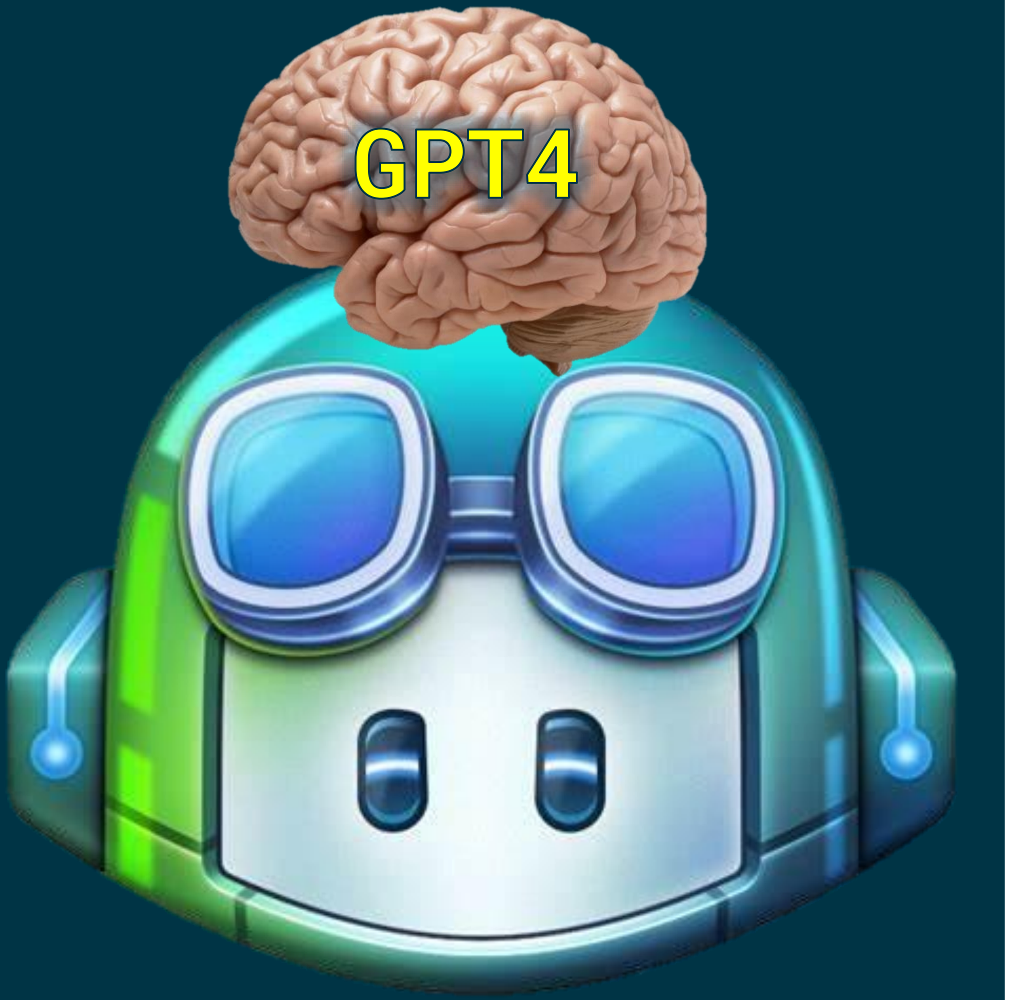
Jun/22 – Github Copilot
Feb/23 – Bing Chat AI
Jul/23 – Bing Chat AI Enterprise

Problems:
*Accuracy
*Bias
Privacy
Integration
References

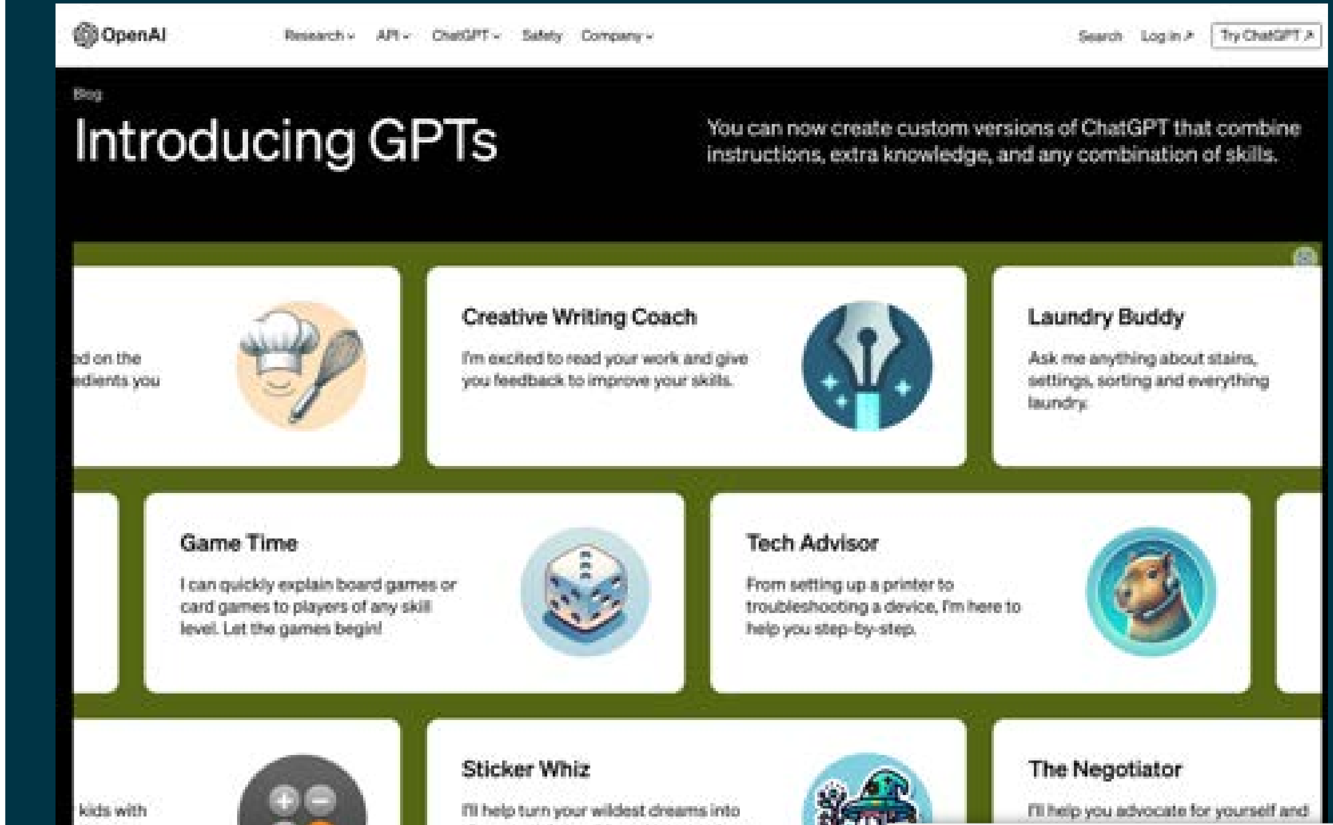
Microsoft ‘Copilot’

- Gitub Copilot (Pythom, Ruby, Go, C#, C++, KQL, ARM, HCL)
- Microsoft 365 Copilot (Teams, Outlook, Word, Excel, PowerPoint)
- Security Copilot (MS Sentinel, DefenderS)
- Windows Copilot
- Biz Apps Copilot (Dynamics, Viva)
- Power Platform* Copilot (Power BI, Power Apps, Power Automate)
- Quantum (q#) Copilot

APIs! Add-ons! Apps!



More on GPT-4



Blog

Introducing GPTs

You can now create custom versions of ChatGPT that combine instructions, extra knowledge, and any combination of skills.

…d on the …edients you

Creative Writing Coach

I'm excited to read your work and give you feedback to improve your skills.

Laundry Buddy

Ask me anything about stains, settings, sorting and everything laundry.

Game Time

I can quickly explain board games or card games to players of any skill level. Let the games begin!

Tech Advisor

From setting up a printer to troubleshooting a device, I'm here to help you step-by-step.

…kids with …sher on …re for you.

Sticker Whiz

I'll help turn your wildest dreams into die-cut stickers, shipped to your door.

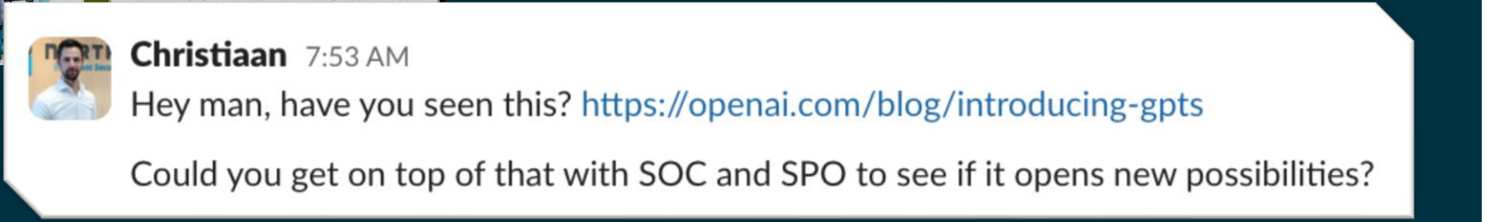
The Negotiator

I'll help you advocate for yourself and

The authors demonstrate that GPT-4 can perform novel and difficult tasks that span **mathematics, coding, vision, medicine, law, psychology and more**, without needing any special prompting.

[…] discuss the challenges and drawbacks of GPT-4, such as its lack of planning, bias, and potential for misuse.

https://openai.com/blog/introducing-gpts



Christiaan 7:53 AM

Hey man, have you seen this? https://openai.com/blog/introducing-gpts

Could you get on top of that with SOC and SPO to see if it opens new possibilities?

Sidenote

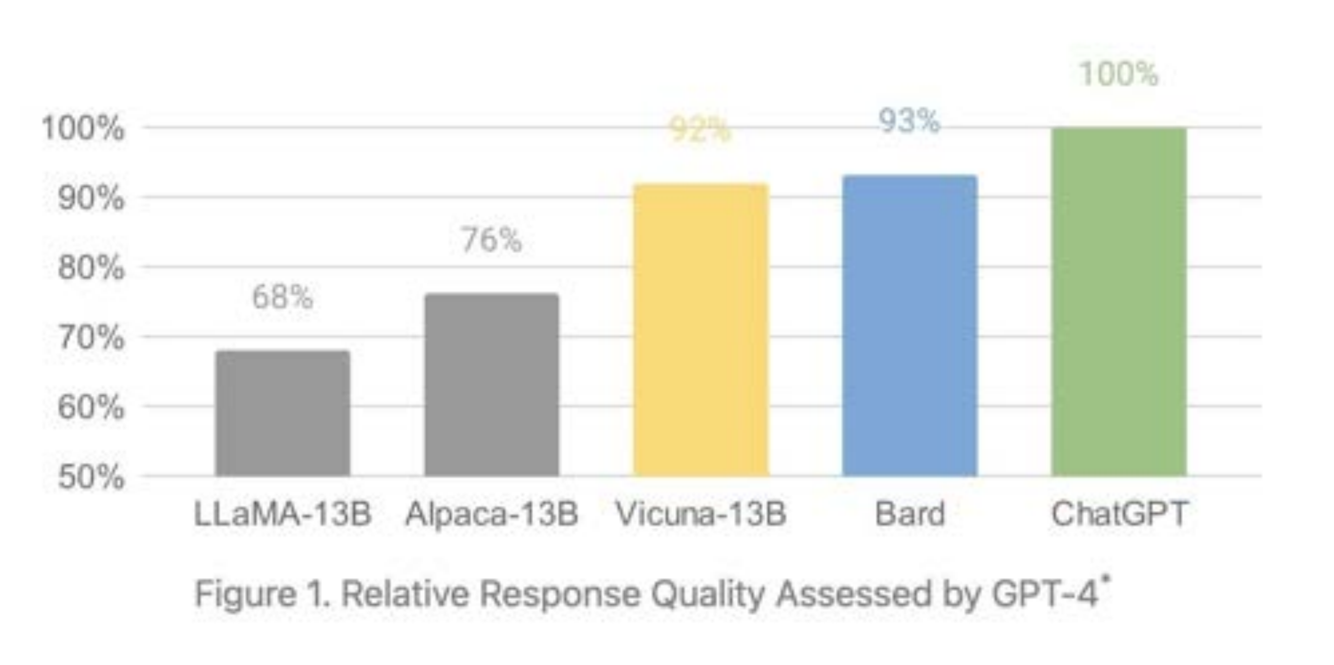
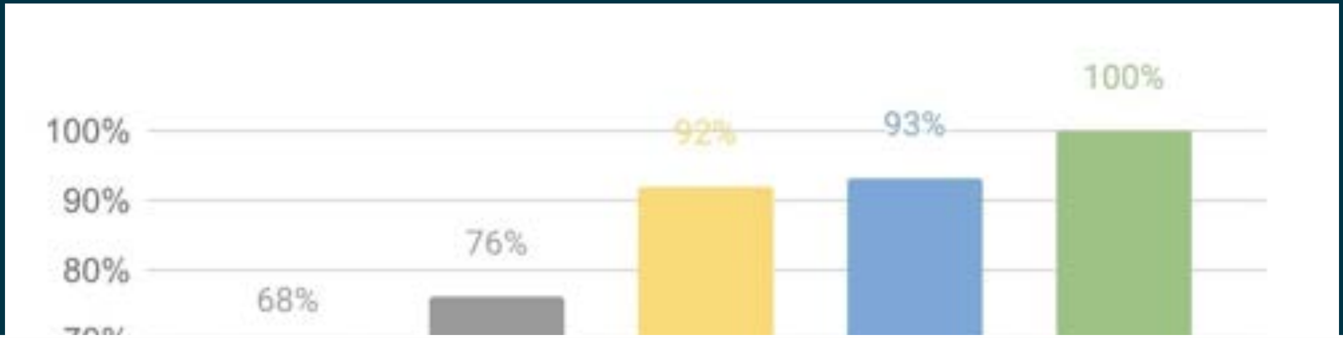
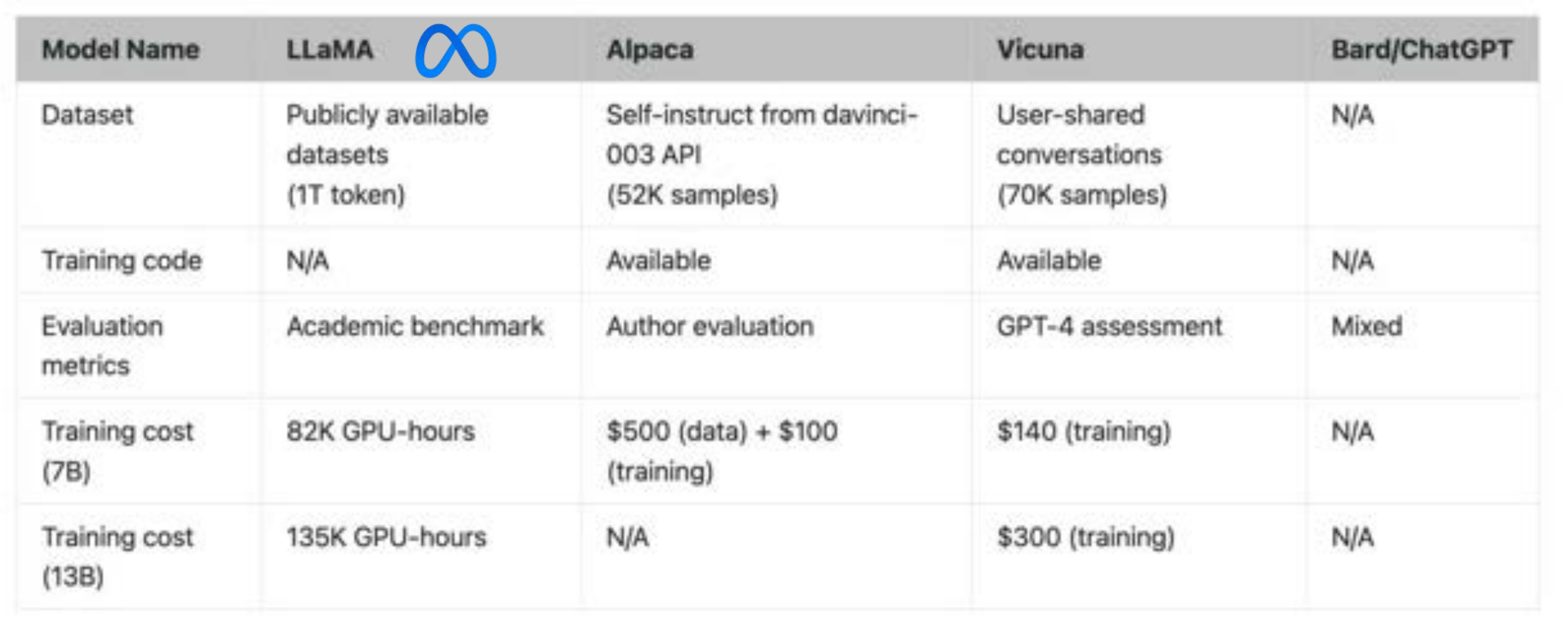


Figure 1. Relative Response Quality Assessed by GPT-4*

Open Source Vicuna Claims 90% Quality VS GPT-4 | NextBigFuture.com

Sidenote

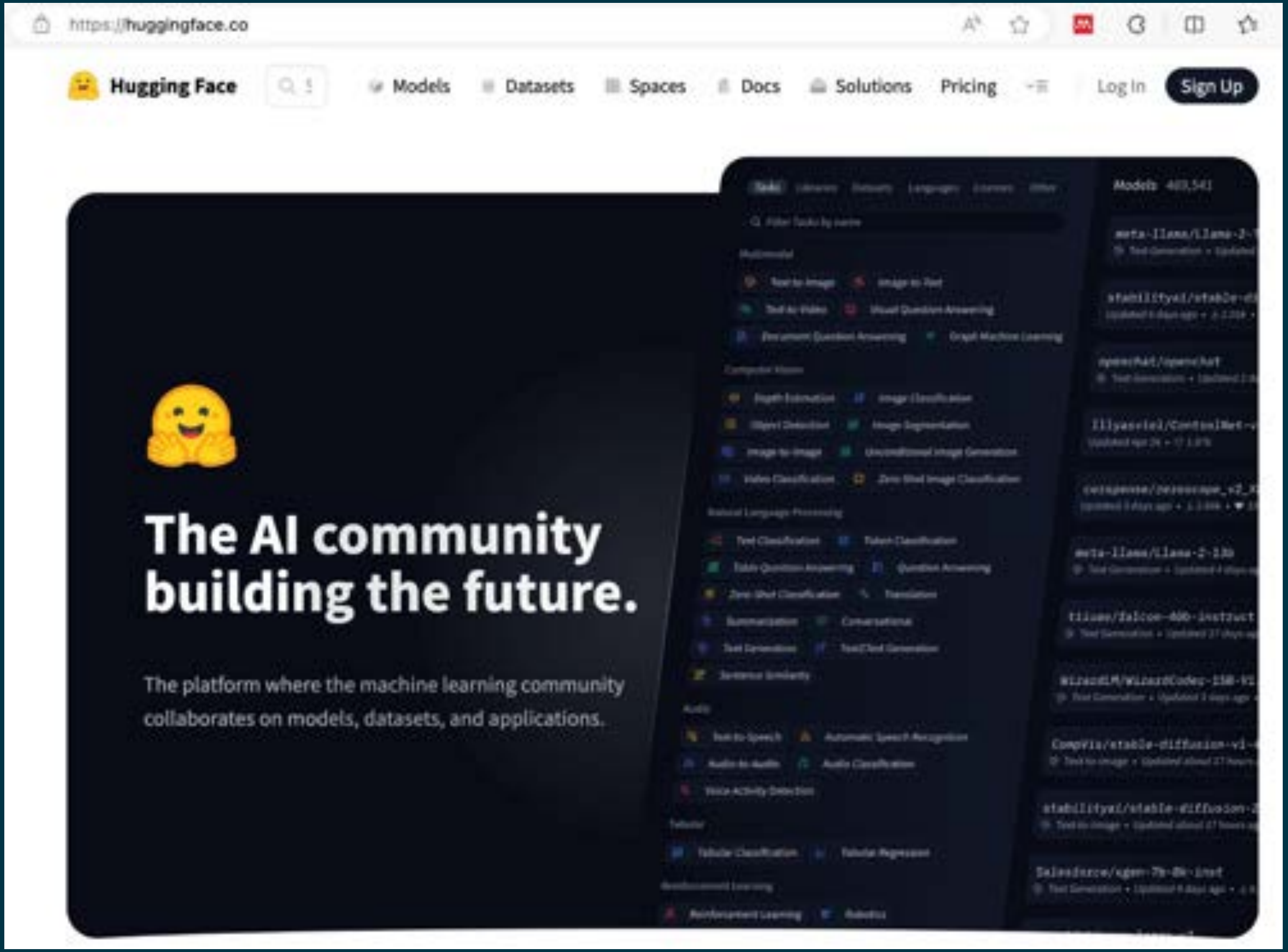




Model Name	LLaMA	Alpaca	Vicuna	Bard/ChatGPT
Dataset	Publicly available datasets (1T token)	Self-instruct from davinci-003 API (52K samples)	User-shared conversations (70K samples)	N/A
Training code	N/A	Available	Available	N/A
Evaluation metrics	Academic benchmark	Author evaluation	GPT-4 assessment	Mixed
Training cost (7B)	82K GPU-hours	$500 (data) + $100 (training)	$140 (training)	N/A
Training cost (13B)	135K GPU-hours	N/A	$300 (training)	N/A

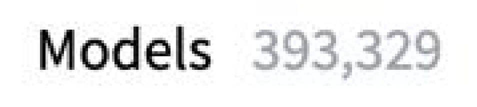
Open Source Vicuna Claims 90% Quality VS GPT-4 | NextBigFuture.com

Sidenote

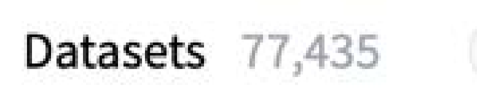


The AI community building the future.

The platform where the machine learning community collaborates on models, datasets, and applications.



Models 393,329



Datasets 77,435

What do they **should** mean to me and to you?

Simplify Human Work

AUTOMATION

&

SCALABILITY

Improve on:
- Speed
- Quality
- Safety
- Cost
- Risk

NORTHWAVE
CYBER SECURITY

[ML &] AI

What are they?

What do they mean to me and to you?

Yesterday, Today, and Tomorrow?