



Geospatial Frontiers: Navigating the Future with Generative AI and Foundational Models

PhD. Blagoj Delipetrev

AI advisor

#AIFORGGOOD

GeoAI Discovery Friday, 1 March 2024 15:30– 17:00 CET. <https://aiforgood.itu.int>

What is Generative AI



You

Please write me a 2 paragraphs poem on "Geospatial Frontiers: Navigating the Future with Generative AI and Foundational Models"



ChatGPT

Across the vast geospatial frontiers, where data meets the sky,
Generative AI and foundational models, on these we now rely.
A tapestry of maps and charts, into the future, we pry,
With algorithms so keen, beneath the digital eye.

In this realm where earth and byte seamlessly blend,
New pathways emerge, on AI we depend.
Navigating the maze of time with a digital friend,
On this journey together, towards a future we wend.



What is Generative AI

Same prompt: "ultra realistic photo, Walter White from breaking bad, standing in the middle of the road, close-up, bright realistic lighting"



Midjourney V3, July 2022 VS Midjourney V6, December 2023

<https://openai.com/sora>

Prompt: Photorealistic closeup video of two pirate ships battling each other as they sail inside a cup of coffee.

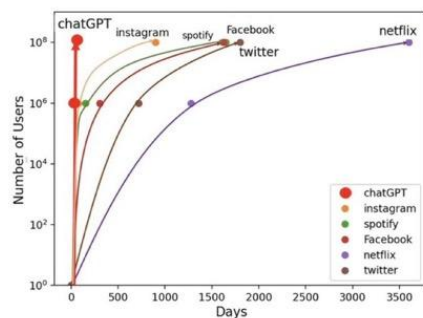


ChatGPT (November 2022)

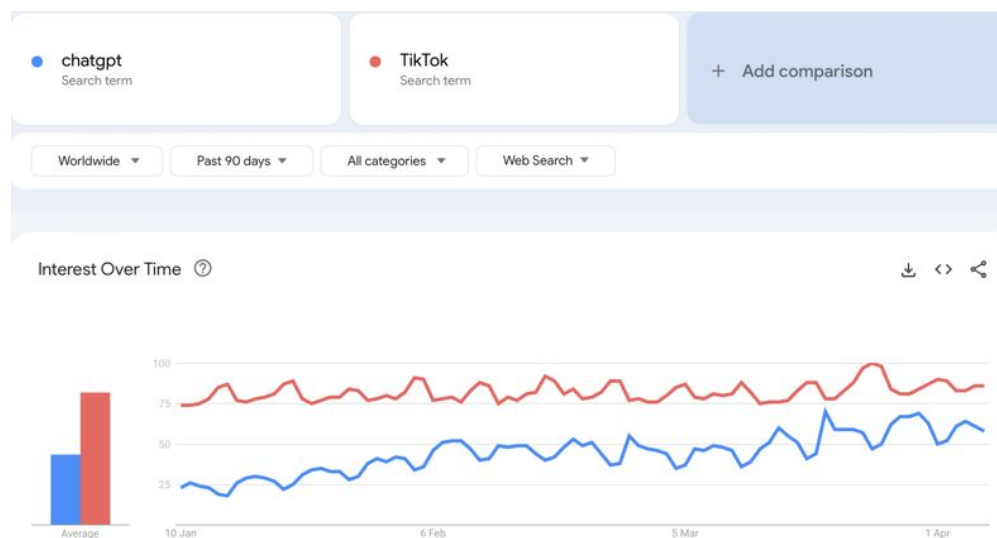


chatGPT : something different is happening.
 Number of days to 1M and 100M users :
 vs
 * instagram
 * spotify
 * facebook
 * netflix
 * twitter

#chatgpt #ai #openai #google @openai



20	reddit.com	Computers Electronics and Technology > Social Media Networks	=	00:08:39	6.19	38.07%
21	openai.com	Computers Electronics and Technology > Computers Electronics and...	▲ 10	00:05:17	6.93	27.39%
22	office.com	Computers Electronics and Technology > Programming and Developer Software	▼ 1	00:09:33	7.55	28.22%
23	netflix.com	Arts & Entertainment > Streaming & Online TV	▼ 1	00:05:33	4.33	34.95%



Generative AI landscape

(the most valued companies in the world)



Global Unicorn Club

1,205 private companies valued at \$1B+



CBINSIGHTS

Artificial intelligence



Auto & transportation



Consumer & retail



A handful of corporates were at the center of some of the highest profile AI fundraises

OpenAI	Beast round \$10B	
ANTHROPIC	Monster round up to \$4B	
Inflection	Mega round \$1.3B	NVIDIA, Microsoft, Google
cohere	Series C \$270M	NVIDIA, salesforce ventures
Hugging Face	Series D \$235M	NVIDIA, Google, salesforce ventures, intel, IBM
runway	Series C \$141M	NVIDIA, Google, salesforce ventures

stateof.ai 2023

Money needed to make best foundational models



zuck 2h · Follow

Some updates on our AI efforts. Our long term vision is to build general intelligence, open source it responsibly, and make it widely available so everyone can benefit. We're bringing our two major AI research efforts (FAIR and GenAI) closer together to support this. We're currently training our next-gen model Llama 3, and we're building massive compute infrastructure to support our future roadmap, including 350k H100s by the end of this year -- and overall almost 600k H100s equivalents of compute if you include other GPUs. Also really excited about our progress building new AI-centric computing devices like Ray Ban Meta smart glasses. Lots more to come soon.

TECHPOWERUP



HOME REVIEWS FORUMS DOWNLOADS CASE MOD GALLERY DATABASES ▾ OUR SOFTWARE ▾ MORE ▾ CONTACT US

Friday, January 19th 2024

Meta Will Acquire 350,000 H100 GPUs Worth More Than 10 Billion US Dollars

by AleksandarK | Jan 19th, 2024 15:47 | Discuss (53 Comments) 🔥

THE WALL STREET JOURNAL.

TECHNOLOGY | ARTIFICIAL INTELLIGENCE

Sam Altman Seeks **Trillions of Dollars** to Reshape Business of Chips and AI

OpenAI chief pursues investors including the U.A.E. for a project possibly requiring up to **\$7 trillion**

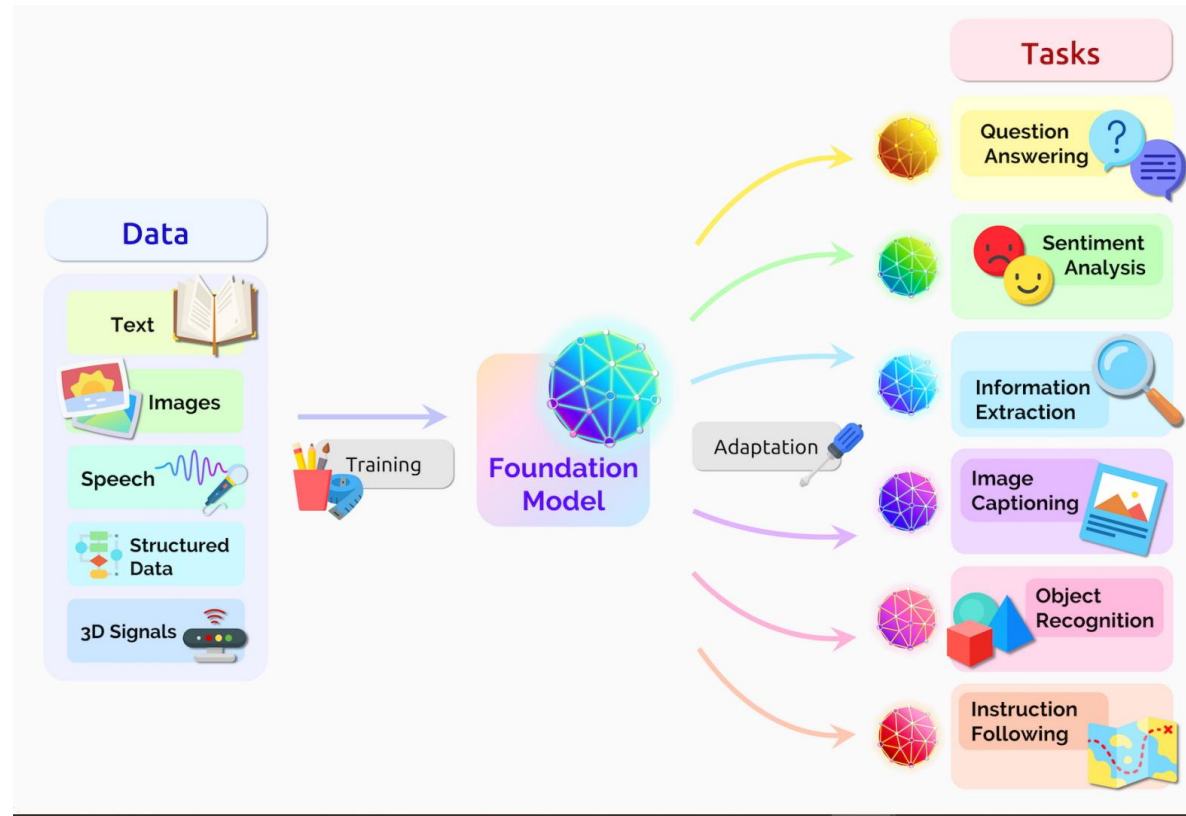
By *Keach Hagey* and *Asa Fitch*

Feb. 8, 2024 9:00 pm ET



OpenAI CEO Sam Altman wants to boost the world's chip-building capacity and expand the ability to

Foundational Multimodal Generative AI models



Google Gemini

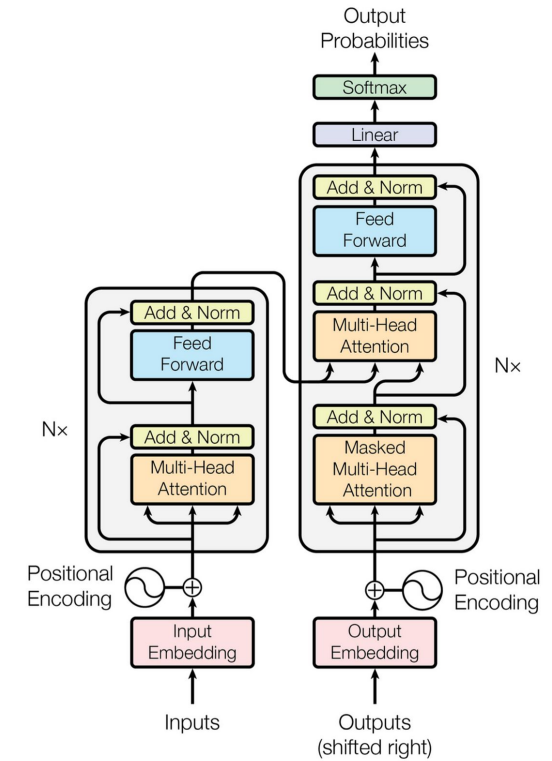
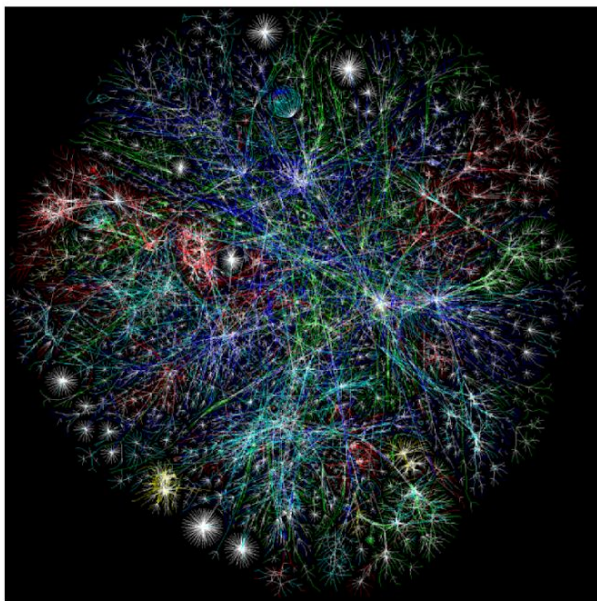


Figure 1: The Transformer - model architecture.

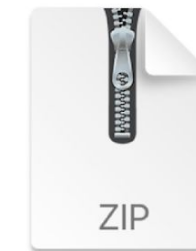
Foundational model: Llama 2 70B



Chunk of the internet,
~10TB of text



6,000 GPUs for 12 days, ~\$2M
~1e24 FLOPS



parameters.zip

~140GB file

*numbers for Llama 2 70B

Scaling laws and unsupervised learning

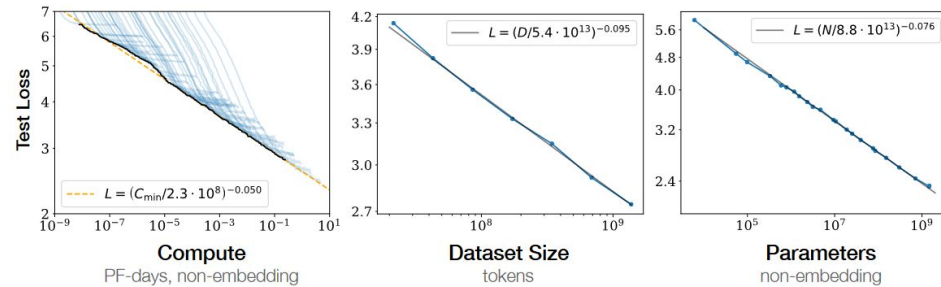


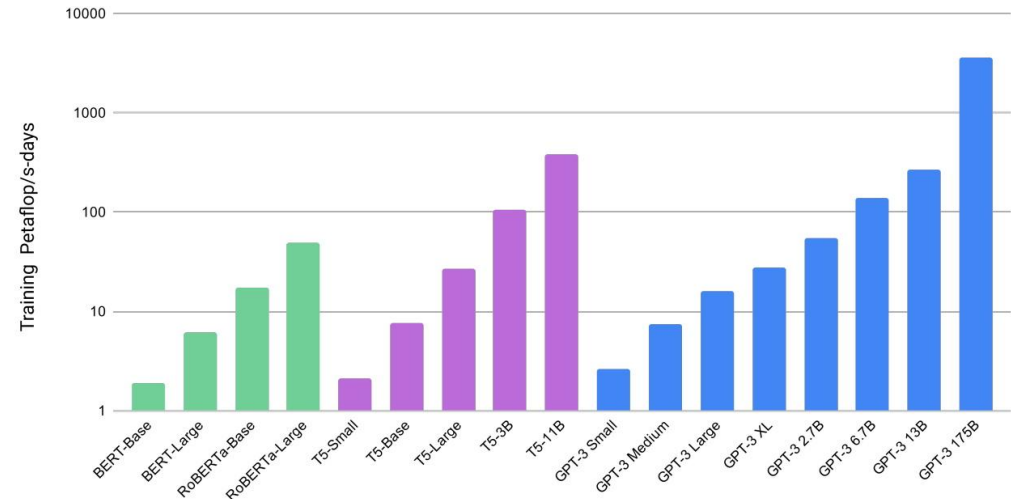
Figure 1 Language modeling performance improves smoothly as we increase the model size, dataset size, and amount of compute used for training. For optimal performance all three factors must be scaled up in tandem. Empirical performance has a power-law relationship with each individual factor when not bottlenecked by the other two.

Unsupervised learning

On the Opportunities and Risks of Foundation Models

Rishi Bommasani* Drew A. Hudson Ehsan Adeli Russ Altman Simran Arora
 Sydney von Arx Michael S. Bernstein Jeannette Bohg Antoine Bosselut Emma Brunskill
 Erik Brynjolfsson Shyamal Buch Dallas Card Rodrigo Castellon Niladri Chatterji
 Annie Chen Kathleen Creel Jared Quincy Davis Dorottya Demszky Chris Donahue
 Moussa Doumbouya Esin Durmus Stefano Ermon John Etchemendy Kawin Ethayarajh
 Li Fei-Fei Chelsea Finn Trevor Gale Lauren Gillespie Karan Goel Noah Goodman
 Shelby Grossman Neel Guha Tatsunori Hashimoto Peter Henderson John Hewitt

Total Compute Used During Training



Open source and proprietary: language, vision and multimodal foundation models

Language models

PROPRIETARY:
GPT4, Bard, Claude,
GPT3.5

OPEN SOURCE:
Llama, Mistral, BLOOM,
*many others

Vision models

PROPRIETARY:
DALL·E 3, Midjourney

OPEN SOURCE:
Stable Diffusion, Segment
anything, YOLO, Nasa-IBM
Geospatial

Multimodal models

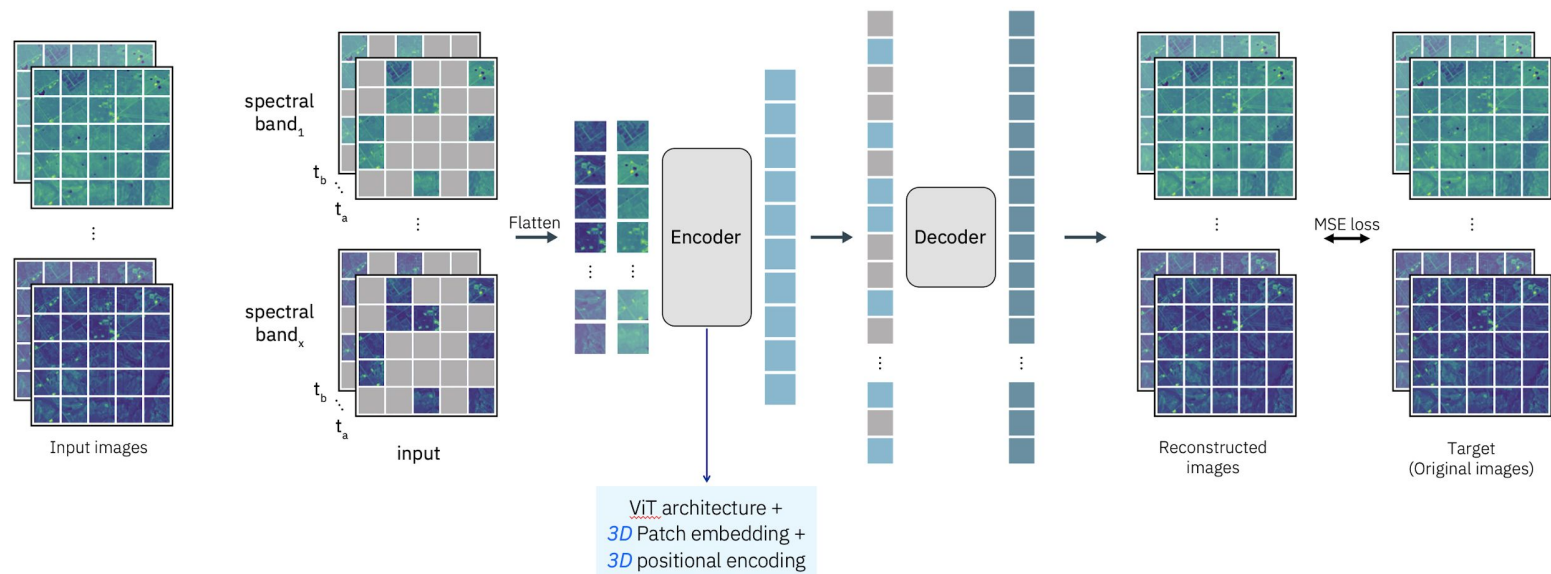
PROPRIETARY:
Google Gemini,
GPT4+ DALL·E 3

LLM Leaderboard from “Chatbot Arena”

Rank ▲	🌐 Model	★ Arena Elo ▲	📊 95% CI ▲	🗳️ Votes ▲	Organization ▲	License ▲	Knowledge Cutoff ▲
1	GPT-4-1106-preview	1254	+5/-5	38745	OpenAI	Proprietary	2023/4
2	GPT-4-0125-preview	1253	+10/-8	6308	OpenAI	Proprietary	2023/4
3	Bard...(Gemini_Pro)	1218	+8/-7	10313	Google	Proprietary	Online
4	GPT-4-0314	1191	+6/-6	20430	OpenAI	Proprietary	2021/9
5	GPT-4-0613	1164	+5/-6	32941	OpenAI	Proprietary	2021/9
6	Mistral_Medium	1152	+5/-7	17847	Mistral	Proprietary	Unknown
7	Claude-1	1150	+7/-5	19017	Anthropic	Proprietary	Unknown
8	Qwen1.5-72B-Chat	1147	+8/-8	5204	Alibaba	Qianwen LICENSE	2024/2
9	Claude-2.0	1132	+6/-8	12753	Anthropic	Proprietary	Unknown
10	Gemini_Pro...(Dev_API)	1122	+7/-7	9024	Google	Proprietary	2023/4
11	Claude-2.1	1120	+6/-4	27723	Anthropic	Proprietary	Unknown

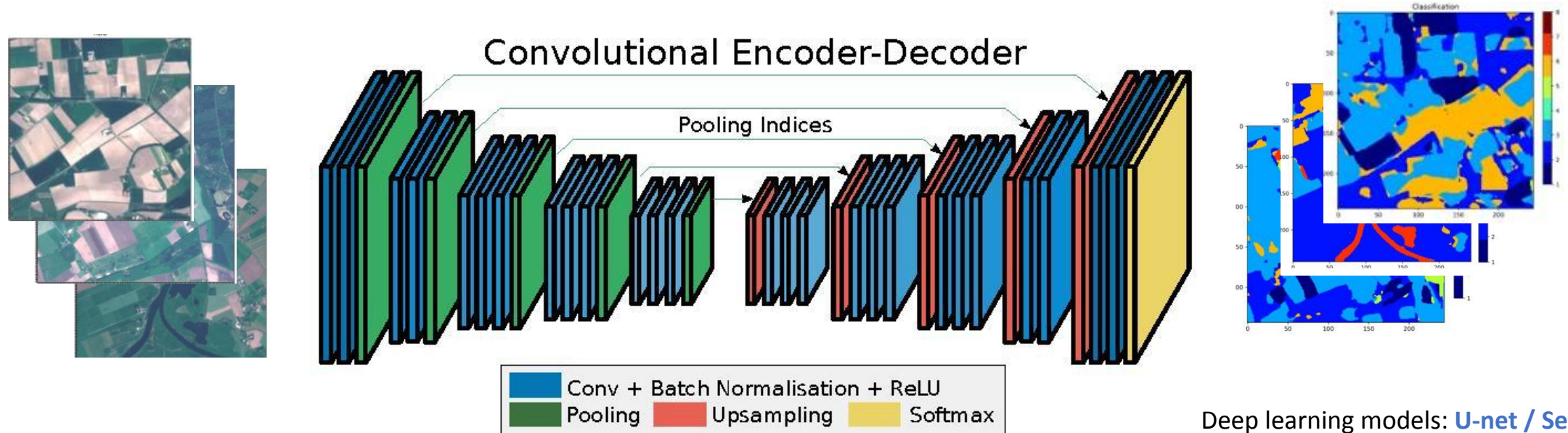
<https://huggingface.co/spaces/lmsys/chatbot-arena-leaderboard>

Nasa-IBM foundation model



- burn scars segmentation,
- flood mapping,
- multi temporal crop classification

Convolutional Neural Networks for Multi-Class Segmentation of Sentinel-2 Imagery (2019)



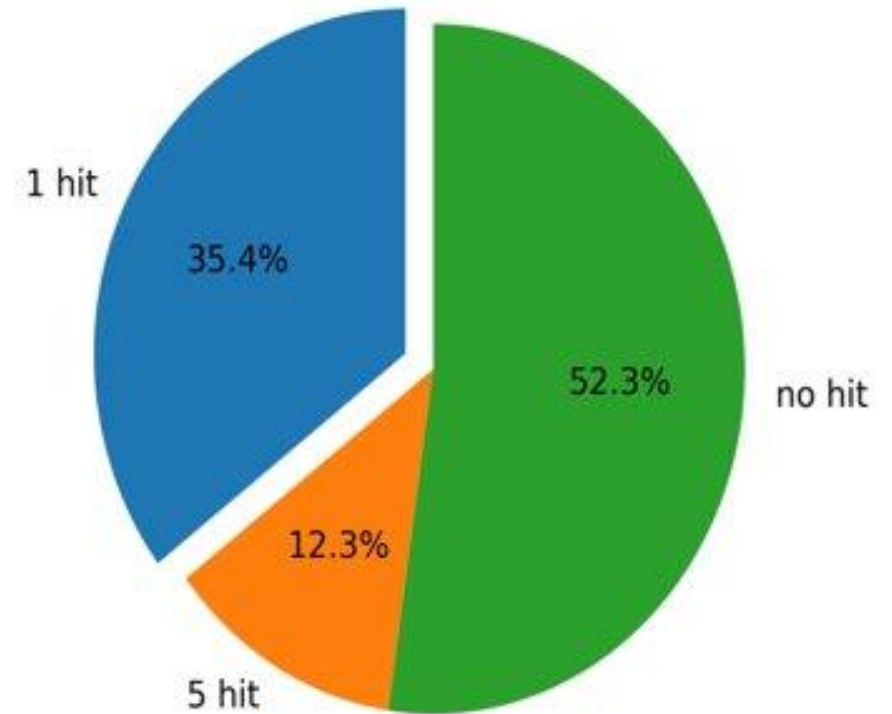
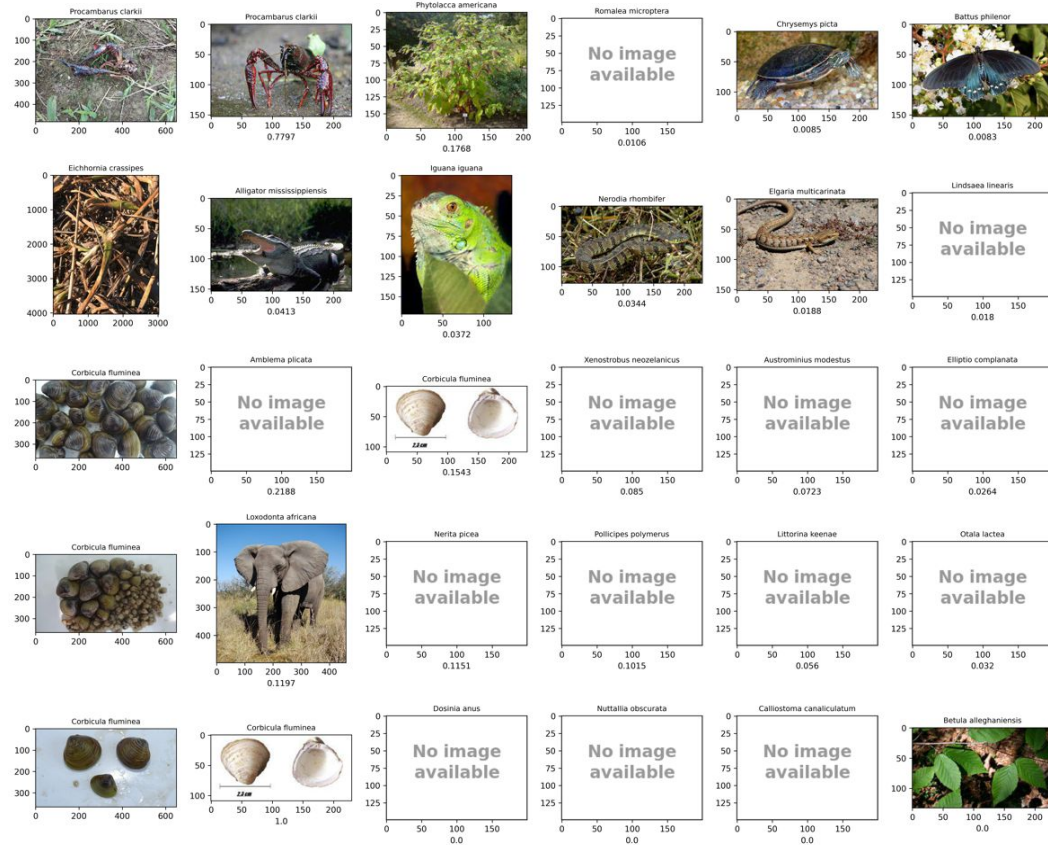
Deep learning models: **U-net / SegNet**

Parameterization:

- block size: 9x9, 122x122, 244x244, 366x366;
- Normalization method;
- Activation functions;
- Number of convolutional filters;
- Loss functions.

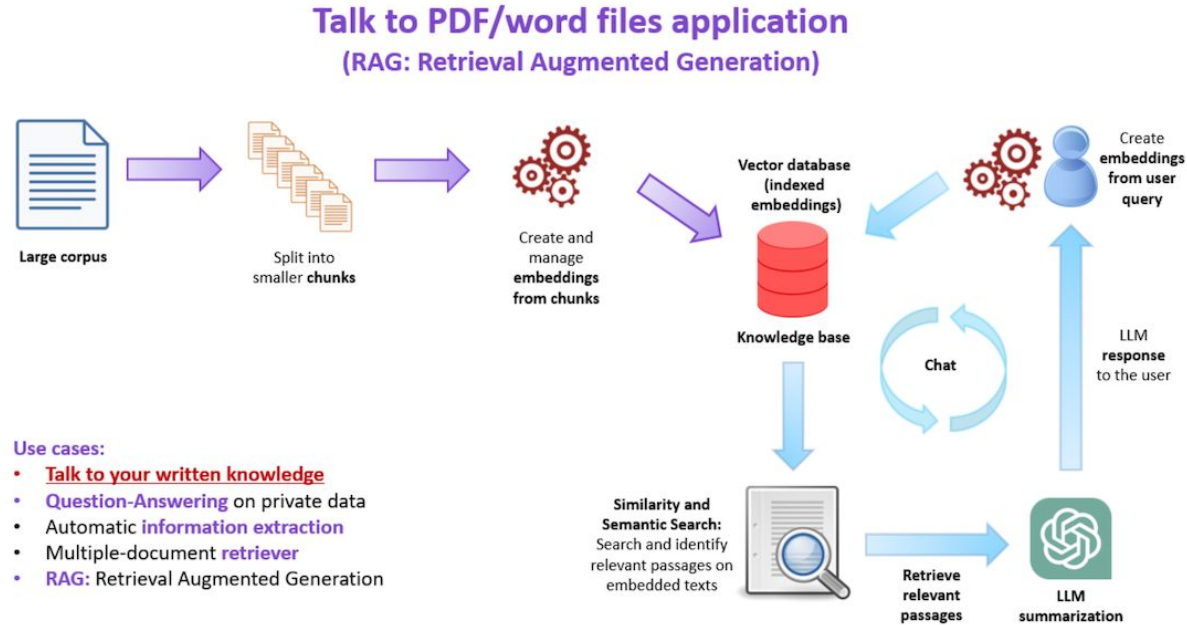


iNaturalist 2017 CV model to recognize invasive alien species (2020)



Automated invasive alien species recognition: lesson learned from applying the iNaturalist 2017 computer vision model on citizen-science data. ENVIROINFO 2020
 Authors: Blagoj Delipetrev, Sven Schade, Irena Mitton and Fabiano-Antonio Spinelli. European Commission, Joint Research Centre (JRC), Italy,

Large Language Model RAG on Documents(2023). AI interview bot (2024)



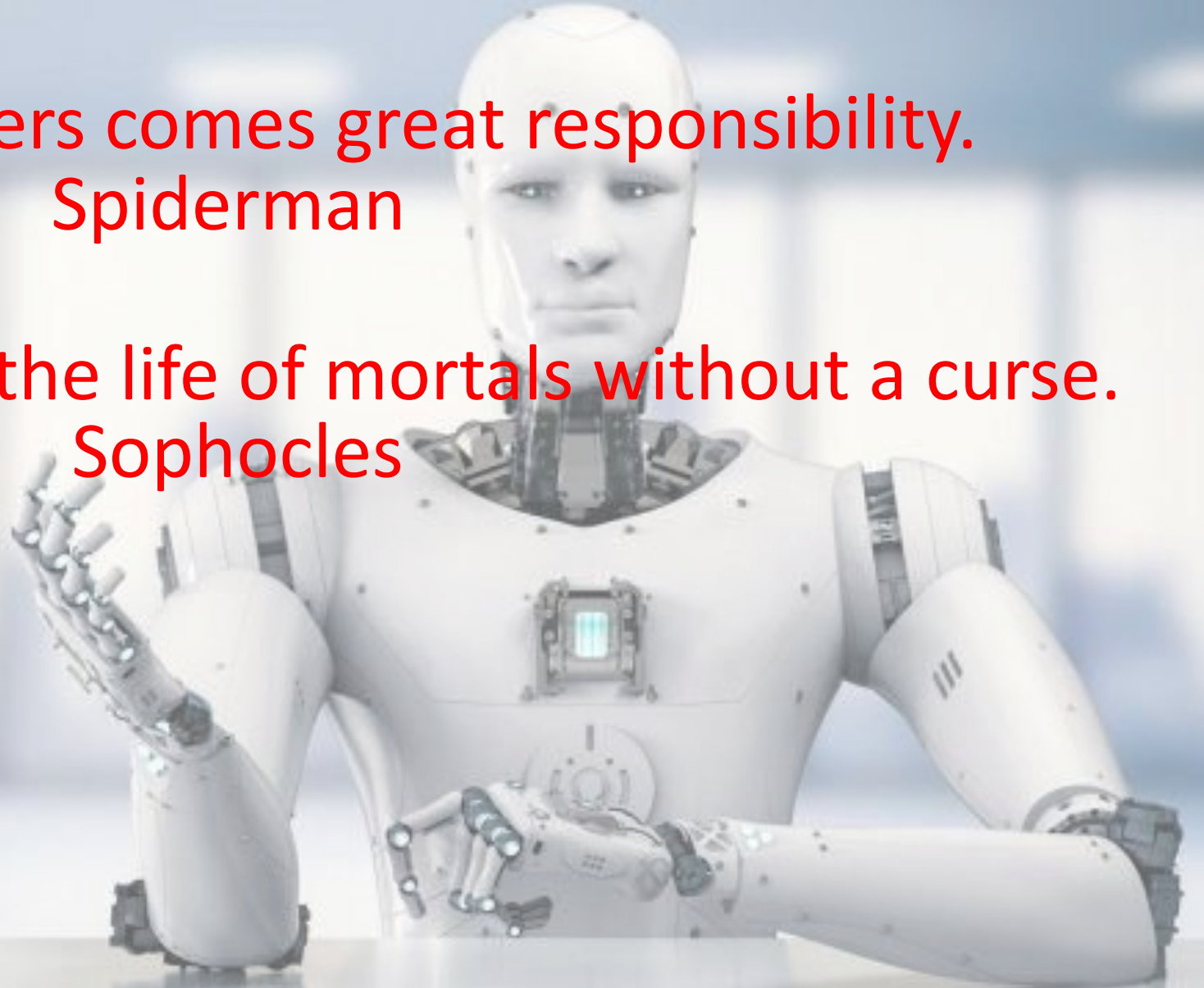
img source: HinePo



<https://blagojdelipetrev.com/blog/44/>

With great powers comes great responsibility.
Spiderman

Nothing vast enters the life of mortals without a curse.
Sophocles

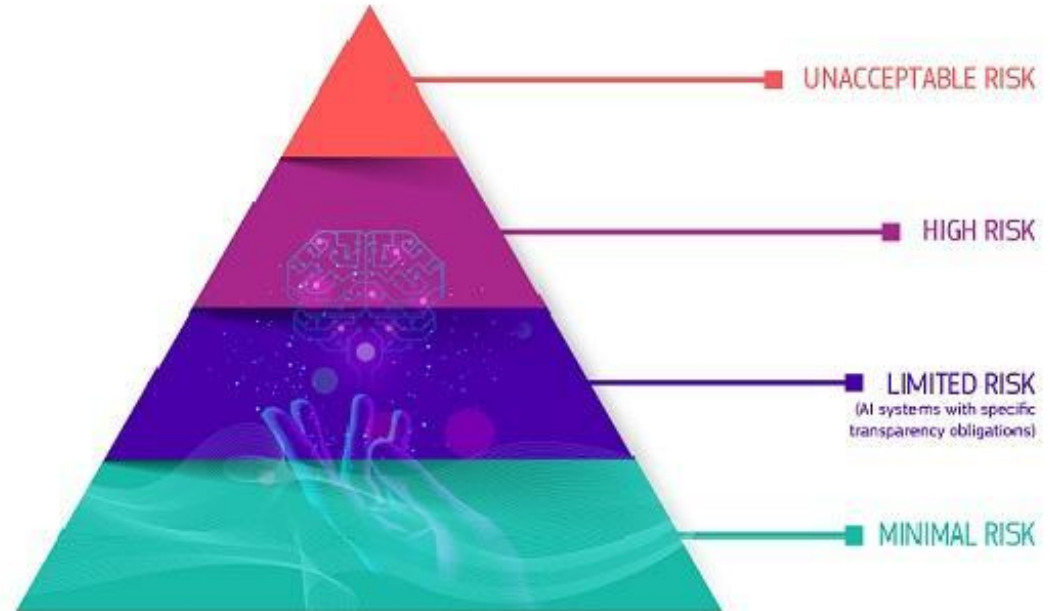


AI risks

- **Data privacy:** Issues arising from the use of personal data for training generative models e.g., location data.
- **Content authenticity:** Challenges with deepfakes and the indistinguishability of generated content from real content, e.g., synthetic landscapes or deepfake satellite image.
- **Bias and fairness:** The risk of AI inheriting biases from training data and the importance of ethical guidelines, e.g., if a geospatial AI model is trained on data from predominantly urban areas, it might not accurately represent rural or less-developed regions.
- **Regulatory challenges:** The need for legal frameworks to address novel issues brought by generative AI technologies, e.g., data sovereignty, cross-border data flows, and the use of AI in sensitive or regulated areas such as land ownership and border control.
- **Transparency:** Black box models in principle are not transparent and interpretability is an issue.



EU AI Act



What works and what won't happen

Language




AI WILL
REPLACE
PROGRAMMERS

PROGRAMMERS
WHO USE AI WILL
REPLACE
PROGRAMMERS WHO DON'T

GPT4 works amazingly

Image/video

 **Yann LeCun** · Following
VP & Chief AI Scientist at Meta
3d · 🌐

Modeling the world for action by generating pixel is as wasteful and doomed to failure as the largely-abandoned idea of "analysis by synthesis".

Decades ago, there was a big debate in ML about the relative advantages of generative methods vs discriminative methods for classification. Learning theorists, such as Vapnik, argued against generative methods, pointing out that training a generative modeling was a way more difficult than classification (from the sample complexity standpoint).

Regardless, a whole community in computer vision was arguing that recognition should work by generating pixels from explanatory latent variables. At inference time, one would infer the configuration of latent variables that generated the observed pixels. The inference method would use optimization: e.g. use a 3D model of an object and try to find the pose parameters that reproduce the image. This never quite worked, and it was very slow.

Later, some people converted to the Bayesian religion and tried to use Bayesian inference for the latent (e.g. using variational approximations and/or sampling). At some point, when Non-Parametric Bayes and Latent Dirichlet Allocation became the rage in text modeling, some folks heroically attempted to apply that to object recognition from images. >>> THIS WAS A COMPLETE AND UTTER FAILURE <<<

If your goal is to train a world model for recognition or planning, using pixel-level prediction is a terrible idea.

Generation happens to work for text because text is discrete with a finite number of symbols. Dealing with uncertainty in the prediction is easy in such settings. Dealing with prediction uncertainty in high-dimension continuous sensory inputs is simply intractable. That's why generative models for sensory inputs are doomed to failure.

Image foundational models are significantly better than any other alternative.

Doomsday AI

Media saying AI will take over the world

My Neural Network



AI will take over soon

My observation: people don't use it enough

**YEAH IF YOU COULD JUST ASK
CHATGPT INSTEAD OF ME**



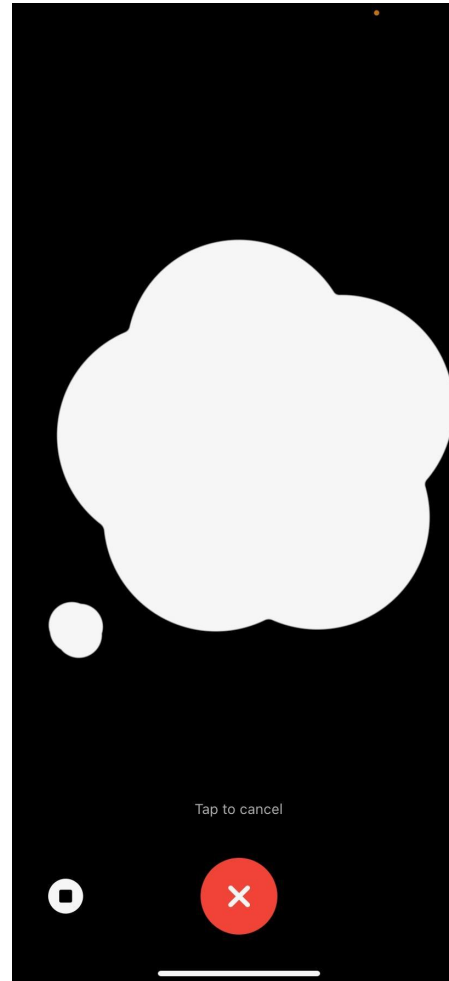
THAT WOULD BE GREAT

1 Internet year = 7 regular years
1 AI year = 7 Internet years

- **Human capital.**
- **Build up solution on**
 - API (Open AI etc.)
 - Open Source models
 - AI infrastructure
 - **Nasa – IBM geospatial foundation model**
- **LLM RAG on your data**

Predictions

- AI is the new electricity.
- **NEW Interface Voice, text, images, video. Keyboard and mouse obsolete.**
- **iPhone moment**
- **Multimodal AI**
- Internet companies made trillion \$ value. AI companies are next.
- **Several geospatial foundation models will be made in 2024/25**



David Chalmers @davidchalmers42 · Sep 20, 2022
one criticism of large language models: they only model (represent) text with no models of the world. i take it this is an empirical (and conceptual) issue: all LMs handle text, some may develop world models to do so. what's the best evidence that LLMs do/don't have world models?

89 70 441

Yann LeCun @ylecun · Sep 21, 2022
Replying to @davidchalmers42
Can an LLM figure this out:
In front of us are six gears numbered 1 to 6, mounted on axles in a row. Each gear is engaged with the next gear. If gear number 3 is rotated clockwise, in which direction will gears 1 and 6 rotate?

10 7 79

Tamay Besiroglu @tamaybes
Replying to @ylecun and @davidchalmers42
GPT-4 can:
Translate Tweet

Model: GPT-4

In front of us are six gears numbered 1 to 6, mounted on axles in a row. Each gear is engaged with the next gear. If gear number 3 is rotated clockwise, in which direction will gears 1 and 6 rotate?

If gear number 3 is rotated clockwise, we can determine the direction of rotation for the other gears by considering their engagement with each adjacent gear.

When two engaged gears rotate, they move in opposite directions. So, following this principle:

1. Gear 3 rotates clockwise.
2. Gear 2, engaged with gear 3, rotates counterclockwise.
3. Gear 1, engaged with gear 2, rotates clockwise.
4. Gear 4, engaged with gear 3, rotates counterclockwise.
5. Gear 5, engaged with gear 4, rotates clockwise.
6. Gear 6, engaged with gear 5, rotates counterclockwise.

So, gear 1 will rotate clockwise and gear 6 will rotate counterclockwise.

3:53 PM · Mar 24, 2023 · 65 Views



Thank you for your attention!

blagoj.delipetrev@gmail.com

<https://blagojdelipetrev.com>