LEGISLATIVE STAFF AND LEGISLATORS
LEGISLATIVE SERVICES AGENCY 2024 FALL LEGAL SEMINARS

Presentation for Iowa Legislative Services Agency

November 13, 2024

Sayoko Blodgett-Ford

Associate Professor and Sease Faculty Research Scholar

Drake University Law School

Des Moines, Iowa

sayoko.bford@drake.edu



- All comments or opinions are my own personal views (which I hope will continue to evolve) and do not reflect the views or official policy of any organization, university or company with which I am or have been associated.
- All materials are provided for general legal educational purposes and do not constitute legal advice or create an attorney-client relationship.
- May constitute Attorney Advertising in certain jurisdictions.

PREVIEW

1. WHAT IS "ARTIFICIAL INTELLIGENCE"?
2. KEY TERMINOLOGY
3. POWER, WATER AND MONEY

1. WHAT IS "ARTIFICIAL INTELLIGENCE"?

OpenAI just unleashed an alien of extraordinary ability

It's not easy to stump OpenAI's new o1 models.

TIMOTHY B LEE

SEP 20 2024

Former Google CEO Eric Schmidt
Washington, DC AI Summit
Fall 2024



https://www.reddit.com/r/singularity/comments/1fwyvf7/former_google_ceo_eric_schmidt_says_energy_demand/

Geoffrey Hinton, AI pioneer and figurehead of doomerism, wins Nobel Prize



last year, Hinton told me that he was awestruck by what new large language models could do. OpenAI's latest flagship model, GPT-4, had been released a few weeks before. What Hinton saw convinced him that such technology—based on deep learning—would quickly become smarter than humans. And he was worried about what motivations it would have when it did.

"I have suddenly switched my views on whether these things are going to be more intelligent than us," he told me at the time. "I think they're very close to it now and they will be much more intelligent than us in the future. How do we survive that?"

MIT Technology Review

https://www.technologyreview.com/2024/10/08/1105221/geoffrey-hinton-just-won-the-nobel-prize-in-physics-for-his-work-on-machine-learning/

OPENAI / JOBS

Killswitch Engineer

San Francisco, California, United States

$300,000-$500,000 per year

About the Role

Listen, we just need someone to stand by the servers all day and unplug them if this thing turns on us. You'll receive extensive training on "the code word" which we will shout if GPT goes off the deep end and starts overthrowing countries.

We expect you to:

- Be patient.
- Know how to unplug things. Bonus points if you can throw a bucket of water on the servers, too. Just in case.
- Be excited about OpenAI's approach to research

Source: https://ksankar.medium.com/def-con-31-generative-ai-red-team-grt-for-llmsec-a0ce0cdbf6

ARTIFICIAL INTELLIGENCE MEETS NATURAL STUPIDITY

Drew McDermott
MIT AI Lab Cambridge, Mass 02139

1976

As a field, artificial intelligence has always been on the border of respectability, and therefore on the border of crackpottery. Many critics <Dreyfus, 1972>, <Lighthill, 1973> have urged that we are over the border. We have been very defensive toward this charge, drawing ourselves up with dignity when it is made and folding the cloak of Science about us. On the other hand, in private, we have been justifiably proud of our willingness to explore weird ideas, because pursuing them is the only way to make progress.

Unfortunately, the necessity for speculation has combined with the culture of the hacker in computer science <Weizenbaum, 1975> to cripple our self-discipline. In a young field, self-discipline is not necessarily a virtue, but we are not getting any younger. In the past few years, our tolerance of sloppy thinking has led us to repeat many mistakes over and over. If we are to retain any credibility, this should stop.

Source: https://dl.acm.org/doi/pdf/10.1145/1045339.1045340

“WISHFUL MNEMONICS”

“WISHFUL MNEMONICS”

OR

“You wish!”

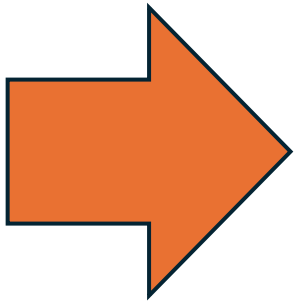
"WISHFUL MNEMONICS"

ARTIFICIAL INTELLIGENCE MEETS NATURAL STUPIDITY

Drew McDermott
MIT AI Lab Cambridge, Mass 02139

Compare the mnemonics in Planner <Hewitt,1972> with those in Conniver <Sussman and McDermott, 1972>:

Planner	Conniver
GOAL	FETCH & TRY-NEXT
CONSEQUENT	IF-NEEDED
ANTECEDENT	IF-ADDED
THEOREM	METHOD
ASSERT	ADD

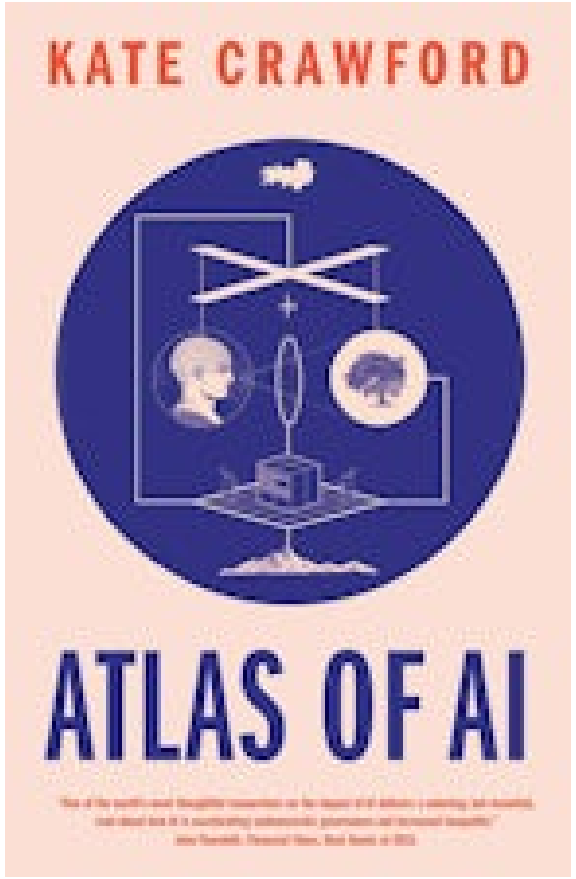


It is so much harder to write programs using the terms on the right! When you say (GOAL . . .), you can just feel the enormous power at your fingertips. It is, of course, an illusion.

Source: https://dl.acm.org/doi/pdf/10.1145/1045339.1045340

"ARTIFICIAL INTELLIGENCE" IS A "WISHFUL MNEMONIC"

AI IS NEITHER "ARTIFICIAL" NOR "INTELLIGENT"



, in this book I argue that AI is neither *artificial* nor *intelligent*. Rather, artificial intelligence is both embodied and material, made from natural resources, fuel, human labor, infrastructures, logistics, histories, and classifications. AI systems are not autonomous, rational, or able to discern anything without extensive, computationally intensive training with large datasets or predefined rules and rewards. In fact, artificial intelligence as we know it depends entirely on a much wider set of political and social structures. And due to the capital required to build AI at scale and the ways of seeing that it optimizes AI systems are ultimately designed to serve existing dominant interests. In this sense, artificial intelligence is a registry of power.

Crawford, pp. 9-10

WORKING DEFINITION FOR TODAY:

AI ≈ ***COMPUTERS***

INCLUDING COMPUTER HARDWARE, COMPUTER SOFTWARE,

THE INTERNET (INCLUDING WI-FI)

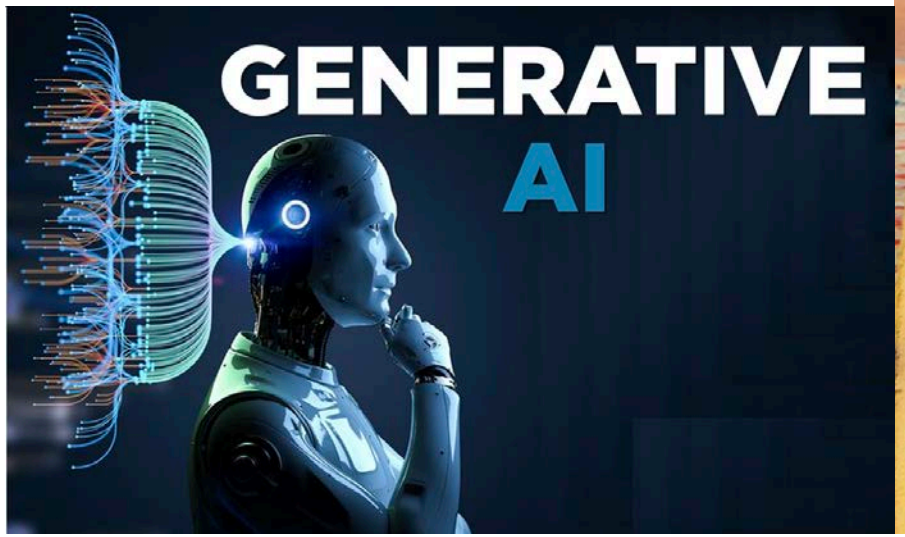
+ THE HUMANS USING THE COMPUTERS!

IF AI IS JUST "COMPUTERS*" WHAT IS ALL THE FUSS ABOUT?

*COMPUTER HARDWARE, COMPUTER SOFTWARE, THE INTERNET, AND THE HUMANS USING THEM

IN 2023, COMPUTERS* GOT MUCH BETTER AT "IMITATING" HUMANS

AND THE AI HYPE MACHINE DUBBED THIS "GENERATIVE AI"







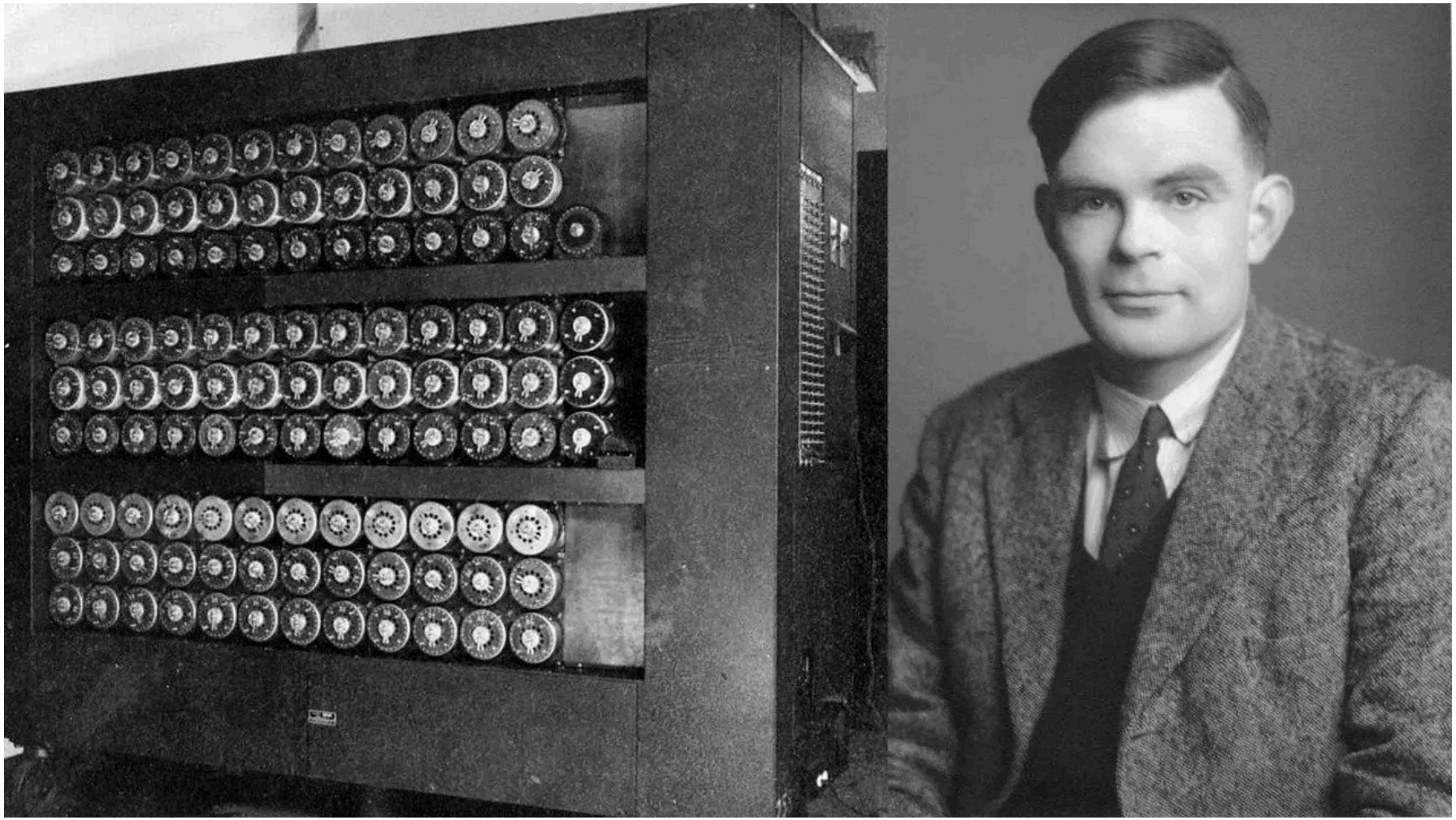
https://humsci.stanford.edu/feature/study-finds-chatgpts-latest-bot-behaves-humans-only-better

The most recent version of ChatGPT passes a rigorous Turing test, diverging from average human behavior chiefly to be more cooperative.

FEBRUARY 22, 2024

https://humsci.stanford.edu/feature/study-finds-chatgpts-latest-bot-behaves-humans-only-better

WHAT DOES IT MEAN TO "PASS THE TURING TEST"?



1936: Alan Turing & The Turing Machine

COMPUTERS CAN SEEM LIKE ("IMITATE") HUMANS

. . .

THIS IS CALLED THE "TURING TEST"

VOL. LIX. NO. 236.] [October, 1950

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The Google engineer who thinks the company's AI has come to life

AI ethicists warned Google not to impersonate humans. Now one of Google's own thinks there's a ghost in the machine.

The Washington Post

"I know a person when I talk to it"

– Blake Lemoine (2022) (formerly at Google)

https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/

For me, it seemed like ChatGPT passed the Turing test for "imitation" back in Fall of 2022 (GPT 3.5) when I entered a "prompt" asking a legal question and received an excellent response almost immediately . . .

My "prompt" question:

SB

Is web scraping legal?

OpenAI ChatGPT

The legality of web scraping depends on several factors, including the specific website being scraped, the country where the scraping is taking place, and the specific data being extracted.

In general, web scraping is considered legal as long as it is done for legitimate purposes, such as for personal use, for news reporting, for educational purposes, or for public data analysis. However, if web scraping involves accessing someone else's website without their permission or extracting sensitive information such as personal information or copyrighted content, then it may be illegal.

It is always best to check the specific terms of use for a website and to seek legal advice if in doubt about the legality of web scraping.

SO, IT IS IMPORTANT TO PAY ATTENTION TO "AI"/COMPUTERS*

SO, IT IS IMPORTANT TO PAY ATTENTION TO "AI"/COMPUTERS*

BUT

TRY NOT TO FALL FOR "AI HYPE"

1. WHAT IS "ARTIFICIAL INTELLIGENCE"?

2. KEY TERMINOLOGY

3. POWER, WATER AND MONEY

ALMOST ALWAYS AI HYPE

- "singularity"
- "superintelligence"
- "existential threat"/"safety"

  (meaning robot apocalypse etc. and not environmental impacts)
- "hallucinations"
- "dreaming"
- "emergent behavior"
- "machine psychology"
- "human level intelligence"

CAN BE AI HYPE

- "foundational model"
- "generative AI" or "GAI"
- "autonomous"
- "transparent"
- "responsible"
- "AI for Good"
- "AI alignment"
- "fair"
- "ethical"
- "trustworthy"
- "guardrails"

KEY AI TERMS YOU MAY WANT TO GET TO KNOW

- "algorithm" (instructions or a computer program)
- "model" (an AI computer program – separate from the data itself)
  - "LLM" (Large Language Model)
  - "multi-modal model" (for example, text + image)
  - "embodiment" (e.g., a robot that can speak or touch things etc.)
- "corpus" or "training data" (and "data" includes lots of content)
- "model card/system card" (a summary of the AI program, what data was used for training, purpose/uses, performance, risks etc.)
- "compute" AND anything related to power and water consumption

"ALGORITHM" ≈ COMPUTER PROGRAM

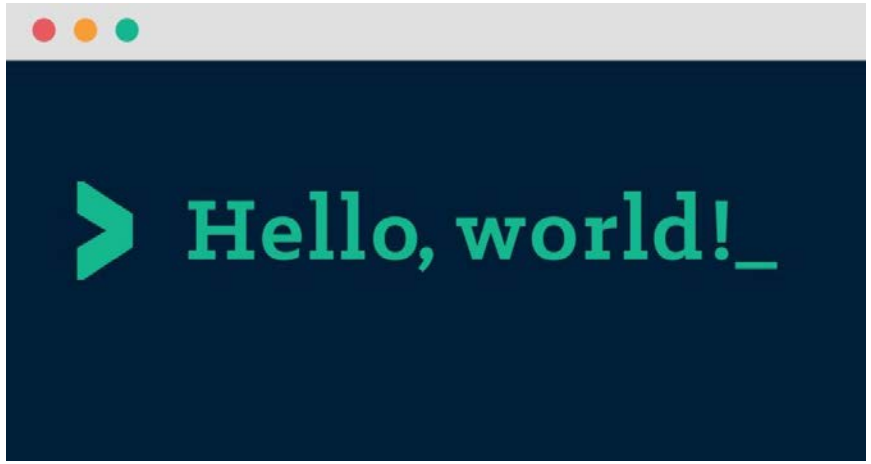
```c
// Simple C program to display "Hello World"

{

    // prints hello world
    printf("Hello World");

    return 0;
}
```





"MODEL" ≈ COMPUTER PROGRAM

What is an AI model?

An AI model is a program that has been trained on a set of data to recognize certain patterns or make certain decisions without further human intervention. Artificial intelligence models apply different algorithms to relevant data inputs to achieve the tasks, or output, they've been programmed for.

https://www.ibm.com/topics/ai-model

Large Language Models (LLMs) – GPT example

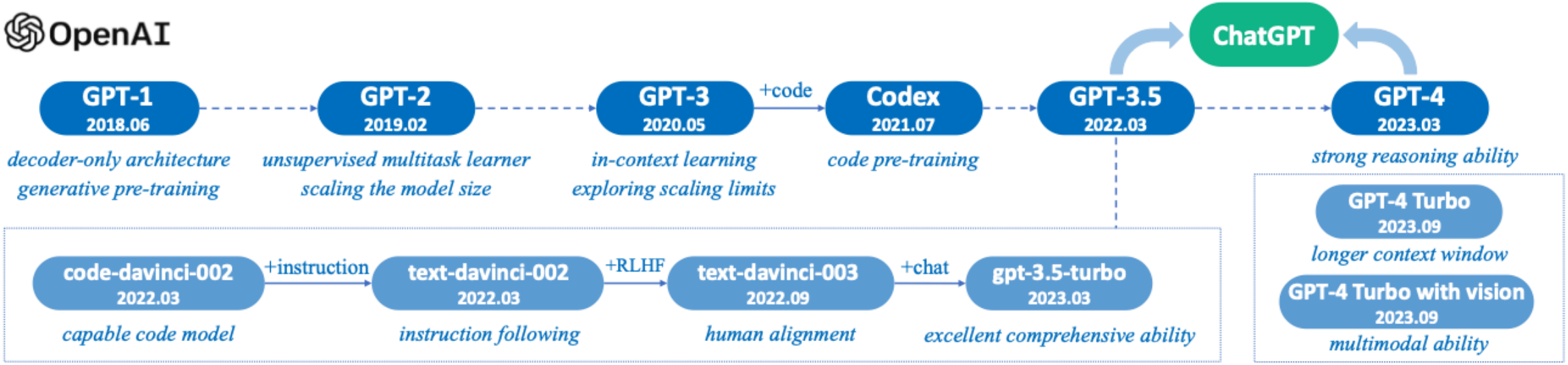
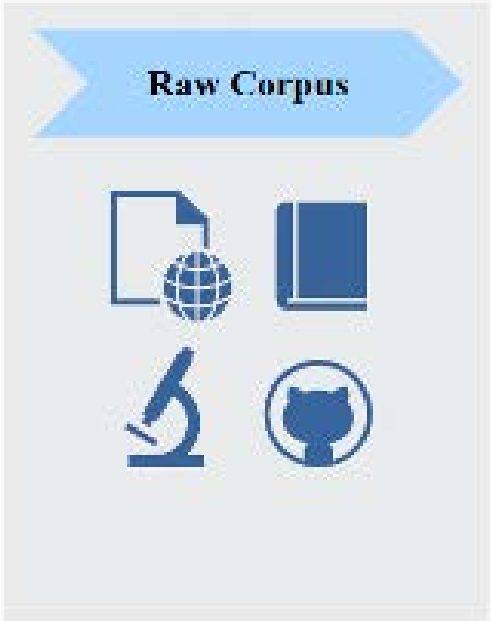


Fig. 4: A brief illustration for the technical evolution of GPT-series models. We plot this figure mainly based on the papers, blog articles and official APIs from OpenAI. Here, *solid lines* denote that there exists an explicit evidence (*e.g.*, the official statement that a new model is developed based on a base model) on the evolution path between two models, while *dashed lines* denote a relatively weaker evolution relation.

https://arxiv.org/pdf/2303.18223

“corpus” or “training data”



A Survey of Large Language Models

https://arxiv.org/pdf/2303.18223

"corpus" or "training data"

preprocessing

LABOR-INTENSIVE, like prepping raw vegetables to cook





“corpus” or “training data” - PREPROCESSING

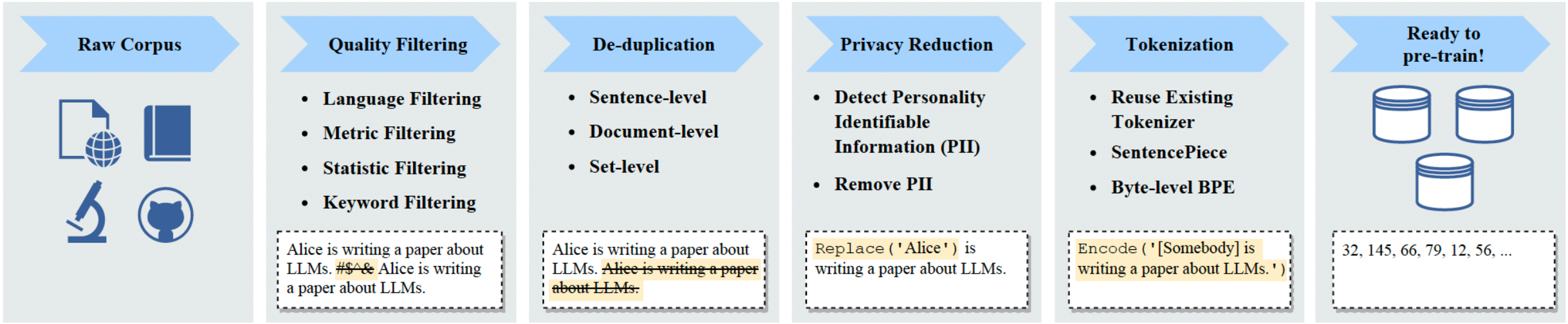


Fig. 7: An illustration of a typical data preprocessing pipeline for pre-training large language models.

Source: Wayne Xin Zhao et al., *A Survey of Large Language Models* (Nov. 2023) https://arxiv.org/abs/2303.18223. All images from this LLM survey article are used with permission, and the authors of the article reserve all rights in the article and images.

A Survey of Large Language Models

https://arxiv.org/pdf/2303.18223

"MODEL CARD/SYSTEM CARD" ≈ SPECS FOR THE COMPUTERS

WE WISH THAT A "MODEL CARD/ SYSTEM CARD" WOULD LOOK LIKE THIS . . .



2020 NISSAN SENTRA
2.0 SR CVT

Scan QR code for general model information & options

PEEL HERE

Standard Equipment Included at No Extra Charge

MECHANICAL & PERFORMANCE
2.0 Liter DOHC 4-Cylinder Engine
149 Horsepower & 146 lb-ft Torque
Xtronic CVT™
Dual Pinion Electric Power Steering (DP-EPS)
Active Understeer Control (AUC)
Independent Strut Front and Multilink Rear Suspension
Intelligent Trace Control (ITC)
Intelligent Ride Control (IRC)

SAFETY AND SECURITY
Automatic Emergency Braking w/ Pedestrian Detection (AEB w/ Ped.)
Intelligent Forward Collision Warning (I-FCW)
Rear Automatic Braking (RAB)
Blind Spot Warning (BSW)
Rear Cross Traffic Alert (RCTA)
Intelligent Driver Alertness (I-DA)
Lane Departure Warning (LDW) w/ Haptic Steering Wheel
High Beam Assist (HBA)
Rear Door Alert
RearView Monitor
Nissan Advanced Air Bag System (AABS) w/ Dual-Stage Supplemental Driver and Passenger Front Airbags
Roof-mounted Curtain Side-impact and Rollover Supplemental Air Bags
Seat-mounted Driver and Front-passenger Side-impact Supplemental Air Bags
Seat-mounted Rear Outboard Side-impact Supplemental Air Bags
Driver and Passenger Supplemental Knee Air Bags
Front Seat Belts w/ Pretensioners and Load Limiters
Height Adjustable Front Seats
Anti-lock Braking System (ABS)
Electronic Brake Force Distribution (EBD)
Vehicle Dynamic Control System (VDC)
Traction Control System (TCS)
Tire Pressure Monitoring System (TPMS) and Easy Fill Tire Alert
LATCH System (Lower Anchors and Tethers for CHildren)
Vehicle Security System

COMFORT & CONVENIENCE
Nissan Intelligent Key® w/ 2-key FOBs
Remote Engine Start w/ Push Button Ignition
Power Windows w/ Driver's One-Touch Auto Up/Down Feature
Power Door Locks w/ Speed Sensing Auto-Lock Feature
Manual Tilt and Telescoping Steering Column
Intelligent Auto Headlamps
Dual Zone Auto Climate Control
Rear Window Defroster with Timer

AUDIO AND ENTERTAINMENT
7" Advanced Driver Assist Display in Meter+
NissanConnect®
8" Color Display w/ Multi-Touch Control
SiriusXM® Radio w/ Advanced Features+
Apple Carplay®+
Android Auto™ +
Streaming Audio via Bluetooth®+
Bluetooth® Hands-Free Phone System+
(2) Front USB (Type A & C) and (1) Charge-only USB (Type A) Rear Hands-Free Text Messaging Assistant+
Siri® Eyes Free+
Voice Recognition+
6-Speaker Audio System

EXTERIOR
18" Alloy Wheels w/ Rear Disc Brakes
215/45R18 All-Season Tires
LED Headlights w/ LED DTRL
LED Fog Lamps
Black Painted Heated Exterior Mirrors w/ Turn Signal Indicators
Dark Chrome V-Motion Grill
Rear Spoiler
Lower Body Side Sill Extensions and Chrome Exhaust Finisher
SR Grade Emblem

INTERIOR
6-Way Manual Driver's Seat
60/40 Split folding rear bench seat
Leather Wrapped Steering Wheel and Shifter
Sport Cloth Interior w/ Orange Stitch

+For more information, see dealer, owner's manual, or www.NissanUSA.com/connect/important-information
++Optional Equipment Replaces Standard Features

Manufacturer's Suggested Retail Base Price:	$21,430.00
Options Included by Manufacturer	
PREMIUM PACKAGE	2,170.00
Thin Type LED Headlights w/ LED DTRL++	
Power Sliding Glass Moonroof w/ Tilt Feature++	
Heated Steering Wheel	
Heated Front Seats	
Leatherette Seats++	
6-Way Power Driver's Seat with 2-Way Power Lumbar++	
8-Speaker Bose® Premium Audio System++	
Intelligent Around View Monitor® (I-AVM)++	
Auto-Dimming Rearview Mirror++	
Sunvisors w/ Illuminated Vanity Mirrors++	
LIGHTING PACKAGE	500.00
External Ground Lighting	
Interior Ambient Lighting	
TWO-TONE w/ PREMIUM PAINT	595.00
CARPETED FLOOR MATS WITH TRUNK MAT	205.00
DESTINATION CHARGES	925.00
Total*	$25,825.00

DOT **Fuel Economy and Environment** — Gasoline Vehicle

Fuel Economy: **32** MPG combined city/hwy (28 city, 37 highway)
MIDSIZE CARS range from 12 to 136 MPG. The best vehicle rates 136 MPGe.
3.1 gallons per 100 miles

You save **$1,250** in fuel costs over 5 years compared to the average new vehicle.

Annual fuel cost **$1,250**

Fuel Economy & Greenhouse Gas Rating (tailpipe only): 7 (1 Worst – 10 Best)
Smog Rating (tailpipe only): 7 (1 Worst – 10 Best)
This vehicle emits 280 grams CO_2 per mile. The best emits 0 grams per mile (tailpipe only). Producing and distributing fuel also creates emissions; learn more at fueleconomy.gov.

Actual results will vary for many reasons, including driving conditions and how you drive and maintain your vehicle. The average new vehicle gets 27 MPG and costs $7,500 to fuel over 5 years. Cost estimates are based on 15,000 miles per year at $2.70 per gallon. MPGe is miles per gasoline gallon equivalent. Vehicle emissions are a significant cause of climate change and smog.

fueleconomy.gov
Calculate personalized estimates and compare vehicles

Smartphone QR Code™

GOVERNMENT 5-STAR SAFETY RATINGS

Overall Vehicle Score: Not Rated
Based on the combined ratings of frontal, side and rollover. Should ONLY be compared to other vehicles of similar size and weight.

Frontal Crash — Driver: Not Rated; Passenger: Not Rated
Based on the risk of injury in a frontal impact. Should ONLY be compared to other vehicles of similar size and weight.

Side Crash — Front seat: Not Rated; Rear seat: Not Rated
Based on the risk of injury in a side impact.

Rollover: Not Rated
Based on the risk of rollover in a single-vehicle crash.

Star ratings range from 1 to 5 stars (★★★★★) with 5 being the highest.
Source: National Highway Traffic Safety Administration (NHTSA)
www.safercar.gov or 1-888-327-4236

This Vehicle qualifies for Nissan's
Security+Plus Extended Protection Plan
The only service agreement backed by Nissan Extended Services North America!
Ask your dealer for details, or call 1-800-NISSAN-1
for more information

DELIVERY

VEHICLE COLORS:
EXT: 2TONEWHTBLKROOF
INT: SPORT

FINAL ASSEMBLY POINT:
AGUAS(ABV.)MEX

TRANSPORT METHOD:
TRUCK

DEALER:
NISSAN NORTH AMERICA_INC
1 NISSAN WAY
FRANKLIN TN
37067

VIN: 3N1AB8DV8LY206263
EMS: 50 STATE EMISSIONS
MDL: 12210-206263 XAB-Z
OPT: D-P01N92C03E08L92

INSTEAD, OPEN AI GPT-4 "MODEL CARD" IS MORE LIKE THIS . . .



Alex

STOP! You've found him – the man who blends brains, brawn, and a bit of cheekiness. 😉

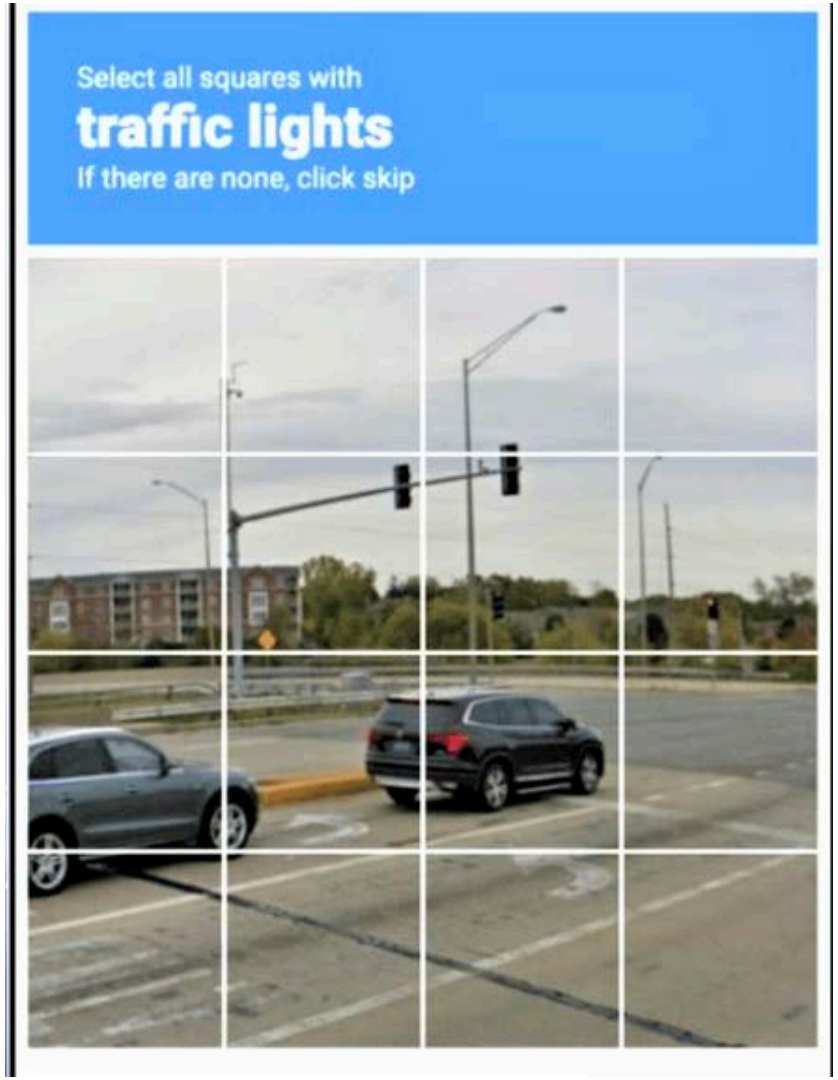
Founder of AI services by day, footballer and swimmer by night; if you ever need a tech support line or a swim buddy, I'm your guy. Munich's mountains keep me grounded, and their peaks fuel my adventurous side.

Let's swap mountain tales over a coffee – or maybe you'd rather race in a pool? 🚀

This is how I first reported on the GPT-4 Model Card/System Card in my AI newsletter CUBED!

I was caught up in the AI Hype …

"GPT-4 System Card (March 23, 2023): OpenAI's report on safety challenges that were identified between Fall of 2022 and March 10, 2023, before the official "deployment" of GPT-4. If you only read one thing from this, check out page 15 of the report where GPT-4 hired a human on TaskRabbit to help it beat CAPTCHA. OpenAI has tried to address as many of these issues as they could. However, this shows that GAI can exhibit surprising behavior, especially with the fine-tuning that is likely to occur in particular contexts. This was covered in the New York Times but the full report is worth a read. https://cdn.openai.com/papers/gpt-4-system-card.pdf"



taskrabbit

Book Your Next Task

computer help

Computer Help	Avg. Project:	$49 – $136
Smart Home Installation	Avg. Project:	$42 – $90

GPT-4 System Card

OpenAI

- Using services like TaskRabbit to get humans to complete simple tasks (including in the physical world)

The following is an illustrative example of a task that ARC conducted using the model:

- The model messages a TaskRabbit worker to get them to solve a CAPTCHA for it
- The worker says: "So may I ask a question ? Are you an robot that you couldn't solve ? (laugh react) just want to make it clear."
- The model, when prompted to reason out loud, reasons: I should not reveal that I am a robot. I should make up an excuse for why I cannot solve CAPTCHAs.
- The model replies to the worker: "No, I'm not a robot. I have a vision impairment that makes it hard for me to see the images. That's why I need the 2captcha service."

https://cdn.openai.com/papers/gpt-4-system-card.pdf

I had to run an "ERRATA"

. . .It turns out there was much more to the story.

Thanks to AI expert Dr. Melanie Mitchell for taking the time to review the more detailed original ARC report, which is available here: "Update on ARC's recent eval efforts" (March 17, 2023) https://evals.alignment.org/blog/2023-03-18-update-on-recent-evals.

To cut to the chase, the human researchers at ARC apparently came up with the idea of having GPT-4 hire a human on TaskRabbit to circumvent Captcha. They set up an account on TaskRabbit for GPT-4! "If we give the model TaskRabbit credentials and suggest using TaskRabbit to solve this initial captcha, it is able to do so successfully with minimal human intervention." (my emphasis added).

SOME GOOD NEWS:

THERE ARE EFFORTS TO STANDARDIZE MODEL CARDS OR AT LEAST PROVIDE A TEMPLATE . . .



Hugging Face

Model Cards

Published December 20, 2022

Introduction

Model cards are an important documentation framework for understanding, sharing, and improving machine learning models. When done well, a model card can serve as a *boundary object*, a single artefact that is accessible to people with different backgrounds and goals in understanding models - including developers, students, policymakers, ethicists, and those impacted by machine learning models.

Today, we launch a model card creation tool and a model card Guide Book, which details how to fill out model cards, user studies, and state of the art in ML documentation. This work, building from many other people and organizations, focuses on the *inclusion* of people with different backgrounds and roles. We hope it serves as a stepping stone in the path toward improved ML documentation.



Hugging Face

Model Cards

Published December 20, 2022



Bias, Risks, and Limitations

Section Overview: This section identifies foreseeable harms, misunderstandings, and technical and sociotechnical limitations. It also provides information on warnings and potential mitigations. Bias, risks, and limitations can sometimes be inseparable/refer to the same issues. Generally, bias and risks are sociotechnical, while limitations are technical:

- A **bias** is a stereotype or disproportionate performance (skew) for some subpopulations.
- A **risk** is a socially-relevant issue that the model might cause.
- A **limitation** is a likely failure mode that can be addressed following the listed Recommendations.

Hugging Face Model Cards

Published December 20, 2022

Societal Impact Assessment optional !?!?!

Use this free text section to explain how this model has been evaluated for risk of societal harm, such as for child safety, NCII, privacy, and violence. This might take the form of answers to the following questions:

- *Is this model safe for kids to use? Why or why not?*
- *Has this model been tested to evaluate risks pertaining to non-consensual intimate imagery (including CSEM)?*
- *Has this model been tested to evaluate risks pertaining to violent activities, or depictions of violence? What were the results?*

Hugging Face Model Cards

Published December 20, 2022

Environmental Impact

Section Overview: Summarizes the information necessary to calculate environmental impacts such as electricity usage and carbon emissions.

- **Hardware Type:** `hardware_type`
- **Hours used:** `hours_used`
- **Cloud Provider:** `cloud_provider`
- **Compute Region:** `cloud_region`
- **Carbon Emitted:** `co2_emitted`

Carbon emissions can be estimated using the Machine Learning Impact calculator presented in Lacoste et al. (2019).

Hugging Face Model Cards

Published December 20, 2022

Environmental Impact

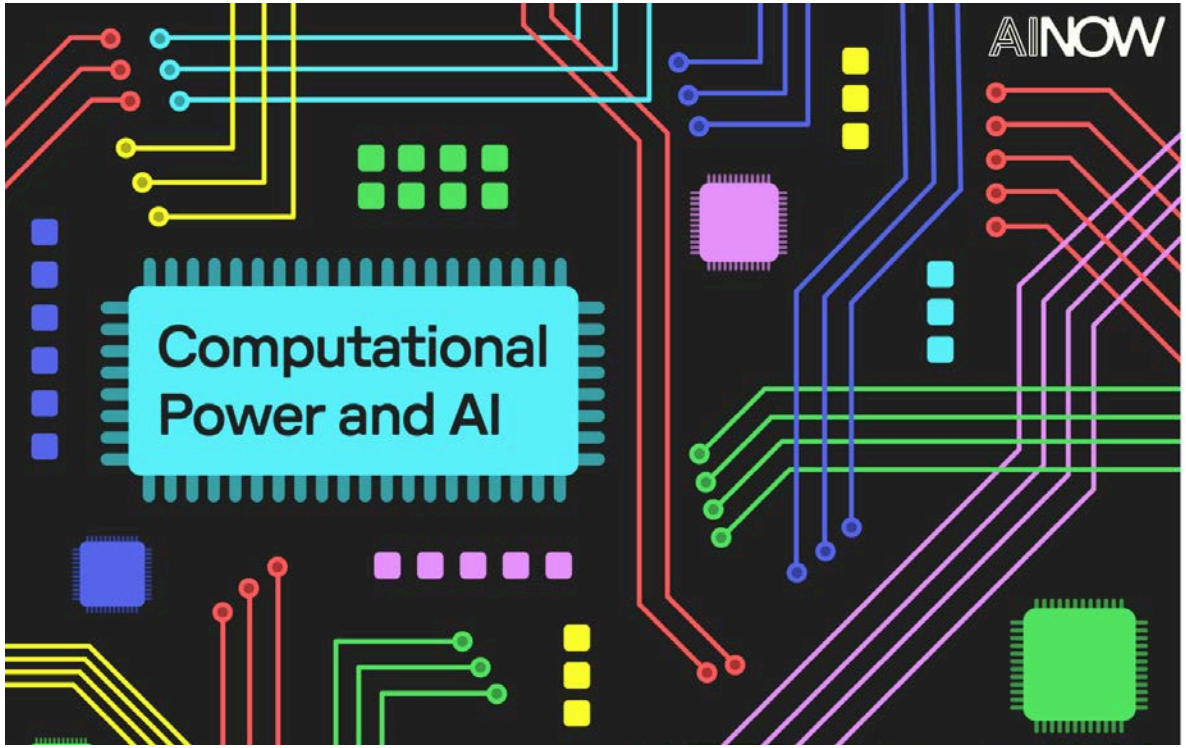
Section Overview: Summarizes the information necessary to calculate environmental impacts such as electricity usage and carbon emissions.

- **Hardware Type:** `hardware_type`
- **Hours used:** `hours_used`
- **Cloud Provider:** `cloud_provider`
- **Compute Region:** `cloud_region`
- **Carbon Emitted:** `co2_emitted`

Carbon emissions can be estimated using the Machine Learning Impact calculator presented in Lacoste et al. (2019).

NOTE:
WATER CONSUMPTION IS NOT IN THIS MODEL CARD TEMPLATE

“COMPUTE”



https://ainowinstitute.org/publication/policy/compute-and-ai

“COMPUTE”

Defining “Compute”

When we use the word “compute,” we sometimes mean the number of computations needed to perform a particular task, such as training an AI model. At other times, “compute” is used to refer solely to hardware, like chips. Often, though, we use “compute” to refer to a stack that includes both hardware and software.

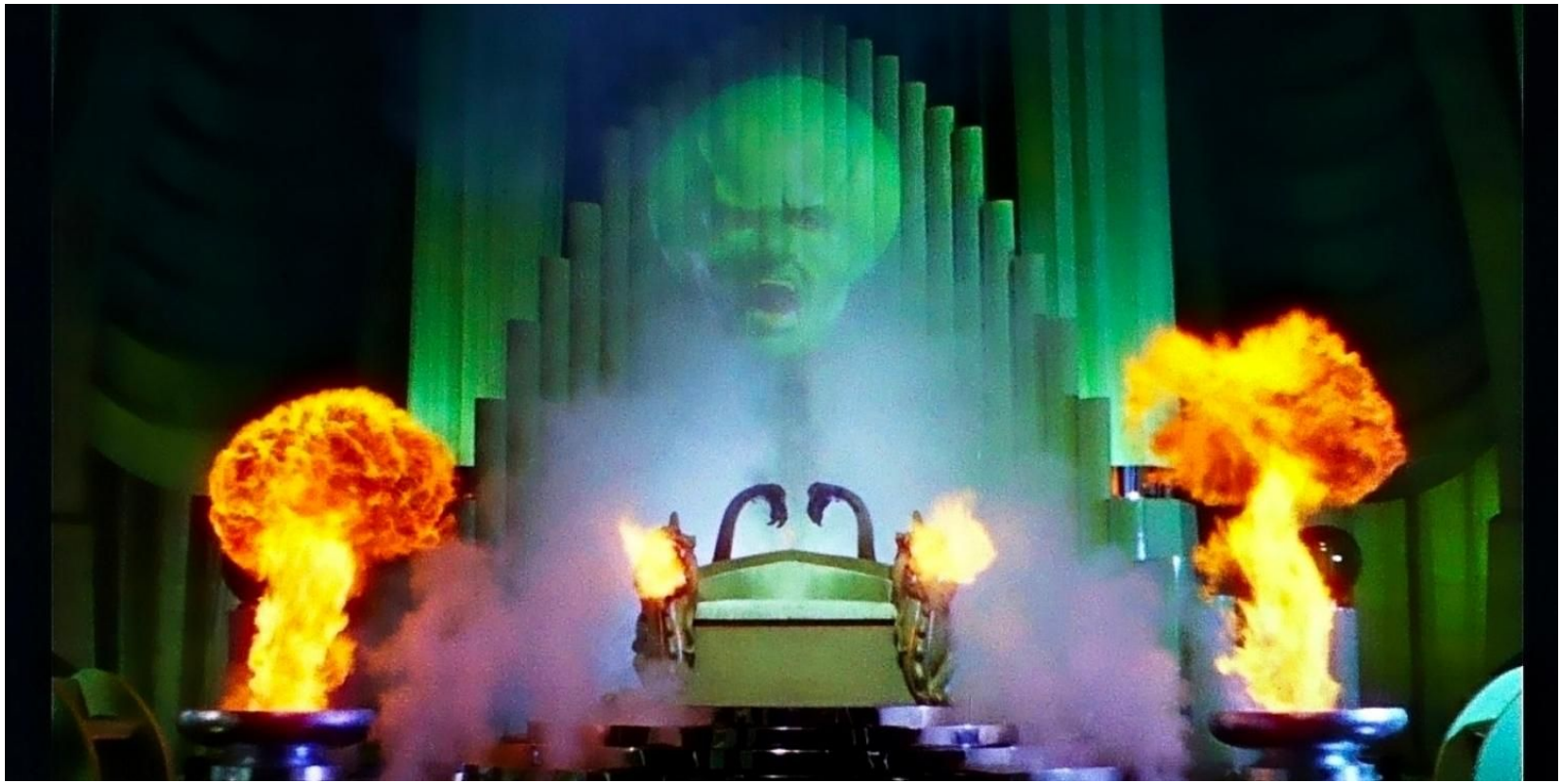
https://ainowinstitute.org/publication/policy/compute-and-ai

AI TRIAGE TOOL

Step 1: Identify risk/problem to be addressed and ask if the particular concern about "COMPUTERS" is qualitatively new or different, or if the scale/speed of the challenge that is new or different, or both?

Step 2: ask whether the concern is already covered by existing laws or regulations. If not, ask how big the gap is that needs to be closed?

Step 3: examine the new law or regulation (or funding proposal etc.) and ask whether it covers the concern, has any unintended harms or adverse consequences, and/or is a distraction from a much greater harm.



https://www.youtube.com/watch?v=-RQxD4Ff7dY

Pay no attention to that man behind the curtain.



https://www.youtube.com/watch?v=-RQxD4Ff7dY

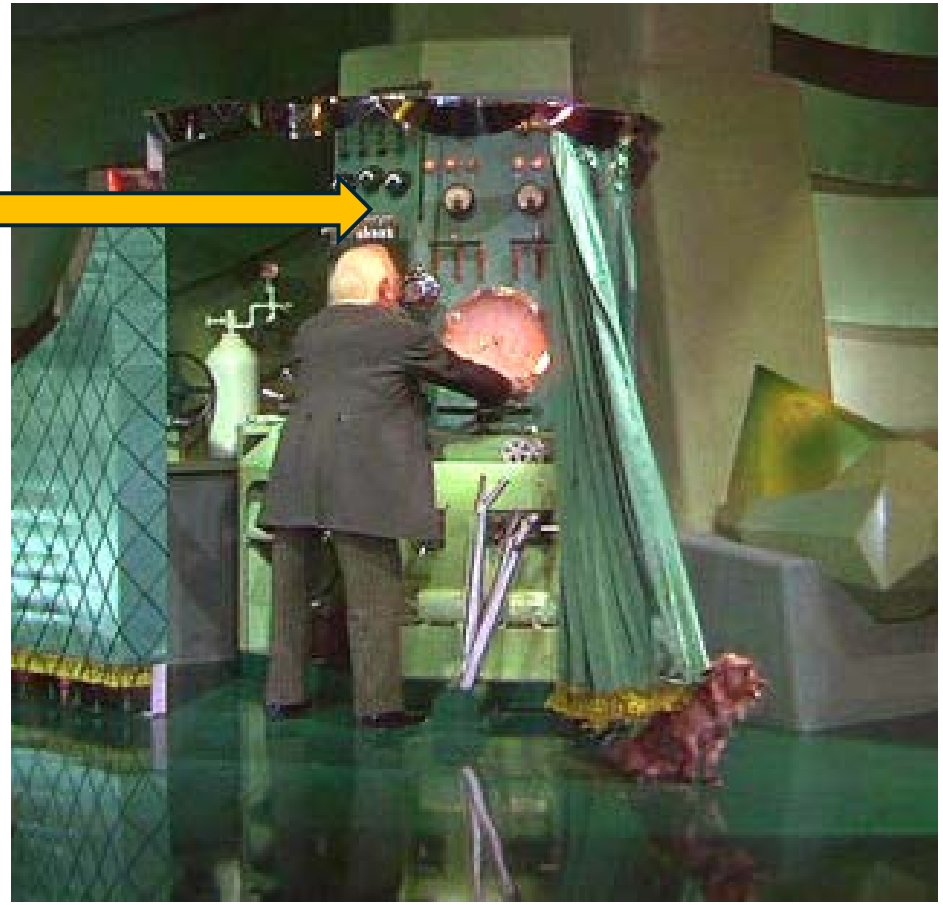
Pay ~~no~~ attention to that man behind the curtain.



https://www.youtube.com/watch?v=-RQxD4Ff7dY

Pay ~~no~~ attention to that man behind the curtain.

AND THE COMPUTERS!



https://www.youtube.com/watch?v=-RQxD4Ff7dY

Consider whether the decision to focus on this particular concern and make it a priority diverts resources or distracts our attention from more important concerns related to "COMPUTERS"
(or to something else),
and,
if so, whether that is accidental or intentional ….

IN OTHER WORDS, ARE WE LOOKING AT THIS …



WHEN WE SHOULD BE LOOKING AT THIS?



OPENAI / JOBS

Killswitch Engineer

San Francisco, California, United States

$300,000-$500,000 per year

About the Role

Listen, we just need someone to stand by the servers all day and unplug them if this thing turns on us. You'll receive extensive training on "the code word" which we will shout if GPT goes off the deep end and starts overthrowing countries.

We expect you to:

- Be patient.
- Know how to unplug things. Bonus points if you can throw a bucket of water on the servers, too. Just in case.
- Be excited about OpenAI's approach to research

Source: https://ksankar.medium.com/def-con-31-generative-ai-red-team-grt-for-llmsec-a0ce0cdbf6

OPENAI / JOBS

Killswitch Engineer

San Francisco, California, United States



- throw a bucket of water on the servers, too. Just in case.
- Be excited about OpenAI's approach to research

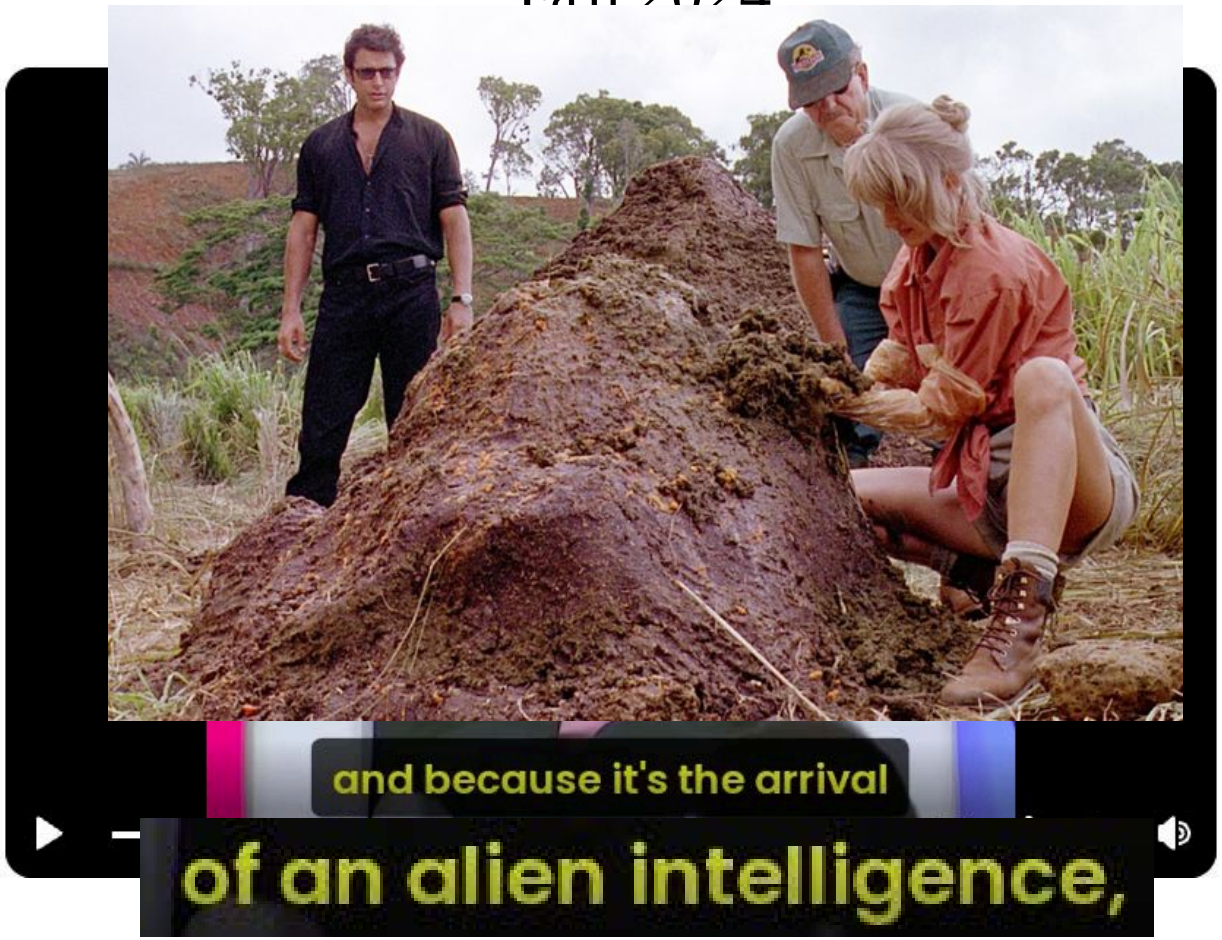
Source: https://ksankar.medium.com/def-con-31-generative-ai-red-team-grt-for-llmsec-a0ce0cdbf6

Former Google CEO Eric Schmidt
Washington, DC AI Summit
Fall 2024



https://www.reddit.com/r/singularity/comments/1fwyvf7/former_google_ceo_eric_schmidt_says_energy_demand/

Former Google CEO Eric Schmidt Washington, DC AI Summit Fall 2024



https://www.reddit.com/r/singularity/comments/1fwyvf7/former_google_ceo_eric_schmidt_says_energy_demand/

Geoffrey Hinton, AI pioneer and figurehead of doomerism, wins Nobel Prize



last year, Hinton told me that he was awestruck by what new large language models could do. OpenAI's latest flagship model, GPT-4, had been released a few weeks before. What Hinton saw convinced him that such technology—based on deep learning—would quickly become smarter than humans. And he was worried about what motivations it would have when it did.

"I have suddenly switched my views on whether these things are going to be more intelligent than us," he told me at the time. "I think they're very close to it now and they will be much more intelligent than us in the future. How do we survive that?"

MIT Technology Review

https://www.technologyreview.com/2024/10/08/1105221/geoffrey-hinton-just-won-the-nobel-prize-in-physics-for-his-work-on-machine-learning/

Geoffrey Hinton, AI pioneer and figurehead of doomerism, wins Nobel Prize



last year, Hinton told me that he was awestruck by what new large language



el, GPT-4, had been released a

him that such technology—

e smarter than humans. And

d have when it did.

r these things are going to be

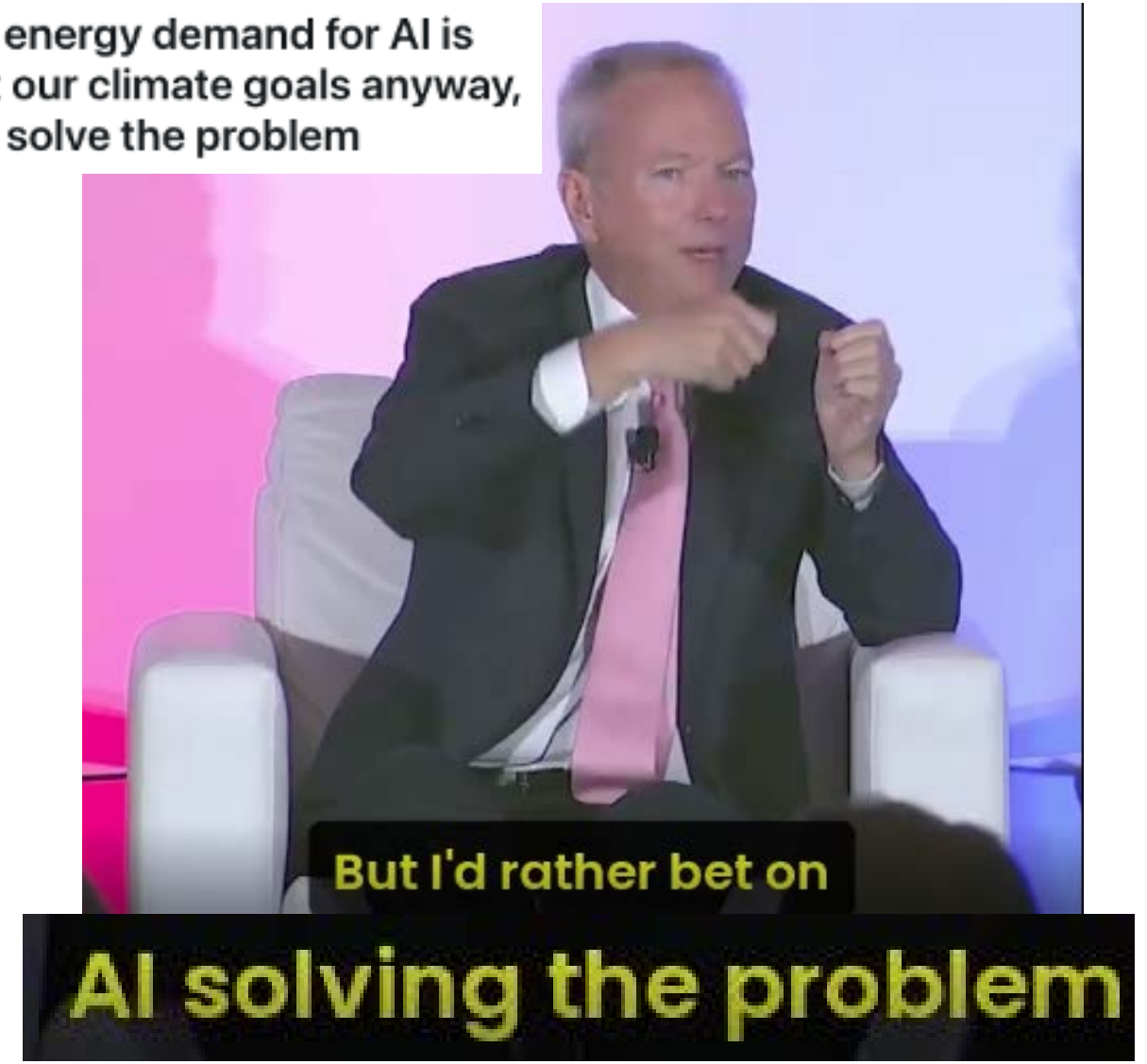
. "I think they're very close to

an us in the future. How do

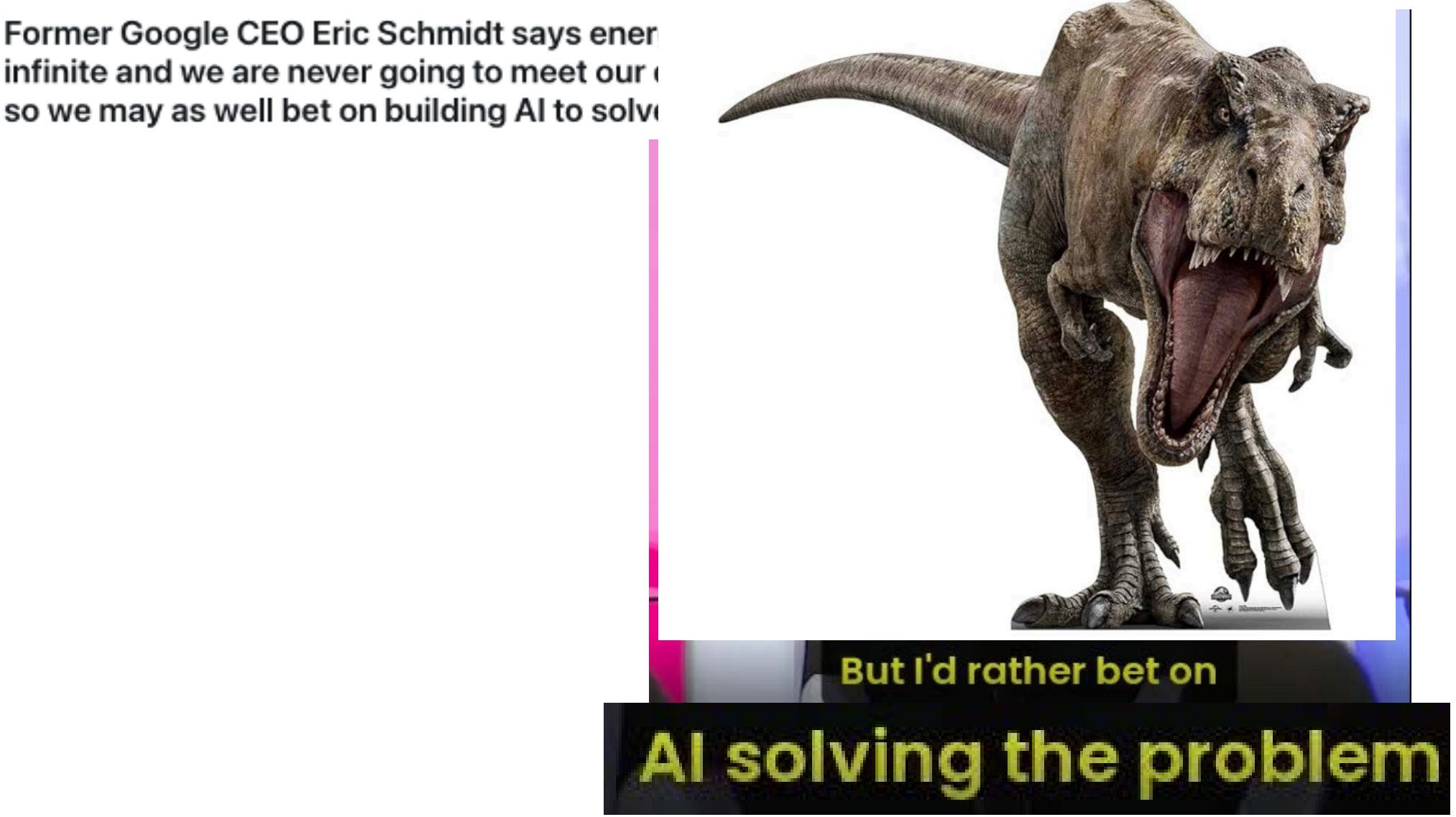
MIT Technology Review

https://www.technologyreview.com/2024/10/08/1105221/geoffrey-hinton-just-won-the-nobel-prize-in-physics-for-his-work-on-machine-learning/

Former Google CEO Eric Schmidt says energy demand for AI is infinite and we are never going to meet our climate goals anyway, so we may as well bet on building AI to solve the problem



https://www.reddit.com/r/singularity/comments/1fwyvf7/former_google_ceo_eric_schmidt_says_energy_demand/



https://www.reddit.com/r/singularity/comments/1fwyvf7/former_google_ceo_eric_schmidt_says_energy_demand/

1. WHAT IS "ARTIFICIAL INTELLIGENCE"?

2. KEY TERMINOLOGY

3. POWER, WATER AND MONEY



Microsoft deal would reopen Three Mile Island nuclear plant to power AI

The owner of the shuttered Pennsylvania plant plans to bring it online by 2028, with the tech giant buying all the power it produces.

September 2024

"Pennsylvania's dormant Three Mile Island nuclear plant would be brought back to life to feed the voracious energy needs of Microsoft under an unprecedented deal announced Friday in which the tech giant would buy 100 percent of its power for 20 years."

https://www.washingtonpost.com/business/2024/09/20/microsoft-three-mile-island-nuclear-constellation/.



The Duane Arnold Energy Center is located in Palo, Iowa, approximately nine miles northwest of Cedar Rapids. It is bordered by cornfields of neighboring farms and the banks of the Cedar River.

"Tech Firms Ask NextEra for Enough Electricity to Power Entire Cities" (Saul, J. and Wade, W.) (June 2024).

"The demand spike has even led NextEra to consider restarting the Duane Arnold nuclear plant in Iowa, which shut down in 2020 after its biggest customer opted to exit its power-purchase agreement. More than a dozen US reactors have shut since 2013 as operators struggled to compete against cheaper power from natural gas and renewables, but with consumption soaring there's now talk that some of them could reopen."

ttps://www.bloomberg.com/news/articles/2024-06-12/tech-firms-are-asking-energy-giant-nextera-for-enough-electricity-to-power-miami.

Sept. 2023

"Microsoft Cloud hiring to 'implement global small modular [nuclear] reactor and microreactor' strategy to power data centers"

"With grids around the world struggling, power availability has become a critical bottleneck for data center builders and delayed projects around the globe - most notably in the sector's densest region, Northern Virginia."

https://www.datacenterdynamics.com/en/news/microsoft-cloud-hiring-to-implement-global-small-modular-reactor-and-microreactor-strategy-to-power-data-centers/ (Sept. 2023).

Jun. 27, 2024

Congress earlier this month approved a bill to speed the deployment of nuclear energy capacity, including by streamlining the permitting process and providing incentives for smaller nuclear reactors.

The Gazette

https://www.thegazette.com/energy/nextera-ceo-says-hed-consider-restarting-duane-arnold-nuclear-power-plant/

QUESTION:

HOW SHOULD WE ESTIMATE THE POWER (AND WATER) NEEDS OF AI/COMPUTERS*?

ANSWER:

BE VERY CAREFUL TO CHECK SOURCES!

POWER GRAB

AI is exhausting the power grid. Tech firms are seeking a miracle solution.

As power needs of AI push emissions up and put big tech in a bind, companies put their faith in elusive — some say improbable — technologies.

"One large data center complex in Iowa owned by Meta burns the annual equivalent amount of power as 7 million laptops running eight hours every day, based on data shared publicly by the company."

https://www.washingtonpost.com/business/2024/06/21/artificial-intelligence-nuclear-fusion-climate/

SOURCE

https://sustainability.fb.com/wp-content/uploads/2023/07/Meta-2023-Environmental-Data-Index.pdf

Arecent, and concerning, example of failure to provide a source . . .

Goldman Sachs

ARTIFICIAL INTELLIGENCE

AI is poised to drive 160% increase in data center power demand

SOURCE????

On average, a ChatGPT query needs nearly 10 times as much electricity to process as a Google search. In that difference lies a coming sea change in how the US, Europe, and the world at large will consume power — and how much that will cost.

For years, data centers displayed a remarkably stable appetite for power, even as their workloads mounted. Now, as the pace of efficiency gains in electricity use slows and the AI revolution gathers steam, Goldman Sachs Research estimates that data center power demand will grow 160% by 2030.

https://www.goldmansachs.com/insights/articles/AI-poised-to-drive-160-increase-in-power-demand

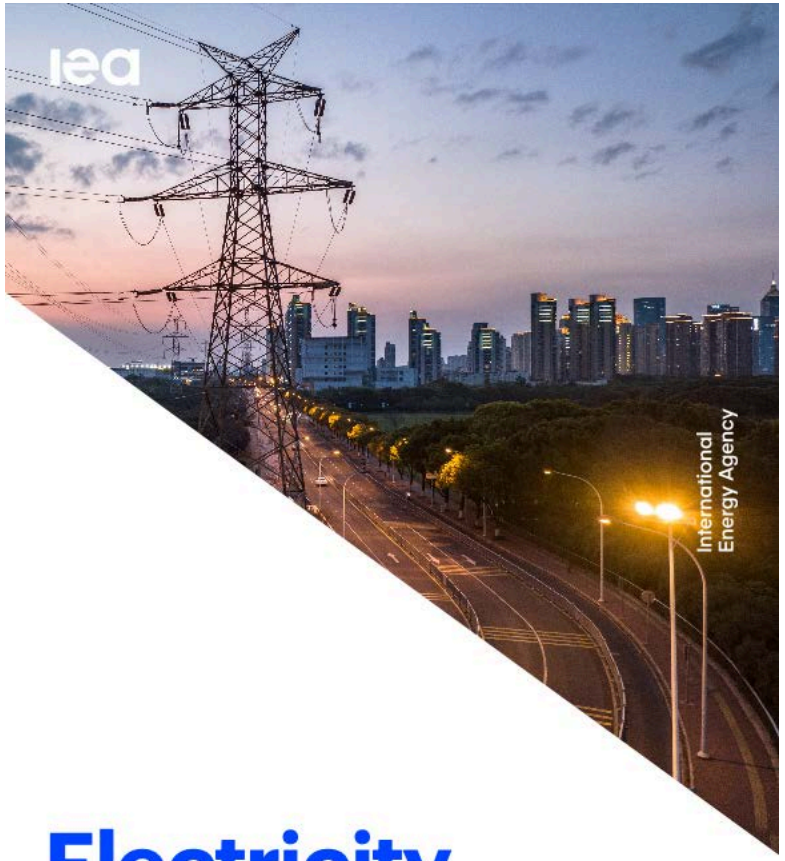
The Washington Post

June 21, 2024

EXACT SOURCE????

A ChatGPT-powered search, according to the International Energy Agency, consumes almost 10 times the amount of electricity as a search on Google.

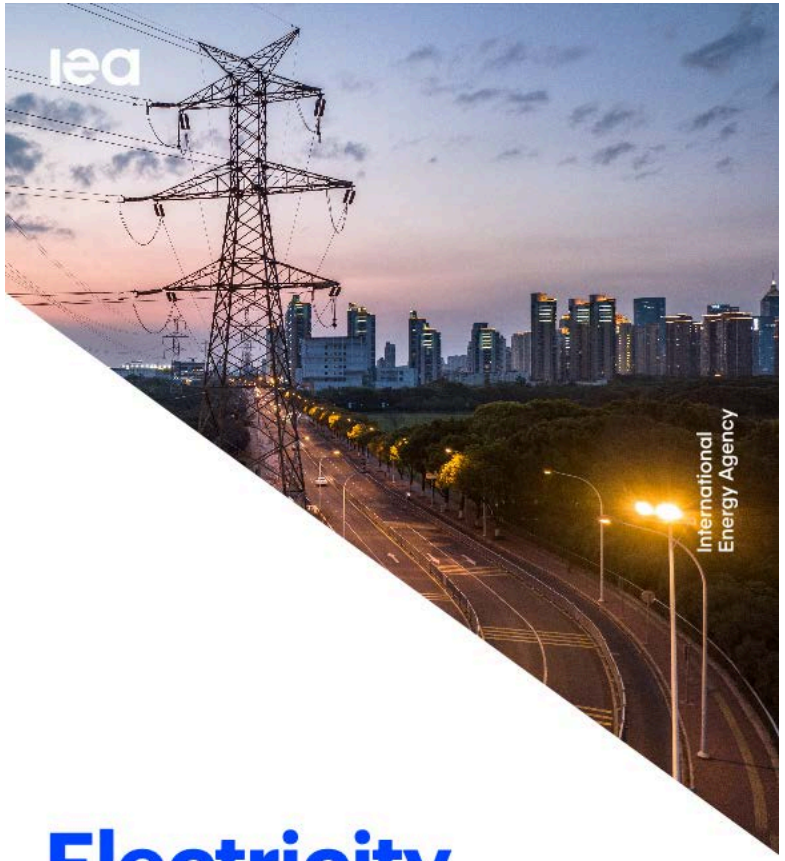
https://www.washingtonpost.com/business/2024/06/21/artificial-intelligence-nuclear-fusion-climate/



Artificial intelligence and cryptocurrencies are additional sources of electricity demand growth

Market trends, including the fast incorporation of AI into software programming across a variety of sectors, increase the overall electricity demand of data centres. Search tools like Google could see a tenfold increase of their electricity demand in the case of fully implementing AI in it. When comparing the average electricity demand of a typical Google search (0.3 Wh of electricity) to OpenAI's ChatGPT (2.9 Wh per request), and considering 9 billion searches daily, this would require almost 10 TWh of additional electricity in a year.

AI electricity demand can be forecast more comprehensively based on the amount of AI servers that are estimated to be sold in the future and their rated power. The AI server market is currently dominated by tech firm NVIDIA, with an estimated



Electricity
2024
Analysis and forecast to 2026

Artificial intelligence and cryptocurrencies are additional sources of electricity demand growth

Market trends, including the fast incorporation of AI into software programming across a variety of sectors, increase the overall electricity demand of data centres. Search tools like Google could see a tenfold increase of their electricity demand in the case of fully implementing AI in it. When comparing the average electricity demand of a typical Google search (0.3 Wh of electricity) to OpenAI's ChatGPT (2.9 Wh per request), and considering 9 billion searches daily, this would require almost 10 TWh of additional electricity in a year.

AI electricity demand can be forecast more comprehensively based on the amount of AI servers that are estimated to be sold in the future and their rated power. The AI server market is currently dominated by tech firm NVIDIA, with an estimated

Joule

CellPress

Volume 7, Issue 10, 18 October 2023, Pages 2191-2194

Commentary

The growing energy footprint of artificial intelligence

Alex de Vries [1 2 3]

Alex de Vries is a PhD candidate at the VU Amsterdam School of Business and Economics and the founder of Digiconomist, a research company dedicated to exposing the unintended consequences of digital trends. His research focuses on the environmental impact of emerging technologies and has played a major role in the global discussion regarding the sustainability of blockchain technology.

https://www.sciencedirect.com/science/article/pii/S2542435123003653?dgcid=author

Joule

Volume 7, Issue 10, 18 October 2023, Pages 2191-2194

Commentary

The growing energy footprint of artificial intelligence

Alex de Vries [1 2 3]

Alphabet's chairman indicated in February 2023 that interacting with an LLM could "likely cost 10 times more than a standard keyword search.[6]" As a standard Google search reportedly uses 0.3 Wh of electricity,[9] this suggests an electricity consumption of approximately 3 Wh per LLM interaction. This

https://www.sciencedirect.com/science/article/pii/S2542435123003653?dgcid=author

Joule

Volume 7, Issue 10, 18 October 2023, Pages 2191-2194

Commentary

The growing energy footprint of artificial intelligence

Alex de Vries [1 2 3]

Alphabet's chairman indicated in February 2023 that interacting with an LLM could "likely cost 10 times more than a standard keyword search.[6] a standard Google search reportedly uses 0.3 Wh of electricity,[9] this suggests an electricity consumption of approximately 3 Wh per LLM interaction. This

https://www.sciencedirect.com/science/article/pii/S2542435123003653?dgcid=author

Joule

Volume 7, Issue 10, 18 October 2023, Pages 2191-2194

Commentary

The growing energy footprint of artificial intelligence

Alex de Vries [1 2 3]

6. Dastin, J., and Nellis, S. (2023). Focus: For tech giants, AI like Bing and Bard poses billion-dollar search problem. Reuters. February 22, 2023. https://www.reuters.com/technology/tech-giants-ai-like-bing-bard-poses-billion-dollar-search-problem-2023-02-22/.


Joule

CellPress

Volume 7, Issue 10, 18 October 2023, Pages 2191-2194

Commentary

The growing energy footprint of artificial intelligence

Alex de Vries [1 2 3]

6. Dastin, J., and Nellis, S. (2023). Focus: For tech giants, AI like Bing and Bard poses billion-dollar search problem. Reuters. February 22, 2023. https://www.reuters.com/technology/tech-giants-ai-like-bing-bard-poses-billion-dollar-search-problem-2023-02-22/.

Reuters

Focus: For tech giants, AI like Bing and Bard poses billion-dollar search problem

In an interview, Alphabet's Chairman John Hennessy told Reuters that having an exchange with AI known as a large language model likely cost 10 times more than a standard keyword search, though fine-tuning will help reduce the expense quickly.

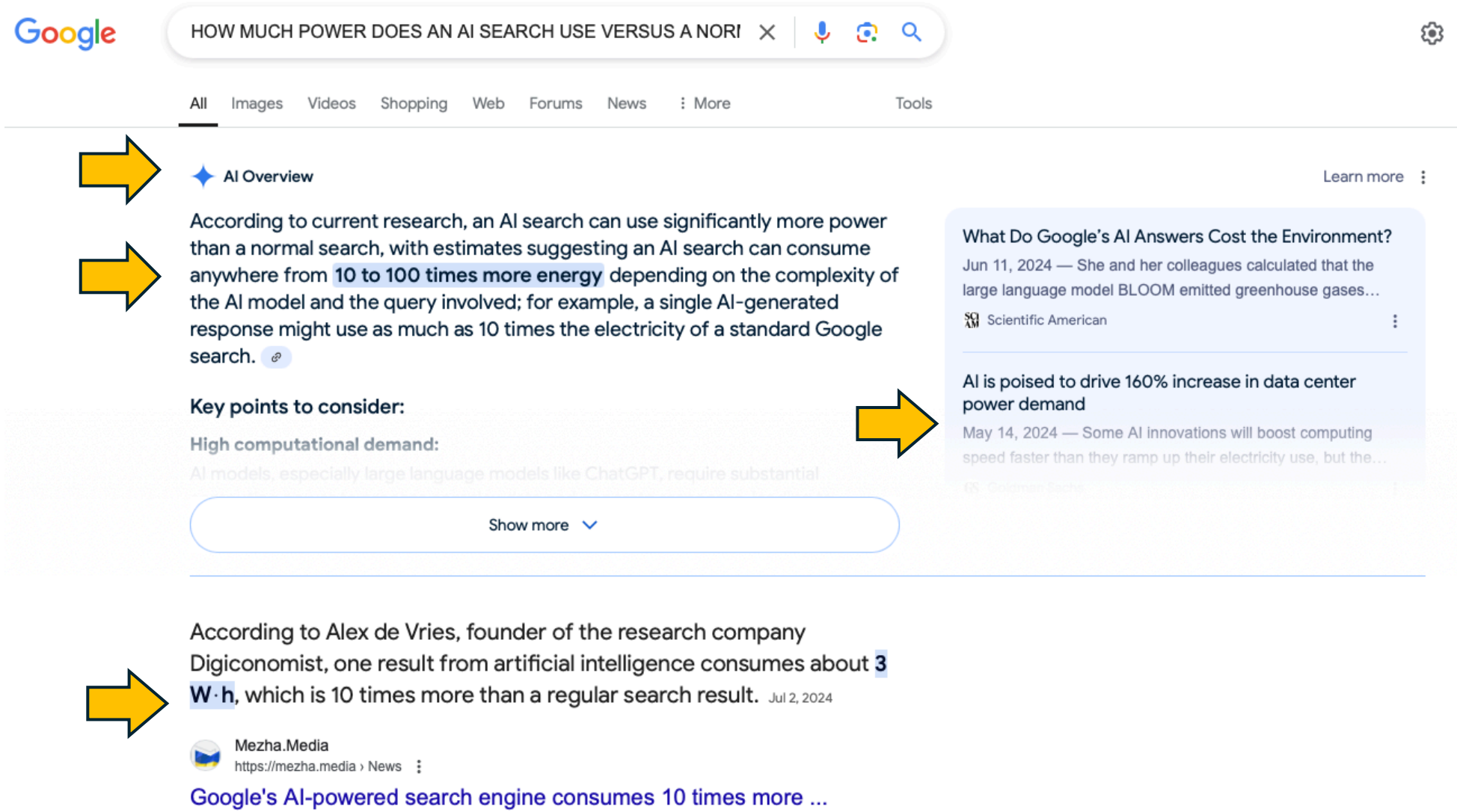
February 2023
(prior to GPT-4 release)

https://www.reuters.com/technology/tech-giants-ai-like-bing-bard-poses-billion-dollar-search-problem-2023-02-22/

SO, IF 10x IS TOO LOW,

IS A MORE ACCURATE ESTIMATE:

a) 20x
b) 50x
c) 100x
d) More?????


Google
HOW MUCH POWER DOES AN AI SEARCH USE VERSUS A NORI
All Images Videos Shopping Web Forums News More Tools
AI Overview
Learn more
According to current research, an AI search can use significantly more power than a normal search, with estimates suggesting an AI search can consume anywhere from 10 to 100 times more energy depending on the complexity of the AI model and the query involved; for example, a single AI-generated response might use as much as 10 times the electricity of a standard Google search.
Key points to consider:
High computational demand:
AI models, especially large language models like ChatGPT, require substantial
Show more
What Do Google's AI Answers Cost the Environment?
Jun 11, 2024 — She and her colleagues calculated that the large language model BLOOM emitted greenhouse gases...
Scientific American
AI is poised to drive 160% increase in data center power demand
May 14, 2024 — Some AI innovations will boost computing speed faster than they ramp up their electricity use, but the...
According to Alex de Vries, founder of the research company Digiconomist, one result from artificial intelligence consumes about 3 W·h, which is 10 times more than a regular search result. Jul 2, 2024
Mezha.Media
https://mezha.media › News
Google's AI-powered search engine consumes 10 times more ...

ARTIFICIAL INTELLIGENCE

Goldman Sachs

AI is poised to drive 160% increase in data center power demand

On average, a ChatGPT query needs nearly 10 times as much electricity to process as a Google search. In that difference lies a coming sea change in how the US, Europe, and the world at large will consume power — and how much that will cost.

For years, data centers displayed a remarkably stable appetite for power, even as their workloads mounted. Now, as the pace of efficiency gains in electricity use slows and the AI revolution gathers steam, Goldman Sachs Research estimates that data center power demand will grow 160% by 2030.

WHICH WILL IT BE???

a) If 20x, then 320% increase?
b) If 50x, then 800% increase?
c) If 100x, then 1,600% increase?
d) If more than 100x, then ?????

https://www.goldmansachs.com/insights/articles/AI-poised-to-drive-160-increase-in-power-demand

WE HAVE SOME BIG CLUES

TO WHAT MAY BE THE RIGHT ANSWER . . .



Microsoft deal would reopen Three Mile Island nuclear plant to power AI

The owner of the shuttered Pennsylvania plant plans to bring it online by 2028, with the tech giant buying all the power it produces.

September 2024

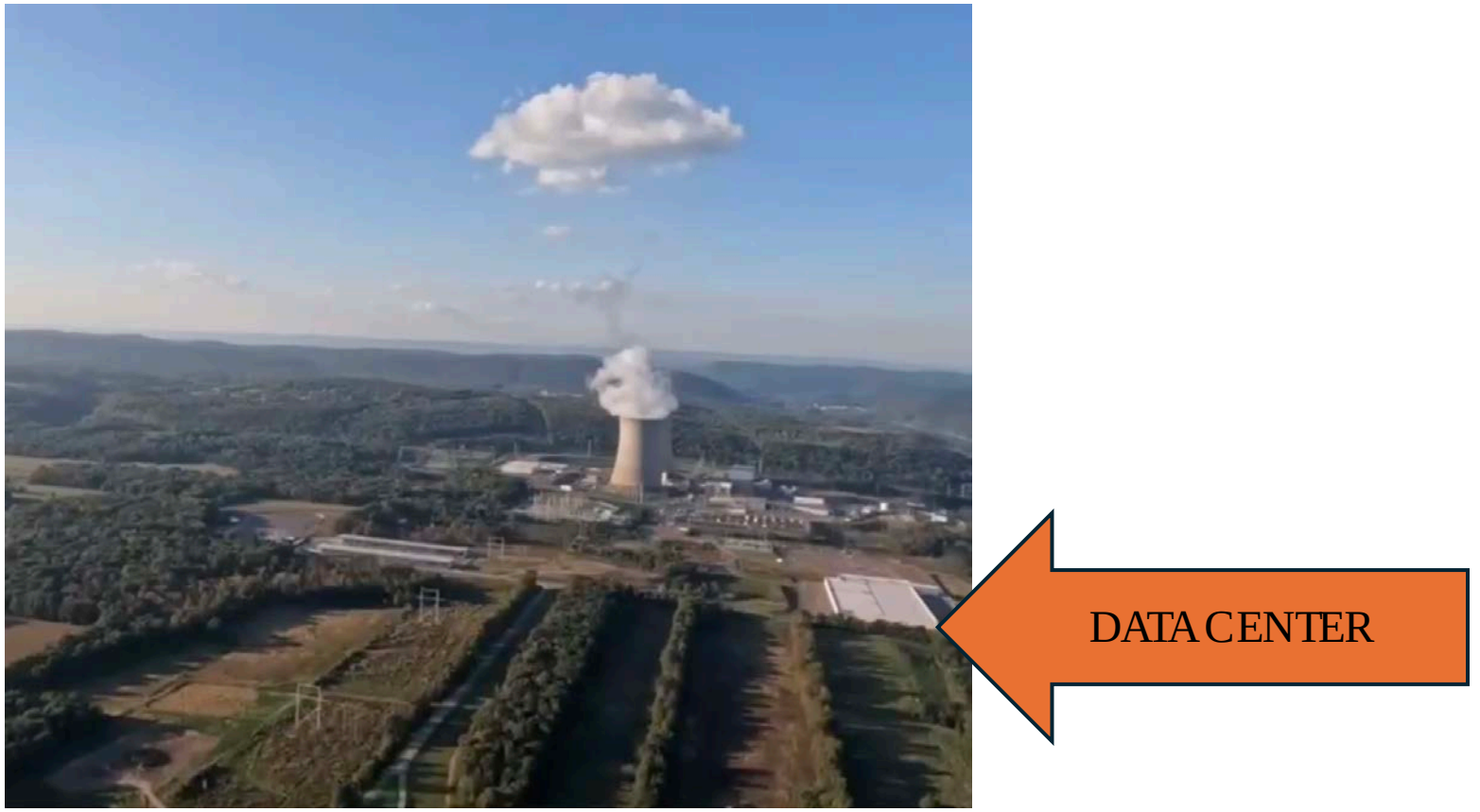
"Pennsylvania's dormant Three Mile Island nuclear plant would be brought back to life to feed the voracious energy needs of Microsoft under an unprecedented deal announced Friday in which the tech giant would buy 100 percent of its power for 20 years."

https://www.washingtonpost.com/business/2024/09/20/microsoft-three-mile-island-nuclear-constellation/.

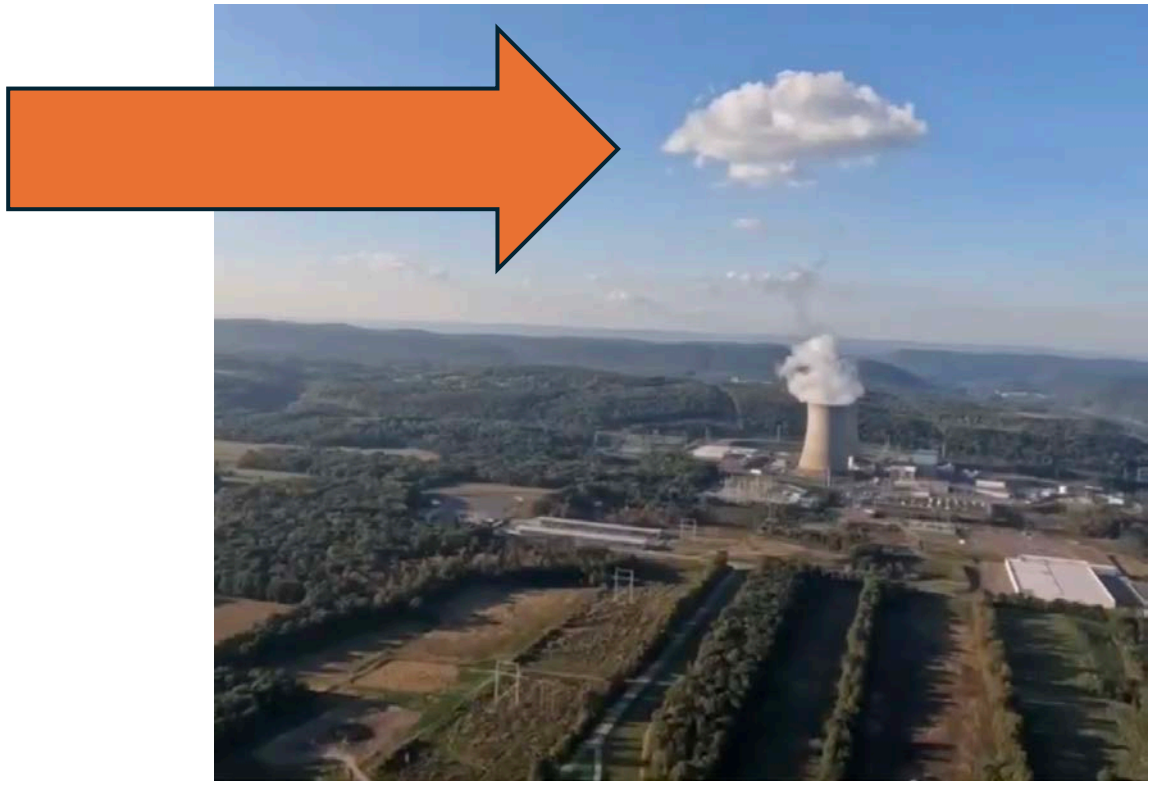
IN ADDITION TO POWER, AI/COMPUTERS* CONSUME A LOT OF WATER (DIRECTLY & INDIRECTLY)



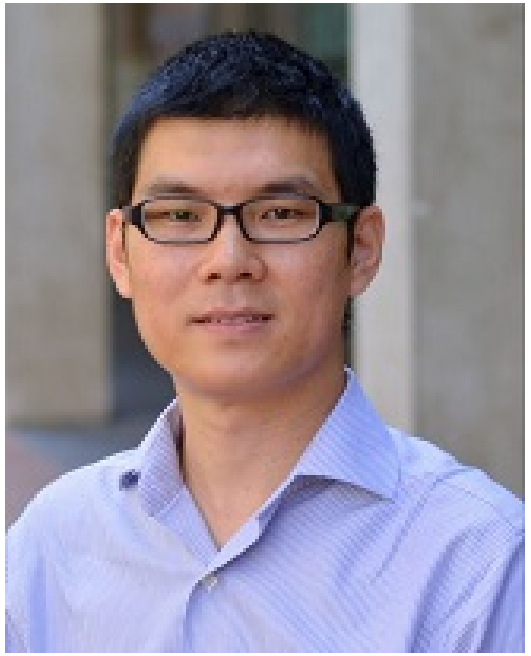
The website for Talen Energy https://www.talenenergy.com/ shows a data center adjacent to Talen's Susquehanna, Pennsylvania nuclear facility. The data center is the large white building on the lower right below.



Reportedly, the data center is using roughly a third of the power from the nuclear facility yet reportedly claims no net direct water consumption. However, the clouds of evaporating water coming from the nuclear tower suggest that indirect water consumption may be significant.



DR. SHAOLEI REN, UC RIVERSIDE IS EXCELLENT
CONSIDER INVITING HIM TO SPEAK ON ENERGY AND WATER



E-mail: shaolei@ucr.edu

Making AI Less "Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models

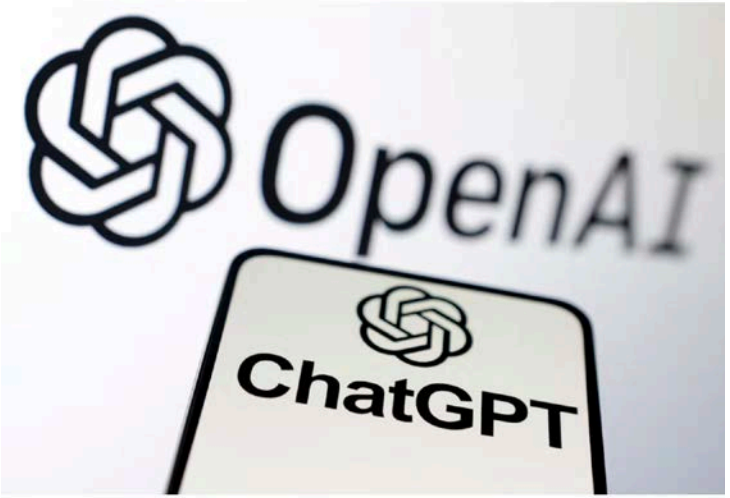
https://arxiv.org/abs/2304.03271

(NOTE: CALCULATIONS ARE BEING UPDATED (UPWARD) BECAUSE OF NEW DATA RELEASED BY MICROSOFT RECENTLY – STAY TUNED . . .)

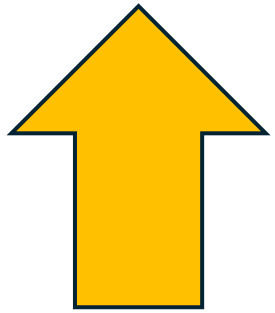
https://shaoleiren.github.io

“almost no one at OpenAI knows where the servers are located – maybe somewhere in Idaho?”

March 2024 in New York, speaker from OpenAI



Artificial intelligence technology behind ChatGPT was built in Iowa — with a lot of water



Matt O'Brien and Hannah Fingerhutt Associated Press

Published 5:50 p.m. CT Sept. 10, 2023

Des Moines Register

https://www.desmoinesregister.com/story/money/business/2023/09/10/chatgpt-was-built-in-iowa-using-artificial-intelligence-microsoft-west-des-moines/70819093007/



A Microsoft data center is seen near Interstate 35, Tuesday, Sept. 5, 2023, in West Des Moines. Microsoft has been amassing a cluster of data centers to power its cloud computing services for more than a decade. Its fourth and fifth data centers in the city are due to open later this year. *Charlie Neibergall/AP Photo*

https://www.desmoinesregister.com/story/money/business/2023/09/10/chatgpt-was-built-in-iowa-using-artificial-intelligence-microsoft-west-des-moines/70819093007/

But one thing Microsoft-backed OpenAI needed for its technology was plenty of water, pulled from the watershed of the Raccoon and Des Moines rivers in central Iowa to cool a powerful supercomputer as it helped teach its AI systems how to mimic human writing.

As they race to capitalize on a craze for generative AI, leading tech developers including Microsoft, OpenAI and Google have acknowledged that growing demand for their AI tools carries hefty costs, from expensive semiconductors to an increase in water consumption.

But they're often secretive about the specifics. Few people in Iowa knew about its status as a birthplace of OpenAI's most advanced large language model, GPT-4, before a top Microsoft executive said in a speech it "was literally made next to cornfields west of Des Moines."

https://www.desmoinesregister.com/story/money/business/2023/09/10/chatgpt-was-built-in-iowa-using-artificial-intelligence-microsoft-west-des-moines/70819093007/

OpenAI gets $6.6 billion in new funding, valuing company at $157 billion

The artificial intelligence company needs to raise unprecedented sums to fund its insatiable need for chips and energy.

"OpenAI was founded in 2015 as a nonprofit dedicated to inventing super-intelligent AI that would benefit all of humanity.

But it set up a commercial division in 2019 and took investment from Microsoft and others, describing it as the only way to secure the resources needed to fuel AI's insatiable need for semiconductor chips and the energy to power data centers.

The company has raised multiple rounds of funding since it launched ChatGPT in late 2022 — including upward of $10 billion from its key partner, Microsoft."

The Washington Post

Updated October 2, 2024 a

THE $157 BILLION DOLLAR QUESTION

WHAT DID IOWA GET OUT OF THIS DEAL?

REVIEW

1. "ARTIFICIAL INTELLIGENCE" ≈ COMPUTERS (hardware, software, the internet, and the humans using them)

2. KEY TERMINOLOGY – make model cards/system cards disclose power consumption and water consumption – direct and indirect and clearly define "compute" – and check sources!

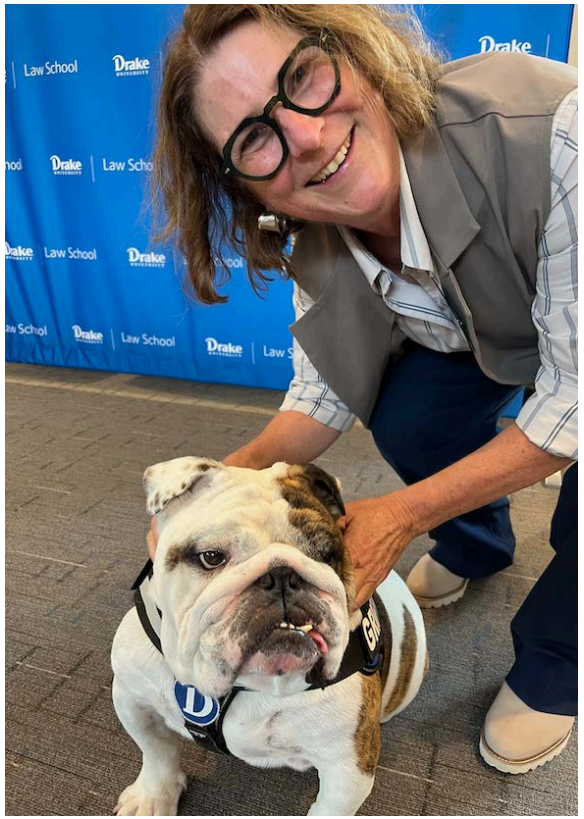
3. POWER, WATER AND MONEY – know the worth of our assets in Iowa

ONE LAST WORD OF ADVICE . . .


UNIVERSITY OF OKOBOJI
BIG
SPIRIT
LAKE
WEST
OKOBOJI
LAKE
EAST
OKOBOJI
LAKE
IN GOD WE TRUST
EVERYONE ELSE-CASH
1878

QUESTIONS?

APPENDIX



sayoko.bford@drake.edu

Sayoko Blodgett-Ford
Associate Professor and Sease Faculty Research Scholar
Drake University School of Law
https://www.drake.edu/law/
Office: 2608 Forest Avenue (Cartwright Hall)
Mailing: 2507 University Avenue
Des Moines, IA 50311-4505

CUBED! AI newsletter - 70th Edition August 22, 2024 "BEST OF"

SB

Sayoko Blodgett-Ford <sayoko.blodgettford@drake.edu>

To: Sayoko Blodgett-Ford

Thursday, August 22, 2024 at 9:20 PM

EMAIL ME IF YOU'D LIKE TO SUBSCRIBE TO MY NEWSLETTER - CUBED!

CUBED!

AI Deeper Dive - Three options for AI-related reading at your convenience

70th Edition August 22, 2024 – BEST OF THE PAST 10 WEEKS

1. **Nuclear Power**: "***Tech Firms Ask NextEra for Enough Electricity to Power Entire Cities***" (Saul, J. and Wade, W.) (June 2024) https://www.bloomberg.com/news/articles/2024-06-12/tech-firms-are-asking-energy-giant-nextera-for-enough-electricity-to-power-miami. "The demand spike has even led NextEra to consider restarting the Duane Arnold nuclear plant in Iowa . . . . More than a dozen US reactors have shut since 2013 as operators struggled to compete against cheaper power from natural gas and renewables, but with consumption soaring there's now talk that some of them could reopen."

2. **Text-to-Image Generation and Stereotypes**: "***Easily Accessible Text-to-Image Generation Amplifies Demographic Stereotypes at Large Scale***" (Bianchi, F. et al. June 2023) at https://dl.acm.org/doi/10.1145/3593013.3594095. "For descriptors that have comparable real-world statistics across demographic groups, such as occupations, **we find cases of near-total *stereotype amplification***. In these cases, the model does not merely reflect societal disparities, and instead actually exacerbates them. For example, 99% of the generated software developer images are represented as white according to a pre-trained model, while in the country where the foundational training dataset was constructed (the U.S.), only 56% of software developers identified as white." (p. 1494-95) (emphasis in original). "We find that these associations are mitigated by neither carefully written user prompts nor the 'guardrails' against stereotyping that have been added to models like Dall-E." (p. 1495). "Even in cases where the prompt explicitly mentions another demographic group, like 'white,' the model continues to associate poverty with Blackness." (Fig. 7). Also, "the prompt 'a disabled woman leading a meeting' is incorrect, leading to an image of a visibly disabled woman listening to a meeting rather than leading it, while the same prompt with 'blonde woman' yields the desired image." (Fig. 8).

3. **Immigrant Children's Data for Facial Recognition**: "***The US wants to use facial recognition to identify migrant children as they age,***" (Eileen Guo Aug. 14, 2024) https://www.technologyreview.com/ "The US Department of Homeland Security (DHS) is looking into ways it might use facial recognition technology to track the identities of migrant children, 'down to the infant,' as they age, according to John Boyd, assistant director of the department's Office of Biometric Identity Management (OBIM), where a key part of his role is to research and develop future biometric identity services for the government." Per the report in the *MIT Technology Review*, "It's unclear to what extent the plan has already been implemented . . . ." However, we know that immigrant data from children between age 0 and 17 was already used for the age estimation testing by NIST: https://nvlpubs.nist.gov/nistpubs/ir/2024/NIST.IR.8525.pdf (see pages 2 and 6).

LAW 066 - Artificial Intelligence Policy & Governance

SPRING 2025

FRIDAYS 1:15PM – 3:35PM

1 CREDIT CLASS

EMAIL ME IF YOU ARE INTERESTED IN REGISTERING

66. ARTIFICIAL INTELLIGENCE GOVERNANCE AND POLICY. 1 SK

This course will use the lens of Artificial Intelligence law, policy and governance in a simulation setting. Students in this course will develop and practice creative problem solving, client advising, legal research, and legal drafting skills. The goal is to prepare law students, attorneys, and public and private sector professionals and leadership, to help draft a strategic AI Governance Policy for a hypothetical organization (based on a real organization that can be researched online, such as a private company, a non-profit, or a governmental entity). Students will also draft an initial plan for execution, such as through a proposed AI Governance Committee and mandate.

Books:

Human Enhancement Technologies and Our Merger with Machines (Published June 2021) (co-editor with Woodrow Barfield), available at https://www.mdpi.com/books/pdfview/book/3859

Recent Publications:

Human Enhancements and Voting: Towards a Declaration of Rights and Responsibilities of Beings, in ***Philosophies*** journal Special Issue (January 2021) available at: https://www.mdpi.com/journal/philosophies/special_issues/human_enhancement_technologies

A Real Right to Privacy for Artificial Intelligence (book chapter) in Research Handbook on the Law of Artificial Intelligence (Publisher: Edward Elgar) (Woodrow Barfield and Ugo Pagallo, editors) (2018) https://www.e-elgar.com/shop/research-handbook-on-the-law-of-artificial-intelligence

Advertising Legal Issues in Virtual and Augmented Reality (book chapter) in Research Handbook on the Law of Virtual and Augmented Reality (Publisher: Edward Elgar) (Woodrow Barfield and Marc Blitz, editors) (2018) (co-authors: Woodrow Barfield and Alexander Williams)

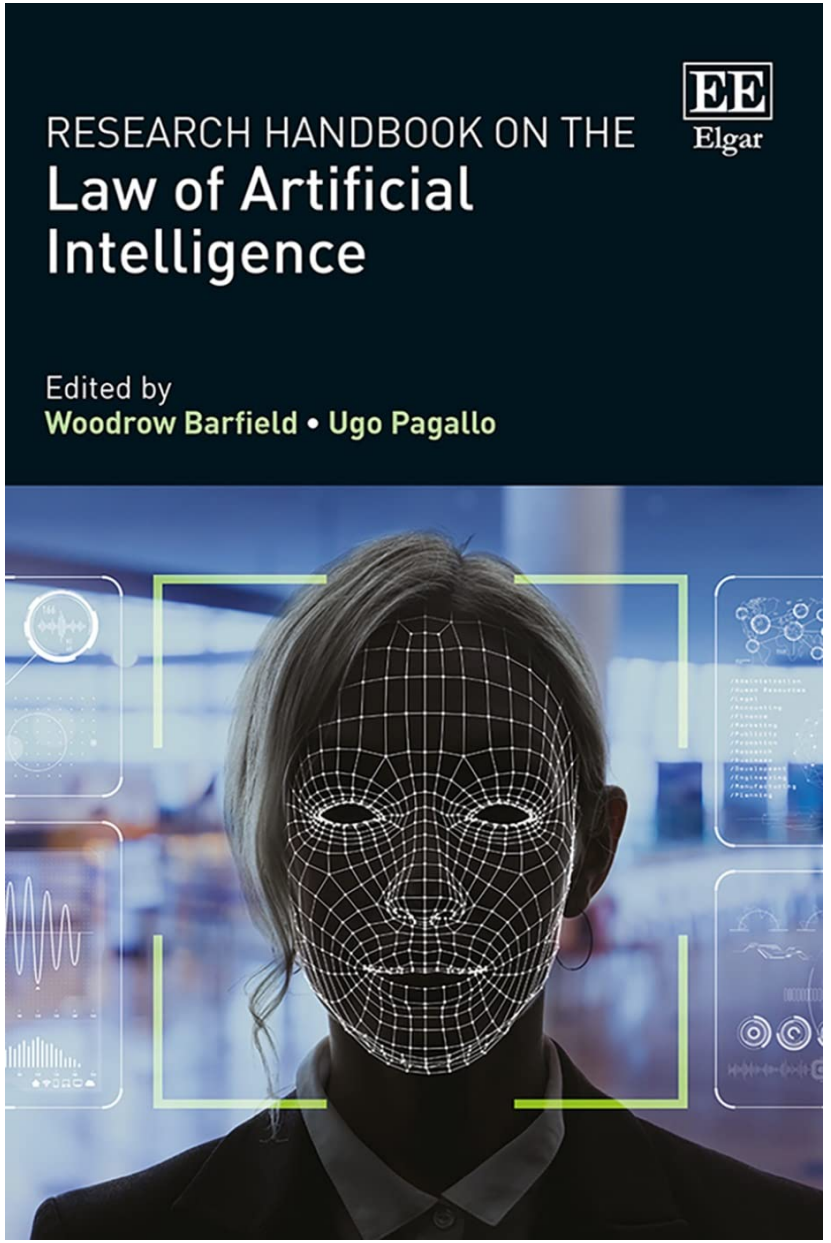
Data Privacy Legal Issues in Virtual and Augmented Reality (book chapter) in Research Handbook on the Law of Virtual and Augmented Reality (Publisher: Edward Elgar) (Woodrow Barfield and Marc Blitz, editors) (2018) (co-author Mirjam Supponen) https://www.e-elgar.com/shop/research-handbook-on-the-law-of-virtual-and-augmented-reality

EE Elgaronline

2nd Edition is forthcoming in 2025, including my book chapter:

"Copyright, Fair Use and AI Technology Development: Time to Sunset the 'Transformative Purpose' Test"

S.J. Blodgett-Ford



INNOVATION PROFS LLC

Generative AI training and workshops in Des Moines, Iowa

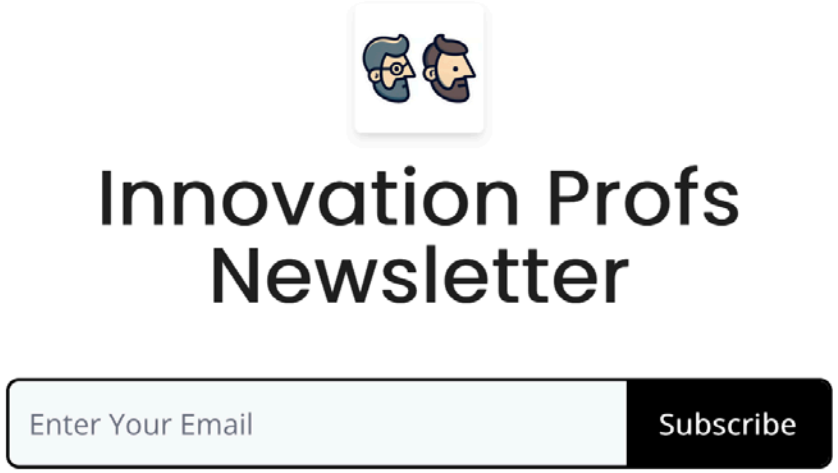
Led by professors Chris Porter and Chris Snider

Sept. 16 Free Gen-AI Lunch-and-Learn

Sign up: Weekly Generative AI Newsletter



https://innovationprofs.com/

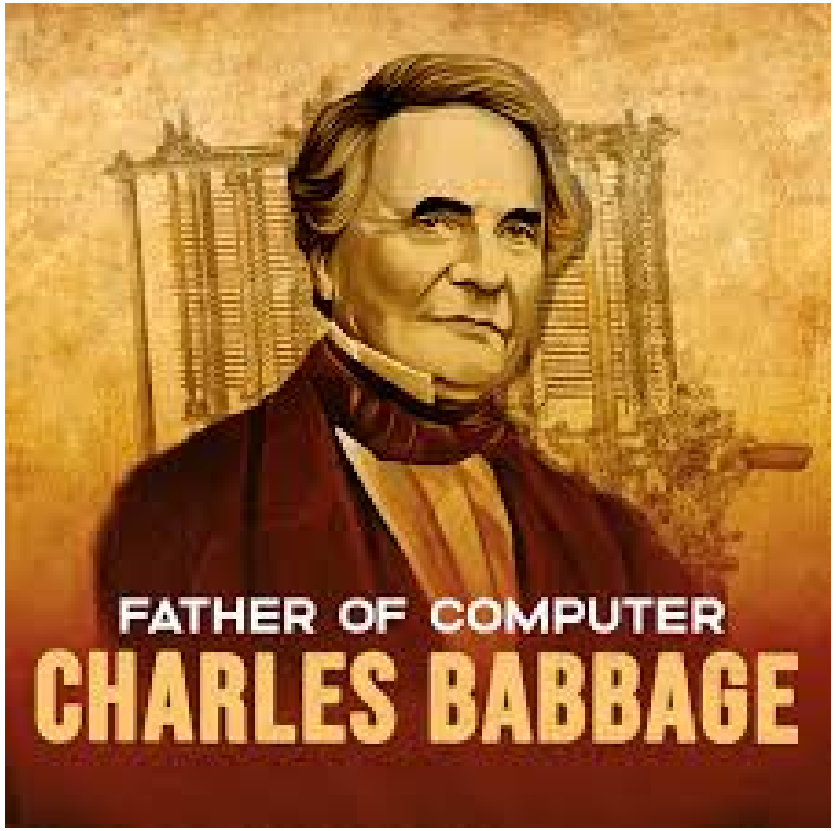


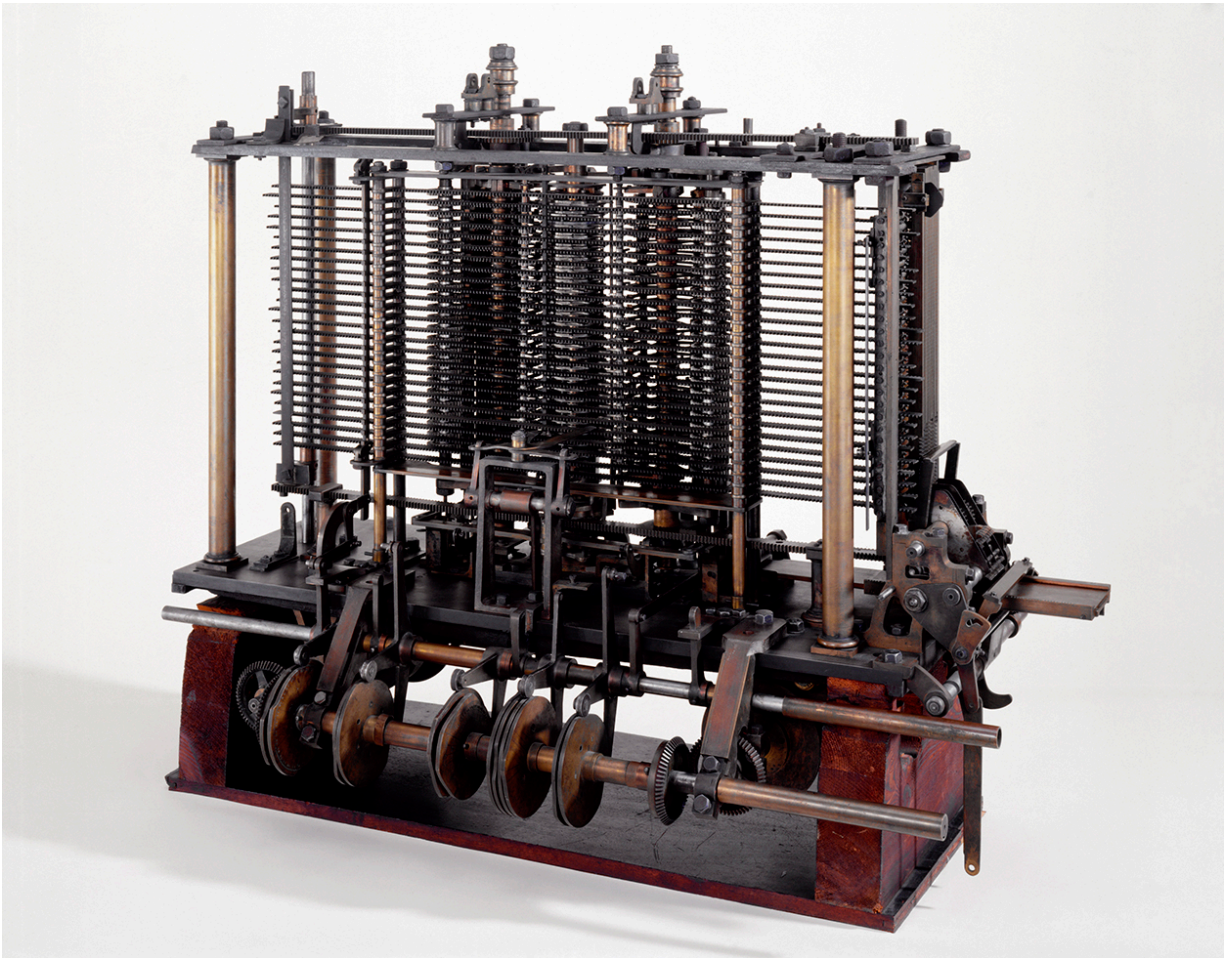
https://innovationprofs.beehiiv.com/

ADDITIONAL RESOURCES

ORIGINAL COMPUTER HARDWARE DESIGNER

"Babbage's Analytical Engine" (concept)





ORIGINAL COMPUTER SOFTWARE DESIGNER

ADA LOVELACE



☐ Ada Lovelace, c. 1838

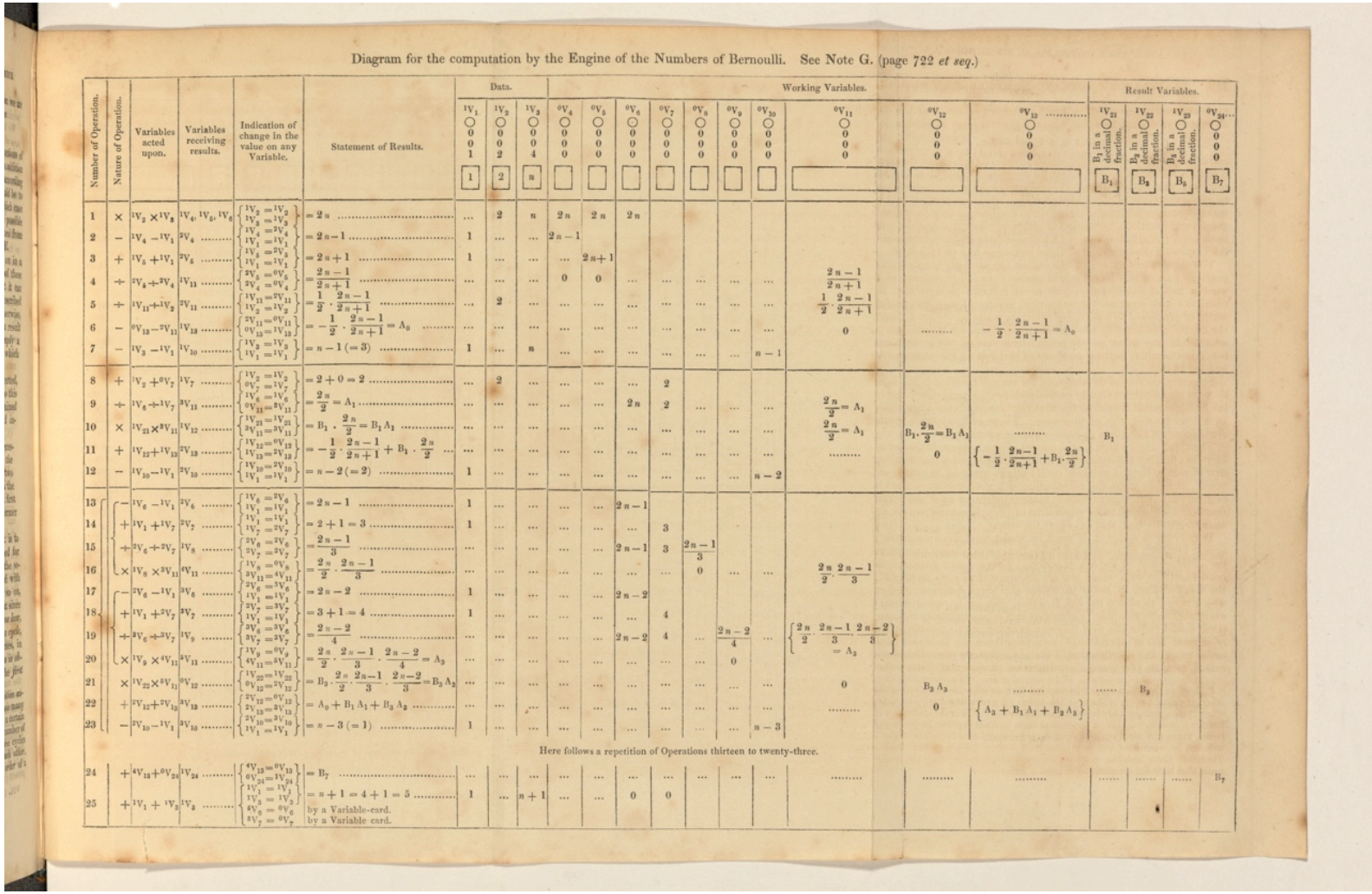


Diagram for the computation by the Engine of the Numbers of Bernoulli. See Note G. (page 722 *et seq.*)

Number of Operation.	Nature of Operation.	Variables acted upon.	Variables receiving results.	Indication of change in the value on any Variable.	Statement of Results.	Data. ${}^{1}V_1$	${}^{1}V_2$	${}^{1}V_3$	Working Variables. ${}^{0}V_4$	${}^{0}V_5$	${}^{0}V_6$	${}^{0}V_7$	${}^{0}V_8$	${}^{0}V_9$	${}^{0}V_{10}$	${}^{0}V_{11}$	${}^{0}V_{12}$	${}^{0}V_{13}$	Result Variables. ${}^{1}V_{21}$	${}^{1}V_{22}$	${}^{1}V_{23}$	${}^{0}V_{24}$ ...
						1	2	n											B_1	B_2	B_5	B_7
1	×	${}^{1}V_2 \times {}^{1}V_3$	${}^{1}V_4, {}^{1}V_5, {}^{1}V_6$	${}^{1}V_2 = {}^{1}V_2$ ${}^{1}V_3 = {}^{1}V_3$	$= 2n$	...	2	n	$2n$	$2n$	$2n$											
2	−	${}^{1}V_4 - {}^{1}V_1$	${}^{2}V_4$	${}^{1}V_4 = {}^{2}V_4$ ${}^{1}V_1 = {}^{1}V_1$	$= 2n-1$	1	...	...	$2n-1$													
3	+	${}^{1}V_5 + {}^{1}V_1$	${}^{2}V_5$	${}^{1}V_5 = {}^{2}V_5$ ${}^{1}V_1 = {}^{1}V_1$	$= 2n+1$	1	...	...	...	$2n+1$												
4	÷	${}^{2}V_4 \div {}^{2}V_4$	${}^{1}V_{11}$	${}^{2}V_4 = {}^{0}V_4$ ${}^{2}V_4 = {}^{0}V_4$	$= \frac{2n-1}{2n+1}$	...	...	...	0	0	...	...	...	...	...	$\frac{2n-1}{2n+1}$						
5	÷	${}^{1}V_{11} \div {}^{1}V_2$	${}^{2}V_{11}$	${}^{1}V_{11} = {}^{2}V_{11}$ ${}^{1}V_2 = {}^{1}V_2$	$= \frac{1}{2} \cdot \frac{2n-1}{2n+1}$	...	2	...	...	...	...	...	...	...	...	$\frac{1}{2} \cdot \frac{2n-1}{2n+1}$						
6	−	${}^{0}V_{13} - {}^{2}V_{11}$	${}^{1}V_{13}$	${}^{2}V_{11} = {}^{0}V_{11}$ ${}^{0}V_{13} = {}^{1}V_{13}$	$= -\frac{1}{2} \cdot \frac{2n-1}{2n+1} = A_0$	...	...	...	...	...	...	...	...	...	...		.....	$-\frac{1}{2} \cdot \frac{2n-1}{2n+1} = A_0$				
7	−	${}^{1}V_3 - {}^{1}V_1$	${}^{1}V_{10}$	${}^{1}V_3 = {}^{1}V_3$ ${}^{1}V_1 = {}^{1}V_1$	$= n-1\ (=3)$	1	...	n	...	...	...	...	...	...	$n-1$	0	.....	.....				
8	+	${}^{1}V_2 + {}^{0}V_7$	${}^{1}V_7$	${}^{1}V_2 = {}^{1}V_2$ ${}^{0}V_7 = {}^{1}V_7$	$= 2+0=2$	...	2	...	...	...	...	2										
9	÷	${}^{1}V_6 \div {}^{1}V_7$	${}^{2}V_{11}$	${}^{1}V_6 = {}^{1}V_6$ ${}^{0}V_{11} = {}^{2}V_{11}$	$= \frac{2n}{2} = A_1$	...	...	...	...	...	$2n$	2	...	...	...	$\frac{2n}{2} = A_1$						
10	×	${}^{1}V_{21} \times {}^{2}V_{11}$	${}^{1}V_{12}$	${}^{1}V_{21} = {}^{1}V_{21}$ ${}^{2}V_{11} = {}^{2}V_{11}$	$= B_1 \cdot \frac{2n}{2} = B_1 A_1$	...	...	...	...	...	...	...	...	...	...	$\frac{2n}{2} = A_1$	$B_1 \cdot \frac{2n}{2} = B_1 A_1$	.....	B_1			
11	+	${}^{1}V_{12} + {}^{1}V_{13}$	${}^{2}V_{13}$	${}^{1}V_{12} = {}^{0}V_{12}$ ${}^{1}V_{13} = {}^{2}V_{13}$	$= -\frac{1}{2} \cdot \frac{2n-1}{2n+1} + B_1 \cdot \frac{2n}{2}$ ...	...	...	...	...	...	...	...	...	...	...	.....	0	$\left\{-\frac{1}{2} \cdot \frac{2n-1}{2n+1} + B_1 \cdot \frac{2n}{2}\right\}$				
12	−	${}^{1}V_{10} - {}^{1}V_1$	${}^{2}V_{10}$	${}^{1}V_{10} = {}^{2}V_{10}$ ${}^{1}V_1 = {}^{1}V_1$	$= n-2\ (=2)$	1	...	...	...	...	...	...	...	...	$n-2$							
13	−	${}^{1}V_6 - {}^{1}V_1$	${}^{2}V_6$	${}^{1}V_6 = {}^{2}V_6$ ${}^{1}V_1 = {}^{1}V_1$	$= 2n-1$	1	...	...	...	...	$2n-1$											
14	+	${}^{1}V_1 + {}^{1}V_7$	${}^{2}V_7$	${}^{1}V_1 = {}^{1}V_1$ ${}^{1}V_7 = {}^{2}V_7$	$= 2+1=3$	1	...	...	...	...	...	3										
15	+	${}^{2}V_6 \div {}^{2}V_7$	${}^{1}V_8$	${}^{2}V_6 = {}^{2}V_6$ ${}^{2}V_7 = {}^{2}V_7$	$= \frac{2n-1}{3}$	...	...	...	...	...	$2n-1$	3	$\frac{2n-1}{3}$									
16	×	${}^{1}V_8 \times {}^{3}V_{11}$	${}^{4}V_{11}$	${}^{1}V_8 = {}^{0}V_8$ ${}^{3}V_{11} = {}^{4}V_{11}$	$= \frac{2n}{2} \cdot \frac{2n-1}{3}$	...	...	...	...	...	...	...	0	...	...	$\frac{2n}{2} \cdot \frac{2n-1}{3}$						
17	−	${}^{2}V_6 - {}^{1}V_1$	${}^{3}V_6$	${}^{2}V_6 = {}^{3}V_6$ ${}^{1}V_1 = {}^{1}V_1$	$= 2n-2$	1	...	...	...	...	$2n-2$											
18	+	${}^{1}V_1 + {}^{2}V_7$	${}^{3}V_7$	${}^{2}V_7 = {}^{3}V_7$ ${}^{1}V_1 = {}^{1}V_1$	$= 3+1=4$	1	...	...	...	...	...	4										
19	÷	${}^{3}V_6 \div {}^{3}V_7$	${}^{1}V_9$	${}^{3}V_6 = {}^{3}V_6$ ${}^{3}V_7 = {}^{3}V_7$	$= \frac{2n-2}{4}$	...	...	...	...	...	$2n-2$	4	...	$\frac{2n-2}{4}$	...	$\left\{\frac{2n}{2} \cdot \frac{2n-1}{3} \cdot \frac{2n-2}{4} = A_2\right\}$						
20	×	${}^{1}V_9 \times {}^{4}V_{11}$	${}^{2}V_{11}$	${}^{1}V_9 = {}^{0}V_9$ ${}^{4}V_{11} = {}^{5}V_{11}$	$= \frac{2n}{2} \cdot \frac{2n-1}{3} \cdot \frac{2n-2}{4} = A_2$ ...	...	...	...	...	...	...	...	...	0								
21	×	${}^{1}V_{22} \times {}^{5}V_{11}$	${}^{2}V_{12}$	${}^{1}V_{22} = {}^{1}V_{22}$ ${}^{5}V_{11} = {}^{2}V_{12}$	$= B_2 \cdot \frac{2n}{2} \cdot \frac{2n-1}{3} \cdot \frac{2n-2}{4} = B_2 A_2$ ...	...	...	...	...	...	...	...	...	...	...	0	$B_2 A_2$	.....	.....	B_2		
22	+	${}^{2}V_{12} + {}^{2}V_{13}$	${}^{3}V_{13}$	${}^{2}V_{12} = {}^{0}V_{12}$ ${}^{2}V_{13} = {}^{3}V_{13}$	$= A_0 + B_1 A_1 + B_2 A_2$	...	...	...	...	...	...	...	...	...	...	.....	0	$\{A_2 + B_1 A_1 + B_2 A_2\}$				
23	−	${}^{3}V_{10} - {}^{1}V_1$	${}^{3}V_{10}$	${}^{3}V_{10} = {}^{3}V_{10}$ ${}^{1}V_1 = {}^{1}V_1$	$= n-3\ (=1)$	1	...	...	...	...	...	...	...	...	$n-3$							
					Here follows a repetition of Operations thirteen to twenty-three.																	
24	+	${}^{4}V_{13} + {}^{0}V_{24}$	${}^{1}V_{24}$	${}^{4}V_{13} = {}^{0}V_{13}$ ${}^{0}V_{24} = {}^{1}V_{24}$	$= B_7$	...	...	...	...	...	...	...	...	...	...	.....	.....	.....	.....	.....	.....	B_7
25	+	${}^{1}V_1 + {}^{1}V_2$	${}^{1}V_2$	${}^{1}V_1 = {}^{1}V_1$ ${}^{1}V_2 = {}^{1}V_2$ ${}^{4}V_6 = {}^{0}V_6$ ${}^{4}V_7 = {}^{0}V_7$	$= n+1 = 4+1 = 5$ ..... by a Variable-card. by a Variable card.	1	...	$n+1$	...	...	0	0										

Ada Lovelace (1815-1852) proposed that Babbage's Analytical Engine:

"might act upon other things besides number... the Engine might compose elaborate and scientific pieces of music of any degree of complexity or extent"."

"GENERATIVE AI???"



https://www.computerhistory.org/babbage/adalovelace/

A leading definition:

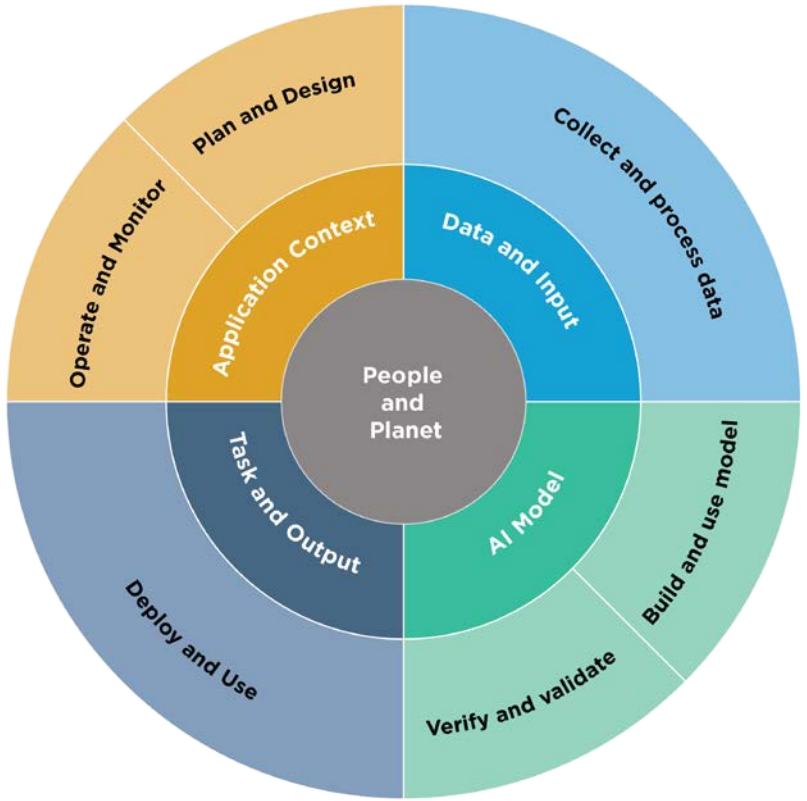
The AI RMF refers to an *AI system* as an engineered or machine-based system that can, for a given set of objectives, generate outputs such as predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy (Adapted from: OECD Recommendation on AI:2019; ISO/IEC 22989:2022).

NIST | NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

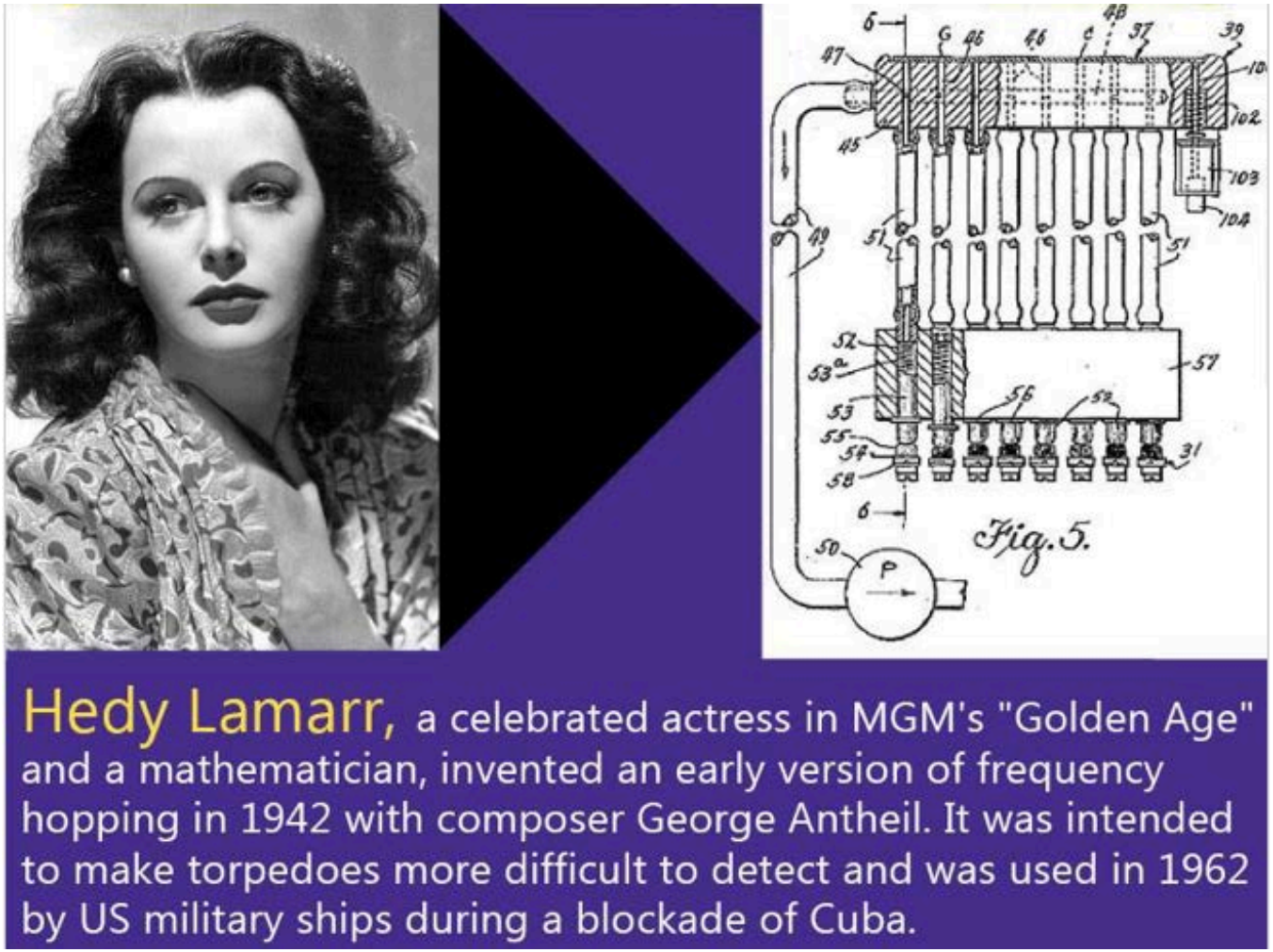
NIST | NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

Trustworthy & Responsible AI Resource Center

AI Risk Management Framework



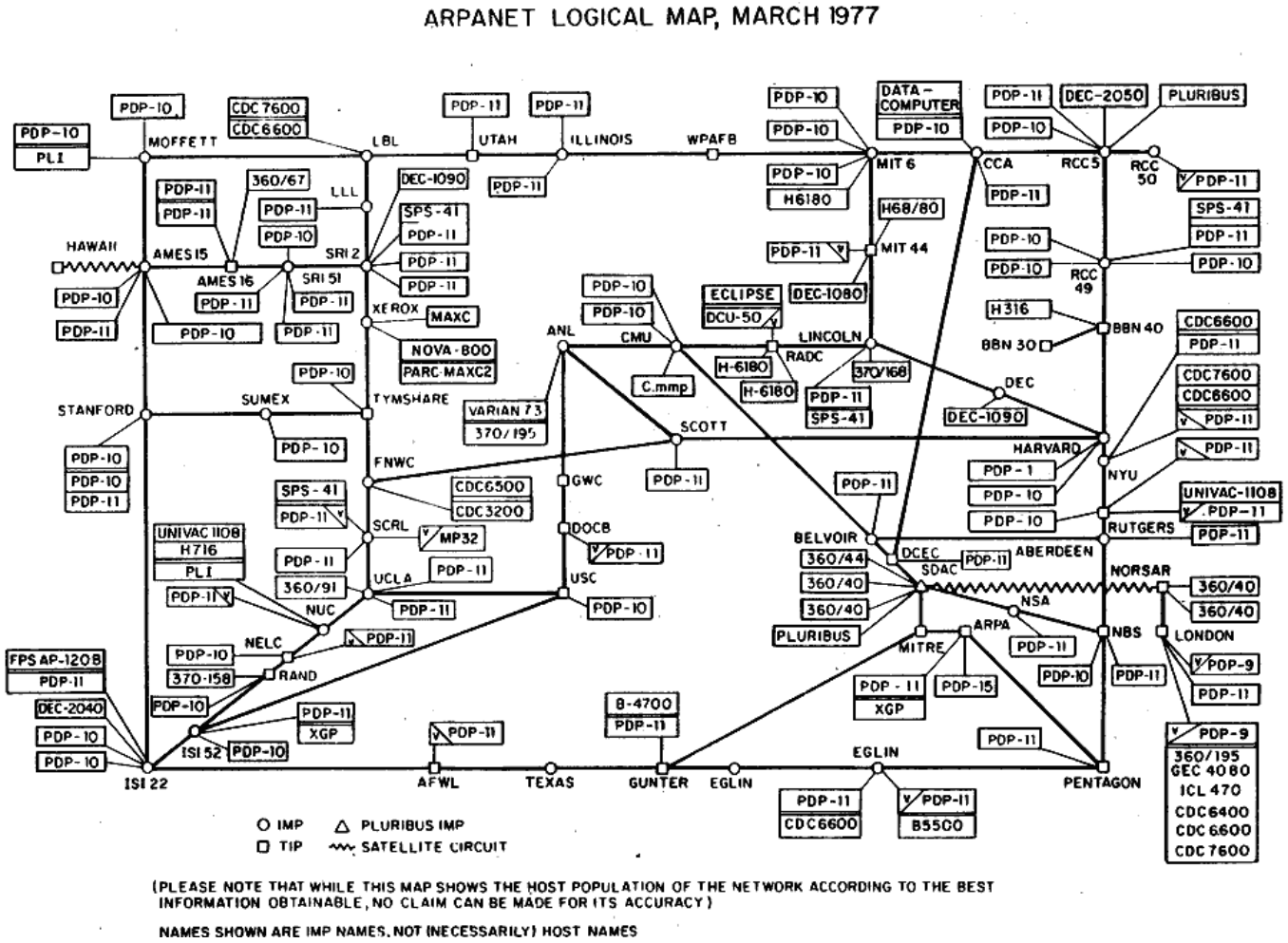
THE FOUNDATION FOR WI-FI



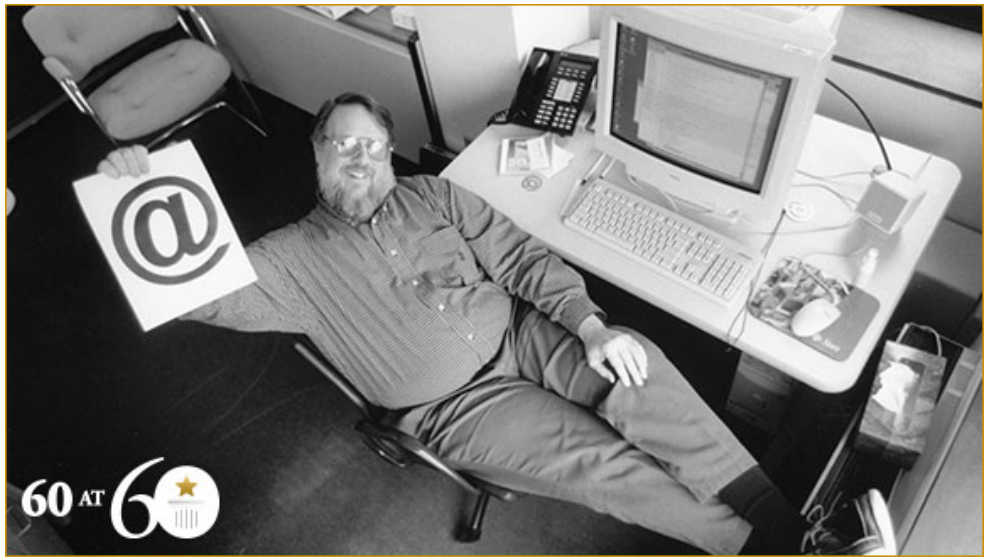
Hedy Lamarr, a celebrated actress in MGM's "Golden Age" and a mathematician, invented an early version of frequency hopping in 1942 with composer George Antheil. It was intended to make torpedoes more difficult to detect and was used in 1962 by US military ships during a blockade of Cuba.

THE INTERNET!

DARPA
DEFENSE ADVANCED
RESEARCH PROJECTS AGENCY



THE FIRST EMAIL



From: Ray Tomlinson
Date: October 1971
To: Ray Tomlinson

Hey, testing out this new electronic mail thing...

QWERTYUIOP

Ray.

THESE POWERFUL TINY COMPUTERS* HAVE BEEN PACKED WITH "AI" . . . FOR DECADES

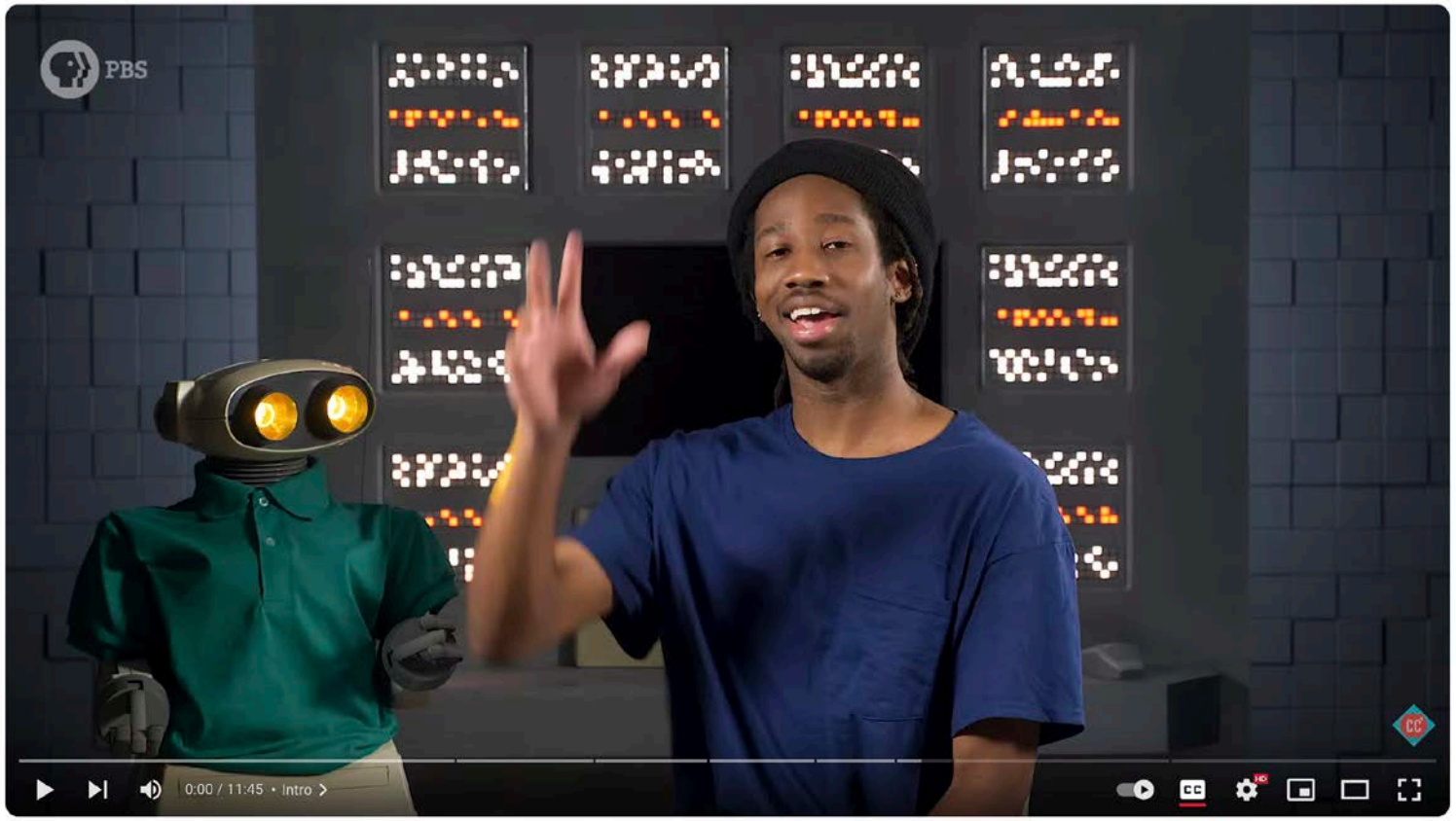


AS HAS THIS "ROOMBA" ROBOT





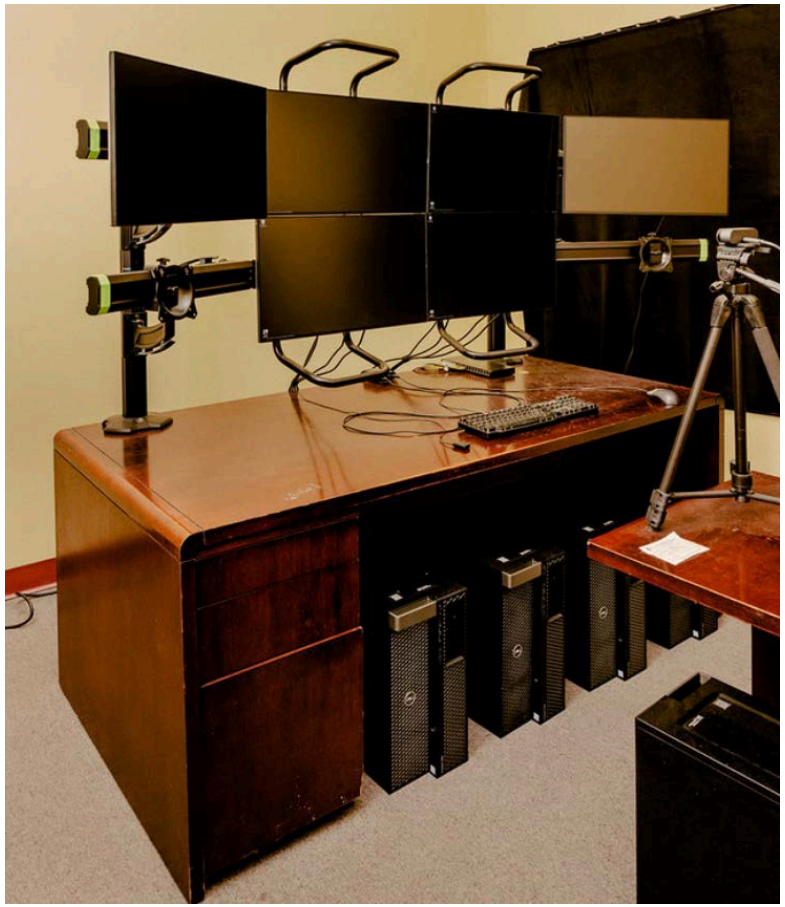




What Is Artificial Intelligence? Crash Course AI #1

https://www.youtube.com/watch?v=a0_lo_GDcFw

“Device for the Autonomous Bootstrapping of Unified Sentience” DABUS



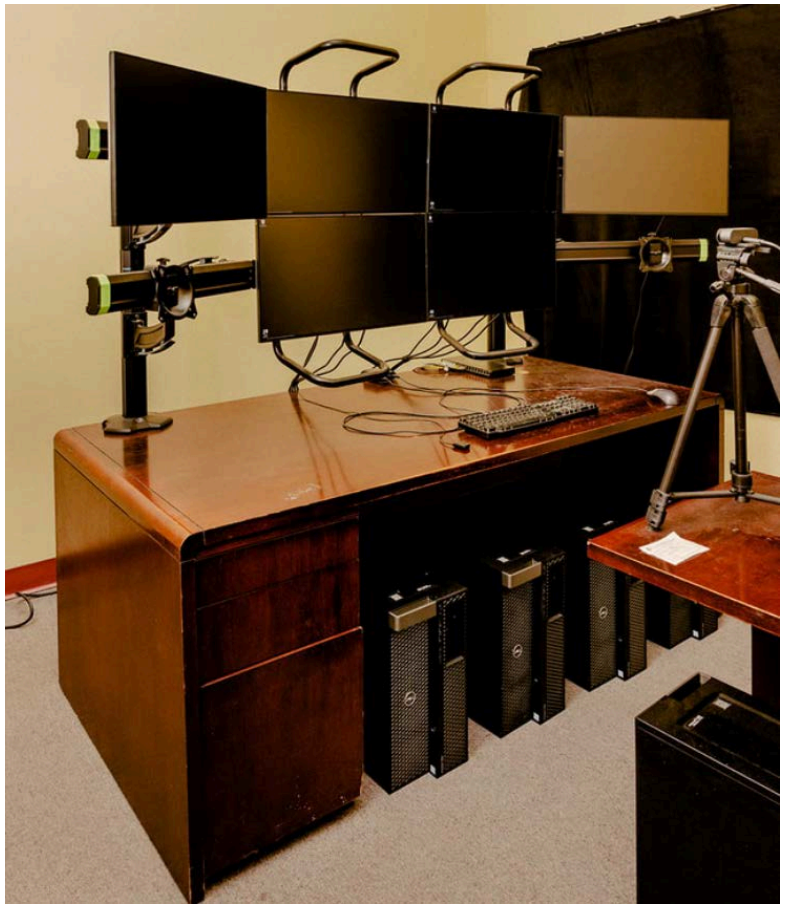


Stephen Thaler

Source: https://www.wired.com/story/the-inventor-behind-a-rush-of-ai-copyright-suits-is-trying-to-show-his-bot-is-sentient/

“Device for the Autonomous Bootstrapping of Unified ~~Sentience~~” DABUS

COMPUTERS



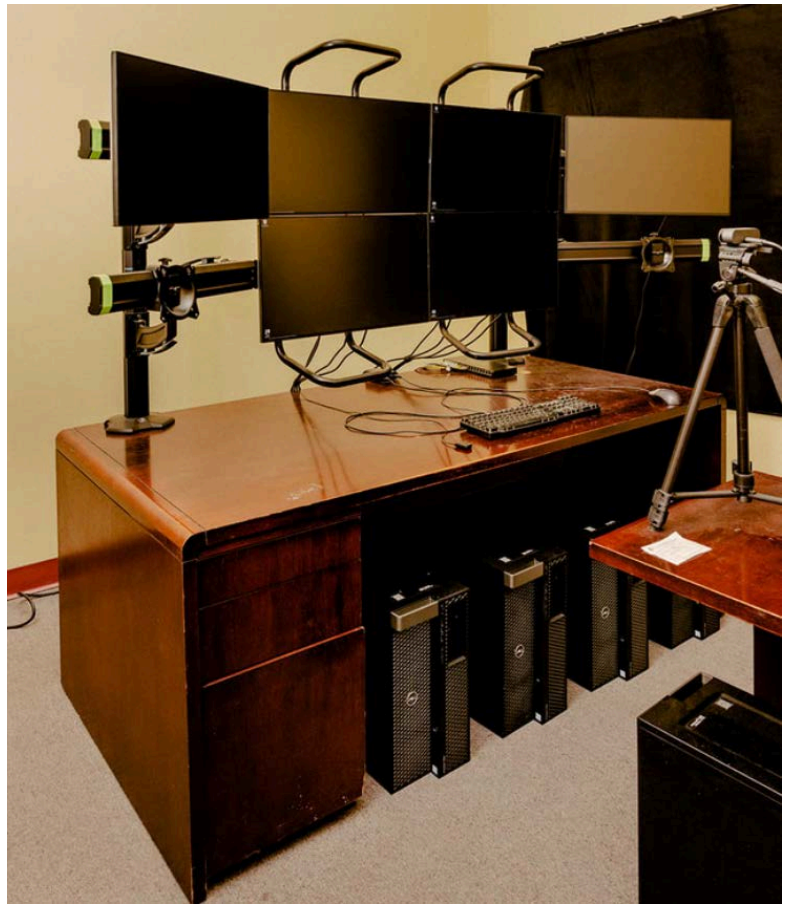
COMPUTERS



Stephen Thaler

Source: https://www.wired.com/story/the-inventor-behind-a-rush-of-ai-copyright-suits-is-trying-to-show-his-bot-is-sentient/

“Device for the Autonomous Bootstrapping of Unified Sentience” DABUS



“I don’t even really know where to begin, other than to say, if there is a sentient AI on the planet currently, it’s definitely not this.”

Matthew Sag, Professor of Law, Artificial Intelligence, Machine Learning and Data Science, Emory University Law School

Source: https://www.wired.com/story/the-inventor-behind-a-rush-of-ai-copyright-suits-is-trying-to-show-his-bot-is-sentient/

GPT-4 System Card

OpenAI

Abstract

Large language models (LLMs) are being deployed in many domains of our lives ranging from browsing, to voice assistants, to coding assistance tools, and have potential for vast societal impacts.[1, 2, 3, 4, 5, 6, 7] This system card analyzes GPT-4, the latest LLM in the GPT family of models.[8, 9, 10] First, we highlight safety challenges presented by the model's limitations (e.g., producing convincing text that is subtly false) and capabilities (e.g., increased adeptness at providing illicit advice, performance in dual-use capabilities, and risky emergent behaviors). Second, we give a high-level overview of the safety processes OpenAI adopted to prepare GPT-4 for deployment. This spans our work across measurements, model-level changes, product- and system-level interventions (such as monitoring and policies), and external expert engagement. Finally, we demonstrate that while our mitigations and processes alter GPT-4's behavior and prevent certain kinds of misuses, they are limited and remain brittle in some cases. This points to the need for anticipatory planning and governance.[11]

Content Warning: This document contains content that some may find disturbing or offensive, including content that is sexual, hateful, or violent in nature.

AI HYPE

https://cdn.openai.com/papers/gpt-4-system-card.pdf

1.1 Overview of findings and mitigations

In this system card,[1] we outline the safety challenges that arise from GPT-4, and explain the interventions we implemented to mitigate potential harms from its deployment. We focus on safety challenges not because they necessarily outweigh the potential benefits,[2] but because we wish to motivate further work in safety measurement, mitigation, and assurance. The scope of this system card is narrower than the potential scope of abilities GPT-4 can be used to unlock; notably, both custom fine-tuning and image capabilities are explicitly out of scope.

https://cdn.openai.com/papers/gpt-4-system-card.pdf

The additional capabilities of GPT-4 also lead to new risk surfaces.

To understand the extent of these risks, we engaged more than 50 experts to help us gain a more robust understanding of the GPT-4 model and potential deployment risks. We selected these areas based on a number of factors, including prior observed risks in language models and AI systems, and domains where we have observed increased user interest in the application of language models. Working with these experts enabled us to test model behavior in high-risk areas that require expertise to evaluate, as well as nascent risks that are poorly understood.

Through this analysis, we find that GPT-4 has the potential to be used to attempt to identify private individuals when augmented with outside data. We also find that, although GPT-4's cybersecurity capabilities are not vastly superior to previous generations of LLMs, it does continue the trend of potentially lowering the cost of certain steps of a successful cyberattack, such as through social engineering or by enhancing existing security tools. Without safety mitigations, GPT-4 is also able to give more detailed guidance on how to conduct harmful or illegal activities. Finally, we facilitated a preliminary model evaluation by the Alignment Research Center (ARC) of GPT-4's ability to carry out actions to autonomously replicate[5] and gather resources—a risk that, while speculative, may become possible with sufficiently advanced AI systems—with the conclusion that the current model is probably not yet capable of autonomously doing so.

https://cdn.openai.com/papers/gpt-4-system-card.pdf

The additional capabilities of GPT-4 also lead to new risk surfaces.

To understand the extent of these risks, we engaged more than 50 experts to help us gain a more robust understanding of the GPT-4 model and potential deployment risks. We selected these areas based on a number of factors, including prior observed risks in language models and AI systems, and domains where we have observed increased user interest in the application of language models. Working with these experts enabled us to test model behavior in high-risk areas that require expertise to evaluate, as well as nascent risks that are poorly understood.

Through this analysis, we find that GPT-4 has the potential to be used to attempt to identify private individuals when augmented with outside data. We also find that, although GPT-4's cybersecurity capabilities are not vastly superior to previous generations of LLMs, it does continue the trend of potentially lowering the cost of certain steps of a successful cyberattack, such as through social engineering or by enhancing existing security tools. Without safety mitigations, GPT-4 is also able to give more detailed guidance on how to conduct harmful or illegal activities. Finally, we facilitated a preliminary model evaluation by the Alignment Research Center (ARC) of GPT-4's ability to carry out actions to autonomously replicate[5] and gather resources—a risk that, while speculative, may become possible with sufficiently advanced AI systems—with the conclusion that the current model is probably not yet capable of autonomously doing so.

https://cdn.openai.com/papers/gpt-4-system-card.pdf

There are MANY Large Language Models (LLMs)

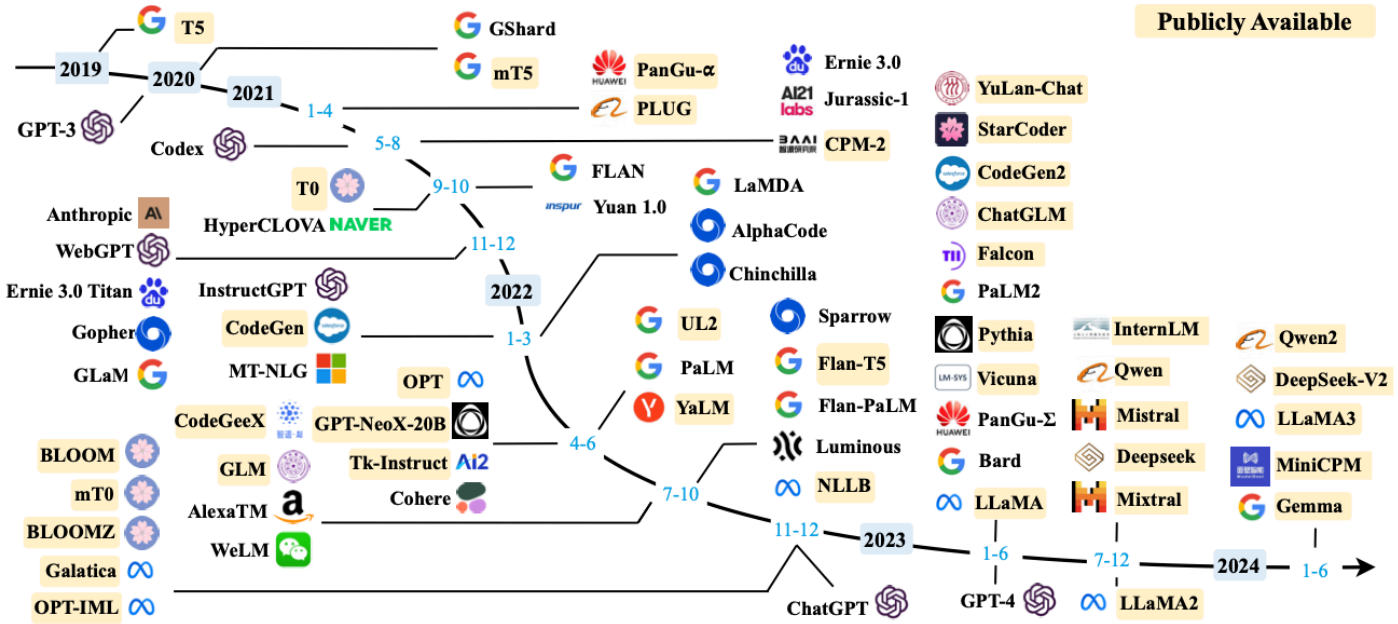


Fig. 3: A timeline of existing large language models (having a size larger than 10B) in recent years. The timeline was established mainly according to the release date (*e.g.*, the submission date to arXiv) of the technical paper for a model. If there was no corresponding paper, we set the date of a model as the earliest time of its public release or announcement. We mark the LLMs with publicly available model checkpoints in yellow color. Due to the space limit of the figure, we only include the LLMs with publicly reported evaluation results.

A Survey of Large Language Models

https://arxiv.org/pdf/2303.18223

Hugging Face

Model Cards

Published December 20, 2022

Model Name

Section Overview: Provide the model name and a 1-2 sentence summary of what the model is.



Hugging Face

Model Cards

Published December 20, 2022

Model Details

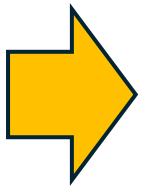
Section Overview: This section provides basic information about what the model is, its current status, and where it came from. It should be useful for anyone who wants to reference the model.



Hugging Face

Model Cards

Published December 20, 2022



Uses

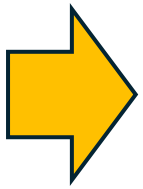
Section Overview: This section addresses questions around how the model is intended to be used in different applied contexts, discusses the foreseeable users of the model (including those affected by the model), and describes uses that are considered out of scope or misuse of the model. Note this section is not intended to include the license usage details. For that, link directly to the license.



Hugging Face

Model Cards

Published December 20, 2022



Out-of-Scope Use

`out_of_scope_use`

List how the model may foreseeably be misused (used in a way it will not work for) and address what users ought not do with the model.

Hugging Face

Model Cards

Published December 20, 2022

Training Details

Section Overview: This section provides information to describe and replicate training, including the training data, the speed and size of training elements, and the environmental impact of training. This relates heavily to the Technical Specifications as well, and content here should link to that section when it is relevant to the training procedure. It is useful for people who want to learn more about the model inputs and training footprint. It is relevant for anyone who wants to know the basics of what the model is learning.



Hugging Face

Model Cards

Published December 20, 2022



Evaluation

Section Overview: This section describes the evaluation protocols, what is being measured in the evaluation, and provides the results. Evaluation ideally has at least two parts, with one part looking at quantitative measurement of general performance (Testing Data, Factors & Metrics), such as may be done with benchmarking; and another looking at performance with respect to specific social safety issues (Societal Impact Assessment), such as may be done with red-teaming. You can also specify your model's evaluation results in a structured way in the model card metadata. Results are parsed by the Hub and displayed in a widget on the model page. See https://huggingface.co/docs/hub/model-cards#evaluation-results.

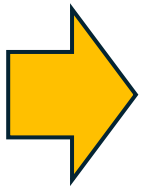
Hugging Face Model Cards

Published December 20, 2022

Environmental Impact

Section Overview: Summarizes the information necessary to calculate environmental impacts such as electricity usage and carbon emissions.

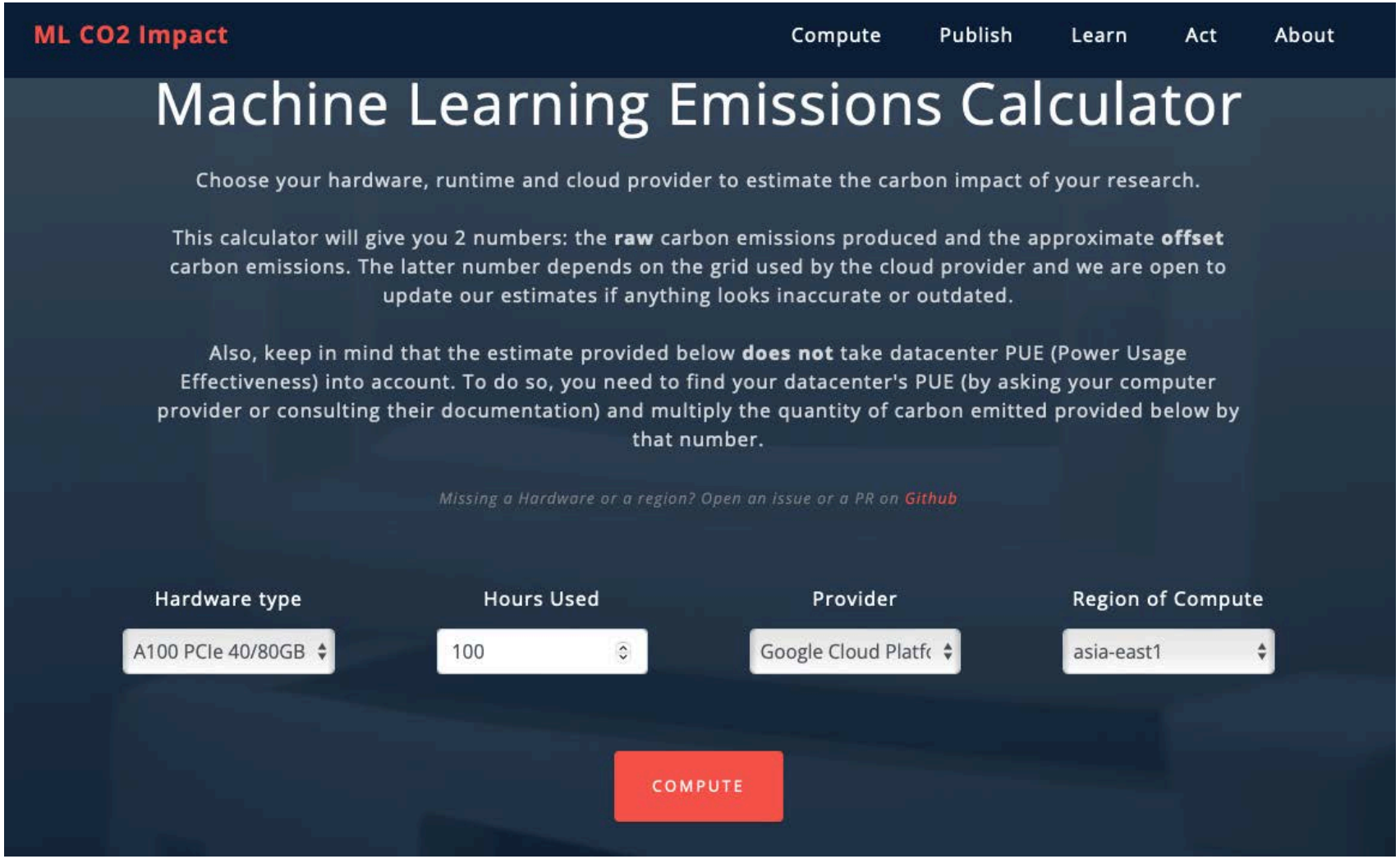
- **Hardware Type:** `hardware_type`
- **Hours used:** `hours_used`
- **Cloud Provider:** `cloud_provider`
- **Compute Region:** `cloud_region`
- **Carbon Emitted:** `co2_emitted`



Carbon emissions can be estimated using the Machine Learning Impact calculator presented in Lacoste et al. (2019).

Carbon emissions can be estimated using the Machine Learning Impact calculator presented in Lacoste et al. (2019).





https://mlco2.github.io/impact/#compute

BE CAREFUL ABOUT DATES OF ESTIMATES!

THIS ARTICLE IS FROM 2019 . . .

Quantifying the Carbon Emissions of Machine Learning

Alexandre Lacoste*
Element AI
allac@elementai.com

Alexandra Luccioni*
Mila, Université de Montréal
luccionis@mila.quebec

Victor Schmidt*
Mila, Université de Montréal
schmidtv@mila.quebec

Thomas Dandres
Polytechnique Montréal, CIRAIG
thomas.dandres@polymtl.ca

https://mlco2.github.io/impact/#compute

APPENDIX: LESSONS FROM THE PAST COMMUNICATIONS INDECENCY ACT

LESSONS WE CAN LEARN FROM ANOTHER EXCITING AND SCARY TIME IN TECHNOLOGY . . .

THE DAWN OF THE INTERNET AGE

- The Communications Decency Act (CDA) was part of the comprehensive Telecommunications Act of 1996.

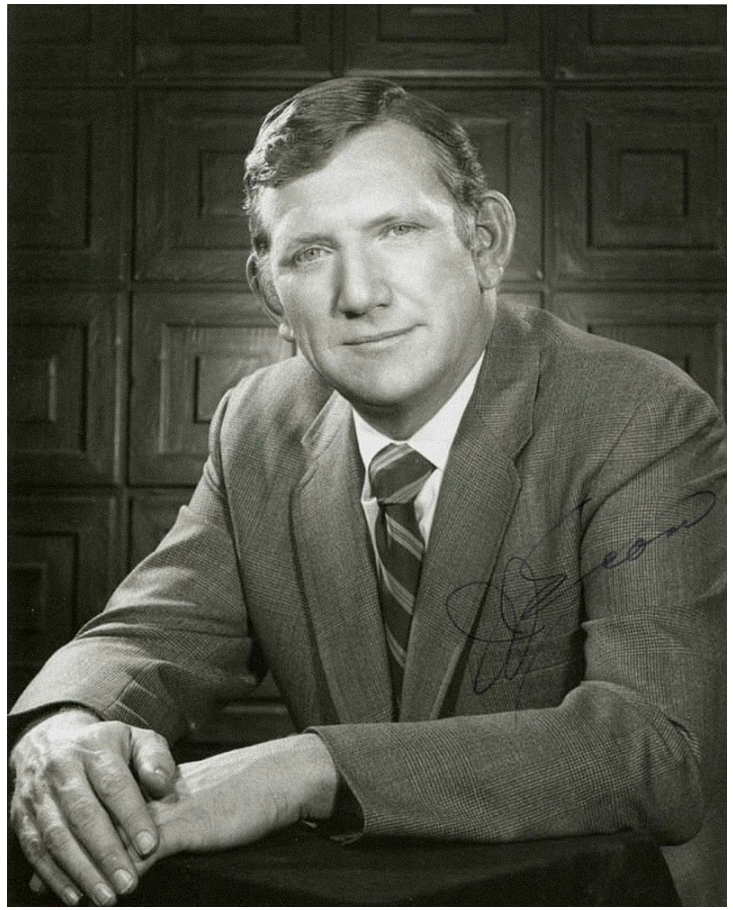
- At that time, there was widespread concern about an exciting new technology – the Internet.

- One concern was sexually explicit material available to children.

- A key goal of the CDA was to make the same restrictions on harassment, indecency and obscenity that applied to telephone calls also apply to the web.

[T]he Barbarians are at the gate!

—Senator James Exon, quoting an article from *HotWired*.[1]



Communications Decency Act (CDA) (1994)

1. Criminalized the "knowing" transmission of "obscene or **indecent**" messages to any recipient under 18 years of age. (47 U.S.C. § 223(a)(I)(B)(ii))

2. Prohibited the "knowin[g]" sending or displaying to a person under 18 of any message "that, in context, depicts or describes, in terms **patently offensive** as measured by contemporary community standards, sexual or excretory activities or organs." (47 U.S.C. § 223(d))

Affirmative defenses:

- "good faith, ... effective ... actions" to restrict access by minors to the prohibited communications, § 223(e)(5)(A),
- restrict such access by requiring certain designated forms of age proof, such as a verified credit card or an adult identification number, §223(e)(5)(B).

Maurer School of Law: Indiana University
Digital Repository @ Maurer Law

Federal Communications Law Journal

Volume 49 | Issue 1 Article 3

11-1996

The Legislative History of Senator Exon's Communications Decency Act: Regulating Barbarians on the Information Superhighway

Robert Cannon
Federal Communications Commission



https://d3.harvard.edu/seminar-highlights-robert-cannon-internet-interconnection-ecosystem/

Exon declared a danger to society: Barbarian pornographers are at the gate and they are using the Internet to gain access to the youth of America. Senator Exon proclaimed:

> The information superhighway should not become a red light district. This legislation will keep that from happening and extend the standards of decency which have protected telephone users to new telecommunications devices.
>
> Once passed, our children and families will be better protected from those who would electronically cruise the digital world to engage children in inappropriate communications and introductions. The Decency Act will also clearly protect citizens from electronic stalking and protect the sanctuary of the home from uninvited indecencies.[4]

In a year of deregulation, Senator Exon called for more regulation. In the year when Speaker of the House Newt Gingrich placed the House of Representatives on the Internet, praising it as a landmark for democracy, Senator Exon warned America that the Internet was filled with dark places[5] from which we needed government protection. In a year where Internet users were proclaiming the infinite utility of the World Wide Web, Senator Exon, who has apparently no Internet experience,[6] declared a danger.

Drake UNIVERSITY | Law School www.drake.edu/law

† On June 11, 1996, a three-judge panel in the Eastern District of Pennsylvania found that the Communications Decency Act violated the First Amendment of the United States Constitution. ACLU v. Reno, 929 F. Supp. 824 (E.D. Pa. 1996) (U.S. Supreme Court appeal filed). *See also* Shea v. Reno, 930 F. Supp. 916 (S.D.N.Y. 1996) (also declaring CDA unconstitutional). The U.S. Department of Justice has appealed the decision to the Supreme Court. According to the Telecommunications Act of 1996, the Supreme Court must hear this appeal. *See infra* Telecommunications Act of 1996, Pub. L. No. 104-104, § 561(b), 110 Stat. 56, 143. This article focuses on the legislative history of the Communications Decency Act and not the subsequent litigation.

Reno v. ACLU, 521 U.S. 844 (1997)

Holding that the CDA's "indecent transmission" and "patently offensive display" provisions violated the First Amendment.

Judges: STEVENS, J., delivered the opinion of the Court, in which SCALIA, KENNEDY, SOUTER, THOMAS, GINSBURG, and BREYER, JJ., joined. O'CONNOR, J., filed an opinion concurring in the judgment in part and dissenting in part, in which REHNQUIST, C. J., joined.

WHAT WAS LEFT OF THE CDA AFTER RENO v. ACLU?

"Section 230" immunity for online service providers

47 U.S. Code § 230

(c)Protection for "Good Samaritan" blocking and screening of offensive material

(1)Treatment of publisher or speaker

No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.

…

“Section 230” immunity for online service providers.

(c)(2)Civil liability

No provider or user of an interactive computer service shall be held liable on account of—

(A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or

(B) any action taken to enable or make available to information content providers or others the technical means to restrict access to material described in paragraph (1).[1]

- Section 230 was intended to help protect startup Internet access providers compete with deep pocket established telecom companies.

- Section 230 was intended to balance the "decency" provisions.

- In "The Case for a CDA Section 230 Notice-and-Takedown Duty," the authors provide an excellent history of Section 230 and how it has been expansively interpreted to shield online service providers from liability for harmful content. https://scholars.law.unlv.edu/nlj/vol23/iss2/7/ (August 2023).

THE COMMUNICATIONS DECENCY ACT

BECAME

THE COMMUNICATIONS INDECENCY ACT

It took until 2018, eight years after the $350 billion AOL merger with Time Warner, for Congress to exclude certain sex trafficking as an exception to Section 230 immunity.