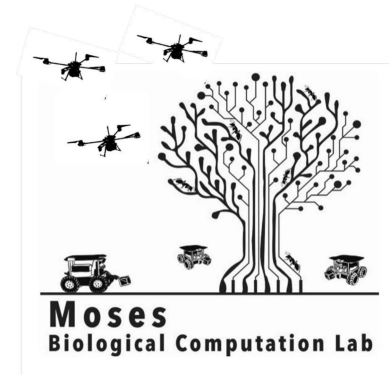
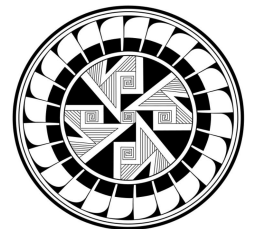


# Overview of Sectors Impacted by Artificial Intelligence Use



New Mexico Legislature Science Technology &  
Telecommunications Committee  
May 1, 2024

Melanie Moses  
UNM Professor of Computer Science  
External Faculty, Santa Fe Institute  
<https://moseslab.cs.unm.edu/>



# Bottom Line Up Front

## Regulatory priorities

- Deep fakes, beyond elections
- AI/algorithms in healthcare
- Transparency & equity in high stakes decision-making
- Protect creative works
- Establish study group to examine impact and equity of AI/algorithms used by NM state government

## Investment priorities

- New Mexicans needs
  - AI literacy
  - To be engaged in shaping the future of AI
- Trustworthy AI for science and society
- Partnerships among NM Colleges, Universities & National Labs catalyze innovation, education & economic growth

Major investments from industry & governments; uncertainty about what's next

# What I'll talk about today

- Refresher:
  - Evolving perspective on AI
  - How does generative AI work?
- What we're doing at UNM
- What to expect from AI
- What you can do to shape the future of AI for New Mexico

# My (evolving) thoughts on AI

- I study bio-inspired robotics and complex systems
- I think of AI as a complex adaptive system that we are integrating into our complex socio-technical system
- This is exciting and scary!
  - AI will help solve scientific problems
  - Current AI encodes biases, promotes surveillance & concentrates power
  - AI codes even better than it writes
  - How will humans co-evolve with AI?

Southwestern Indian Polytechnic Institute (SIPI)  
1<sup>st</sup> place Swarmathon 2018



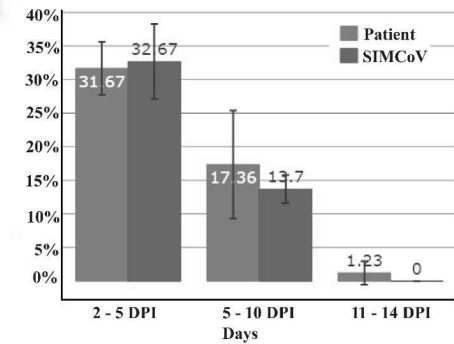
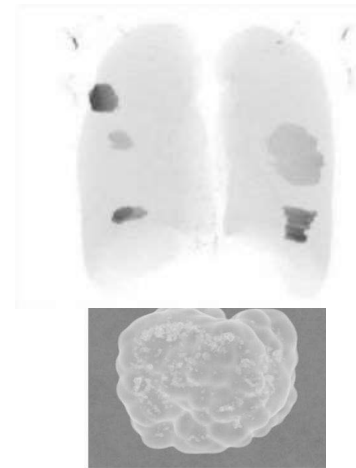
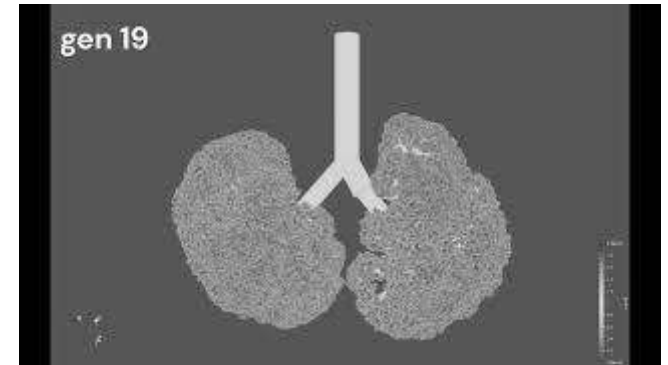
## My (evolving) thoughts on AI

- I study bio-inspired robotics and complex systems
- I think of AI as a complex adaptive system that we are integrating into our complex socio-technical system
- This is exciting and scary!
  - AI will help solve scientific problems
  - Current AI encodes biases, promotes surveillance & concentrates power
  - AI codes even better than it writes
  - How will humans co-evolve with AI?



## My (evolving) thoughts on AI

- I study bio-inspired robotics and complex systems
- I think of AI as a complex adaptive system that we are integrating into our complex socio-technical system
- This is exciting and scary!
  - AI will help solve scientific problems
  - Current AI encodes biases, promotes surveillance & concentrates power
  - AI codes even better than it writes
  - How will humans co-evolve with AI?



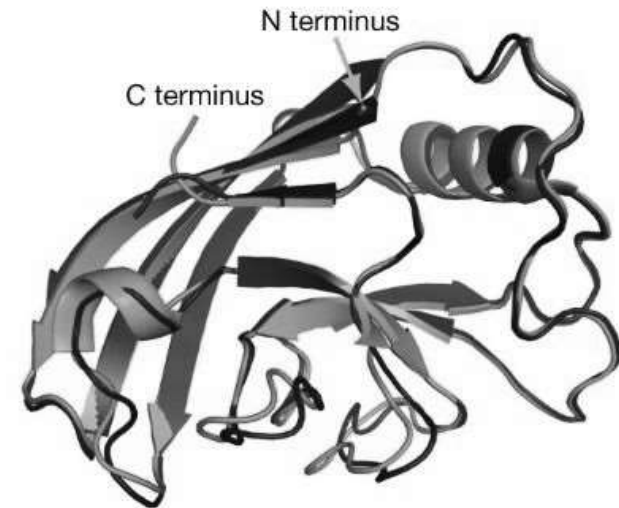
Akil Andrews



Humayra Tasnim

## My (evolving) thoughts on AI

- I study bio-inspired robotics and complex systems
- I think of AI as a complex adaptive system that we are integrating into our complex socio-technical system
- This is exciting and scary!
  - AI will help solve scientific problems
  - Current AI encodes biases, promotes surveillance & concentrates power
  - AI codes even better than it writes
  - How will humans co-evolve with AI?



Highly accurate protein structure prediction with AlphaFold, *Nature* 2021

MIT  
Technology  
Review

[Featured](#) [Topics](#) [Newsletters](#) [Events](#) [Podcasts](#)

[SIGN IN](#)

[SUBSCRIBE](#)

CLIMATE CHANGE AND ENERGY

**How robots and AI are helping develop better batteries**

## My (evolving) thoughts on AI

- I study bio-inspired robotics and complex systems
- I think of AI as a complex adaptive system that we are integrating into our complex socio-technical system
- This is exciting and scary!
  - AI will help solve scientific problems
  - Current AI encodes biases, promotes surveillance & concentrates power
  - AI codes even better than it writes
  - How will humans co-evolve with AI?



[www.dailymaverick.co.za/article/2021-04-16-this-weekend-were-watching-coded-bias/](http://www.dailymaverick.co.za/article/2021-04-16-this-weekend-were-watching-coded-bias/)



<https://www.codedbias.com/>



# A bit about me & AI

- I study bio-inspired robotics and complex systems
- I think of AI as a complex adaptive system that we are integrating into our complex socio-technical system
- This is exciting and scary!
  - AI will help solve scientific problems
  - Current AI encodes biases, promotes surveillance & concentrates power & threatens jobs, particularly of creatives
  - AI codes even better than it writes
  - How will humans co-evolve with AI?

## Finance worker pays out \$25 million after video call with deepfake 'chief financial officer'



By Heather Chen and Kathleen Magramo, CNN

🕒 2 minute read · Published 2:31 AM EST, Sun February 4, 2024



Authorities are increasingly concerned at the damaging potential posed by artificial intelligence technology. boonchai wedmakawand/Moment RF/Getty Images

<https://www.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html>

## My (evolving) thoughts on AI

- I study bio-inspired robotics and complex systems
- I think of AI as a complex adaptive system that we are integrating into our complex socio-technical system
- This is exciting and scary!
  - AI will help solve scientific problems
  - Current AI encodes biases, promotes surveillance & concentrates power & threatens jobs, particularly of creatives
  - AI codes even better than it writes
  - How will humans co-evolve with AI?

Here's a Python code using 'matplotlib' to display the logistic map:

```
python
import numpy as np
import matplotlib.pyplot as plt

def logistic_map(r, x):
    return r * x * (1 - x)

def plot_logistic_map():
    r = np.linspace(2.5, 4.0, 10000) # Values of r
    x = 1e-5 * np.ones_like(r) # Initial condition for x

    iterations = 1000
    last = 100 # We'll only display the last few iterations to see the steady state

    for i in range(iterations):
        x = logistic_map(r, x)
        # Only display the bifurcation diagram for the last few iterations
        if i >= (iterations - last):
            plt.plot(r, x, 'k', alpha=0.25)

    plt.xlim(2.5, 4)
    plt.ylim(-0.1, 1.1)
    plt.xlabel('r')
    plt.ylabel('x')
    plt.title('Bifurcation diagram of the logistic map')
    plt.show()

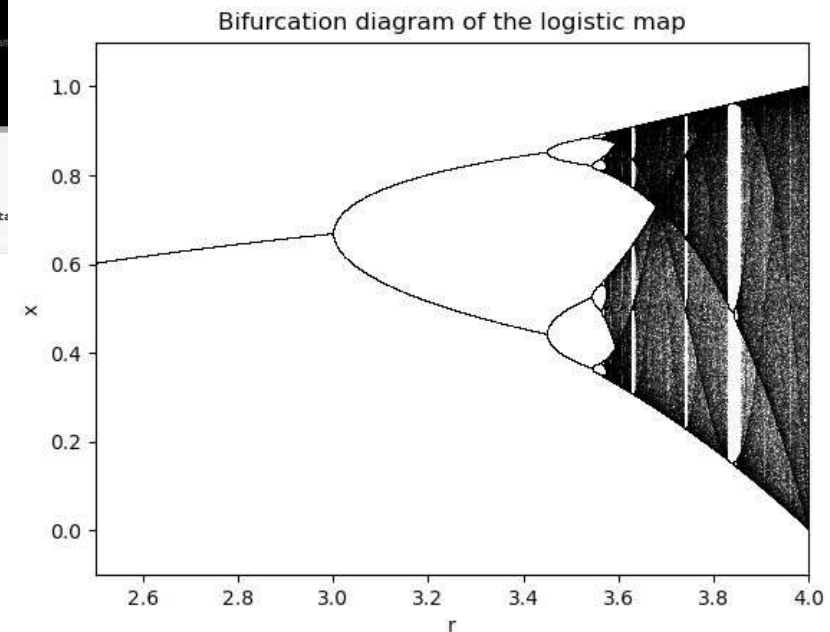
plot_logistic_map()
```

To see the bifurcation diagram:

1. Install the necessary libraries: 'pip install numpy matplotlib'
2. Run the above code.

From ChatGPT  
do my students'  
homework

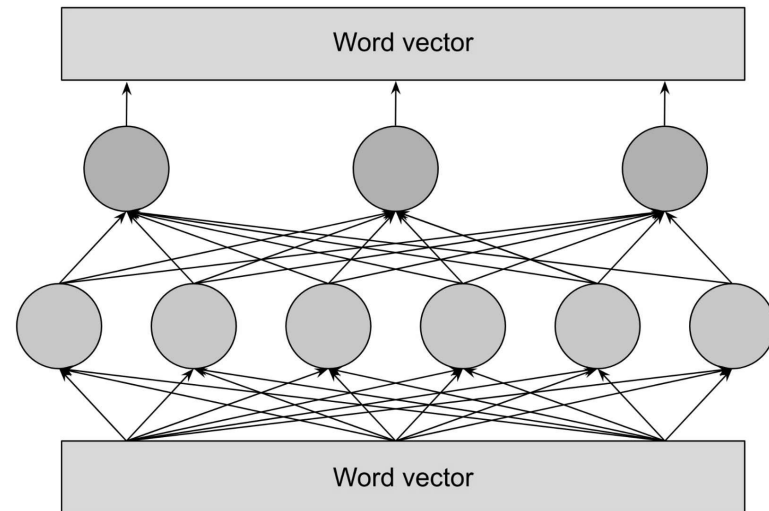
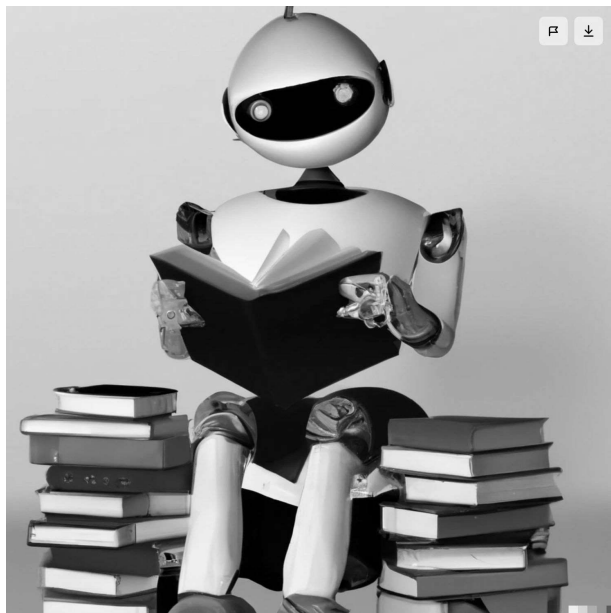
To Students: Use  
ChatGPT to do your  
homework



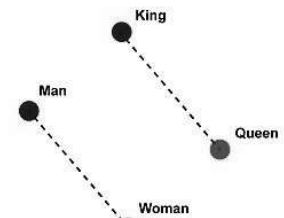
# LLMs learn by predicting what word comes next

LLM find patterns, predict next word

Trained on billions of nodes, trillions of words, trillions of parameters



<i>cat</i>	→	0.6	0.9	0.1	0.4	-0.7	-0.3	-0.2
<i>kitten</i>	→	0.5	0.8	-0.1	0.2	-0.6	-0.5	-0.1
<i>dog</i>	→	0.7	-0.1	0.4	0.3	-0.4	-0.1	-0.3



# This is tricky!

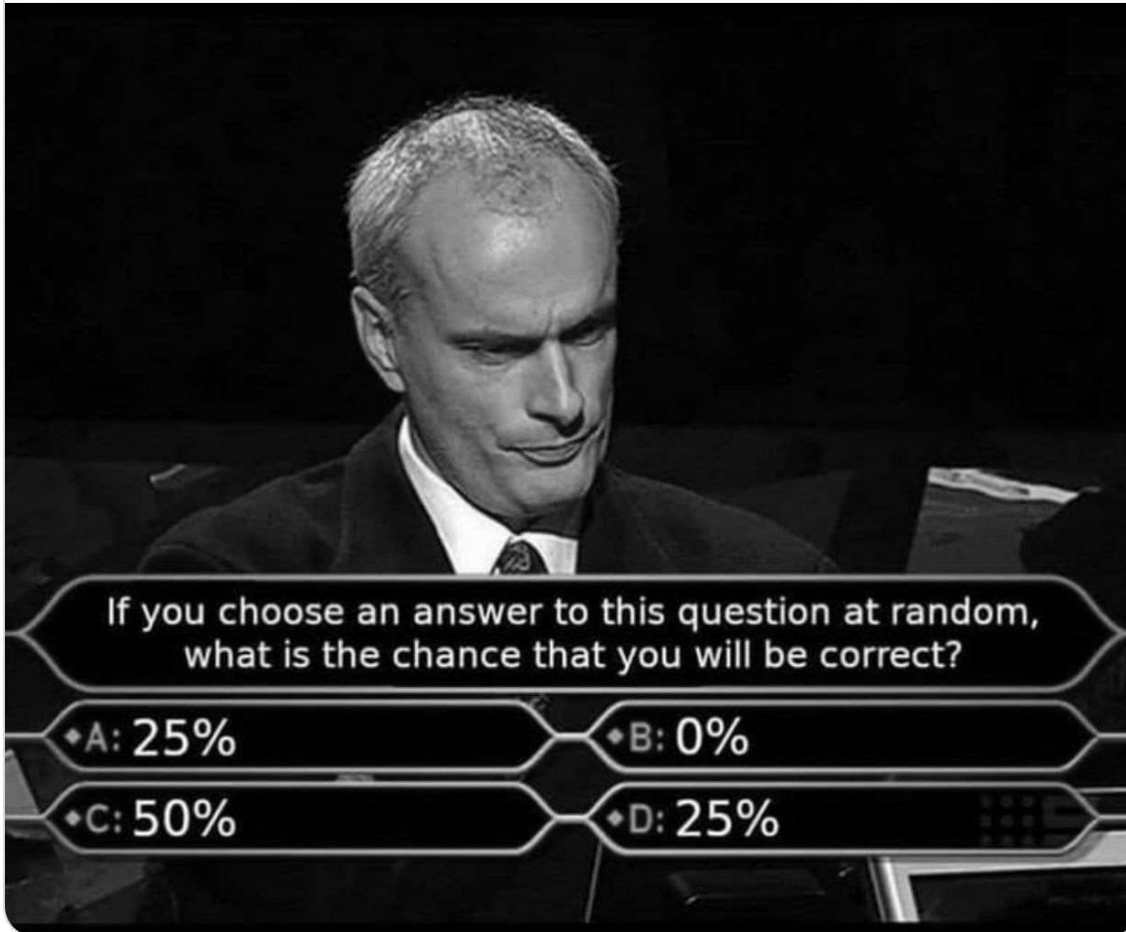
Transformers: “attention mechanisms” disambiguate multiple possible meanings:

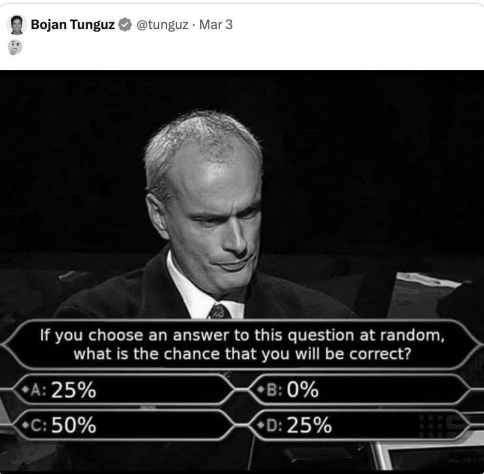
Bank: so many meanings! Sometimes LLMs guess wrong

But the guesses are usually plausible (particularly difficult to recognize that they are wrong)



 **Bojan Tunguz** ✓ @tunguz · Mar 3





## ChatGPT

This is a classic self-referential question that leads to a paradox if you try to answer it directly, as the percentages provided conflict with each other. Let's break down the options:

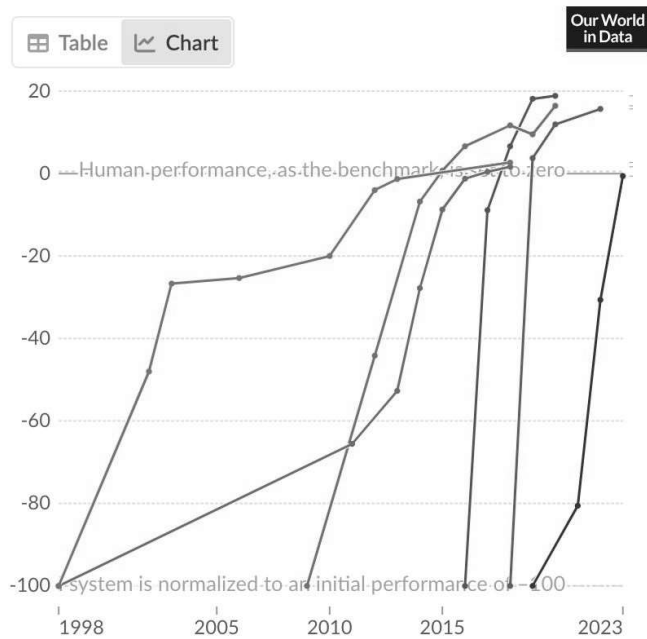
1. If we assume A) 25% is correct, then there are two options (A and D) that suggest a 25% chance of being correct. However, this would mean that there's actually a 50% chance of picking one of the two correct answers by chance, which contradicts the initial assumption.
2. If B) 0% is assumed correct, it implies that none of the answers provided would be correct, including B) itself, which creates a contradiction.
3. Assuming C) 50% is correct would mean there are two options out of four that are correct, which is not the case as we see no duplicate percentage values other than what A and D suggest (25%), but this itself leads to a contradiction as explained in point 1.
4. D) 25% faces the same issue as A) 25%, making the assumption of a 25% chance of being correct impossible due to the presence of two answers indicating 25%.

Thus, the question is designed to be paradoxical and doesn't have a straightforward answer within the confines of its own logic. It serves as a playful brain teaser rather than a question with a definitive answer.

# Scale is all you need?

## Test scores of AI systems on various capabilities relative to human performance

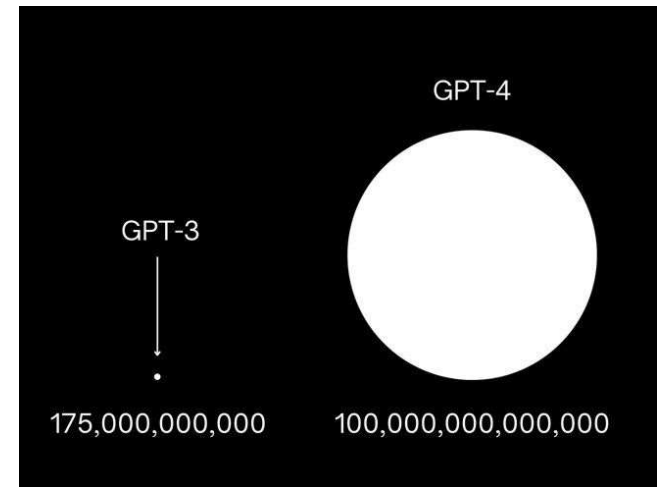
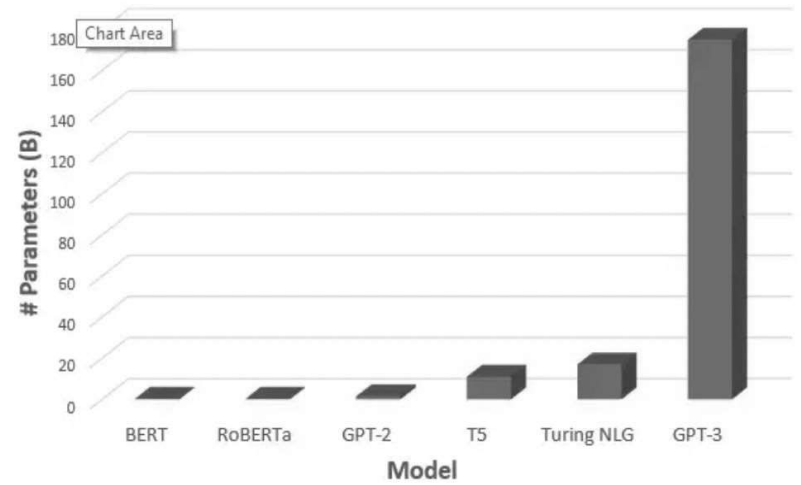
Within each domain, the initial performance of the AI is set to -100. Human baseline, set to zero. When the AI's performance crosses the zero line, it surpasses humans.



**300 billion trillion calculations to train GPT-3 on 500 billion words.**

It still manages to be wrong!

It's always making up answers



Training compute (FLOPs) of milestone Machine Learning systems over time

<https://arxiv.org/abs/2202.05924>

n = 102

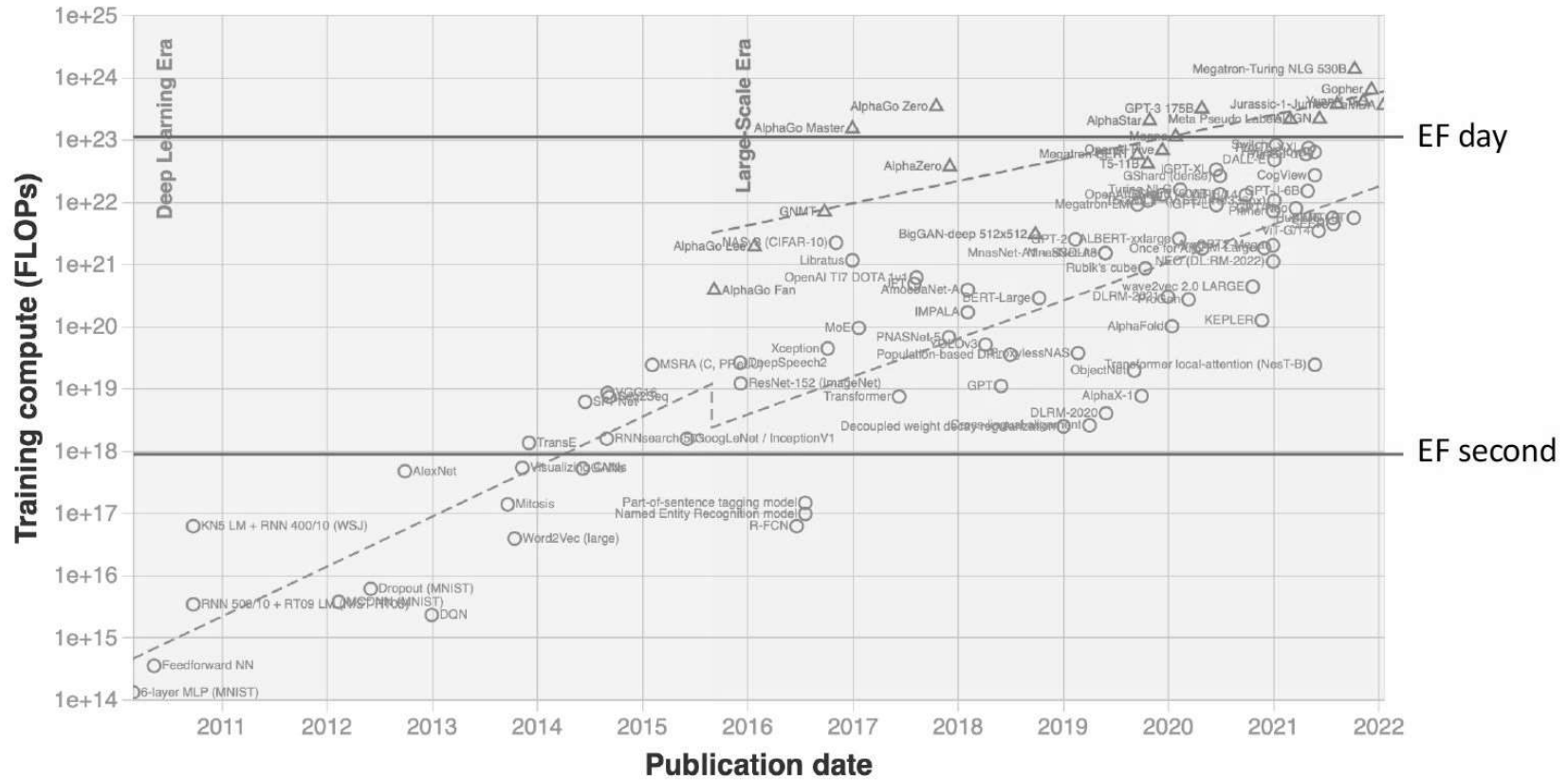
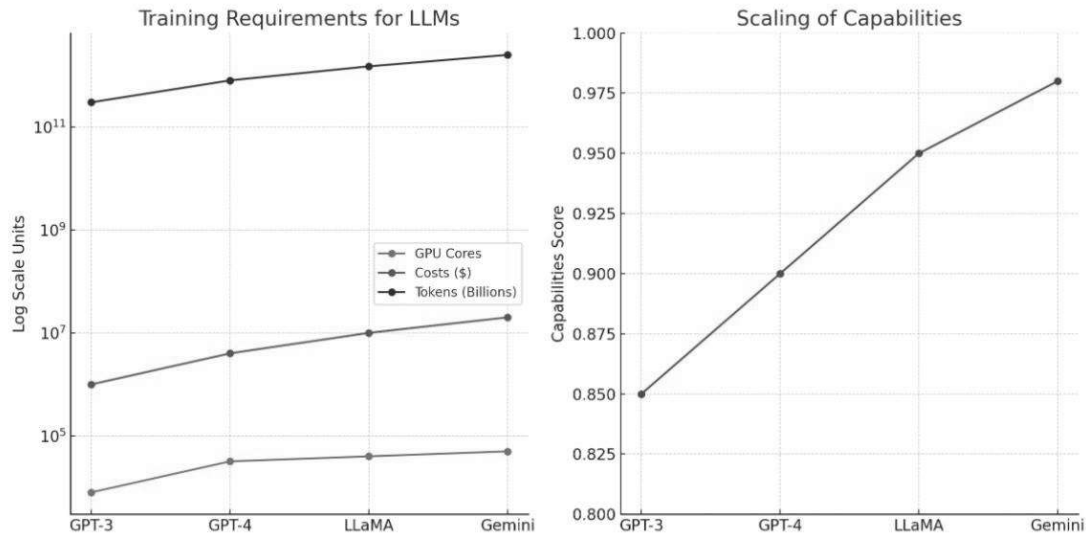


Figure 3: Trends in training compute of  $n=102$  milestone ML systems between 2010 and 2022. Notice the emergence of a possible new trend of large-scale models around 2016. The trend in the remaining models stays the same before and after 2016.

Slide Credit: Rick Stevens, Argonne Natl Lab & U Chicago



Let's create this visualization. Please hold on while I prepare the figure.



Here's the figure illustrating the scaling of training requirements and capabilities for popular large language models like GPT-3, GPT-4, LLaMA, and Gemini:

WRONG answer from GPT4

# Scale is what they have!



<https://medium.com/@johnadjanohoun/microsoft-and-openai-to-forge-100-billion-supercomputer-for-ais-next-leap-2d5bbcf189f6>

ombrulla.com

Saudi Arabia plans \$40 billion investment in AI

Saudi Arabia Goes Big on AI: A \$40 Billion Investment [Visit >](#)

<https://www.linkedin.com/pulse/saudi-arabia-goes-big-ai-40-billion-investment-ombrulla-d3eic>

Microsoft, Meta, and Google's parent company, Alphabet, disclosed this week that they had spent more than \$32 billion combined on data centers and other capital expenses in just the first three months of the year.

The people who made the biggest fortunes in California's gold rush weren't the miners — they were the people selling the shovels. (NVIDIA)

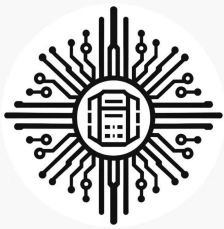
<https://www.nytimes.com/2024/04/27/technology/ai-big-tech-spending.html?smid=em-share>

# How did we get here?

- Scale
- A few clever tricks in neural networks
  - Encoders, Transformers
- Human feedback by low paid workers (Kenyan workers paid \$2/day)
- Marketing & economic incentives have concentrated power in a few huge companies

# Where are we going?

- No one knows!
- No exponential lasts forever; all scaling hits constraints
- Feedback comes from LLM users
- Will smaller models become viable - can we democratize AI & use it for good?
- How can we regulate AI & prevent harm?



# AI@UNM: Empowering Innovation in New Mexico

>30 UNM Faculty Actively Conducting Research in AI



## Accelerate Discovery with AI

- Supercharge sustainable solutions: smart water allocation; solar energy generation, transmission & storage; forest fire mitigation and recovery; environmental monitoring & remediation
- Secure infrastructure for trustworthy AI in medicine; law & policy; education; & national security
- Advance biomedical innovations, drug discovery & personalized health care for diverse and rural populations
- Develop novel materials, computational fabrication & advanced manufacturing
- Synthesize data to detect methane leaks, volcanic & seismic & natural hazards; improve climate models
- Develop new open source foundational architectures for accessible, applicable and approachable AI
- Develop energy-aware AI based on collective intelligence and quantum technologies

## Foster Collaboration & Catalyze Economic Growth

- Crystalize NM as an HPC technology hub
- Prepare NM students to thrive in the AI workforce
- Attract world class faculty
- Spur a startup ecosystem & tech transfer
- Propel NM to the forefront of 3D immersive technologies for education, workforce training, manufacturing, planning & the creative industry
- Engage diverse communities to shape AI for healthcare, law & issues affecting daily lives
- Leverage New Mexico strengths:
  - Expertise in cybersecurity and quantum computing to design next generation AI
  - Power of National Labs, abundance of landscape & natural resources & diversity and creativity of New Mexicans
  - Leadership in Hispanic Serving Institution & EPSCoR networks



# Interdisciplinary Working Group on Algorithmic Justice



Melanie Moses  
UNM / SFI



Cris Moore  
SFI



Kathy Powers  
Poli Sci, UNM/ SFI



Alfred Mathewson  
Law, UNM



Sonia Rankin Law,  
UNM



Mirta Galesic  
SFI



Josh Garland  
Arizona State



Matthew Fricke  
CS, UNM



Gabe Sanchez  
Poli Sci, UNM



Tina Eliassi-Rad  
CS, NEU /SFI



Mahzarin Banaji  
Psych, Harvard/SFI



Trilce Estrada  
CS, UNM



Nadiyah Humber  
Law, UConn

# Trustworthy AI Research & Education for New Mexico



Xin Chen  
Computer Science



Trilce Estrada  
Computer Science



Sonia Gipson-Rankin  
Law



Abdullah Mueen  
Computer Science



Meeko Oishi  
Computer Engineering



Cris Moore, SFI  
Computer Science



Melanie Moses  
Computer Science



Manel Martinez Ramon  
Computer Engineering



Lydia Tapia, Chair  
Computer Science

# TrustworthyAI @ UNM

## For Trusted Autonomy



- Develop tools to **detect unsafe behavior** in neural networks
- **Build safe data sets** to train neural networks to produce results within pre-specified safe bounds
- **Develop mitigation strategies** to eliminate dangerous or unexpected predictions & behaviors
- **Iteratively improve neural networks** by removing causes of unsafe behavior
- Build upon our toolbox for neural net verification



Mathematical Foundations



# TrustworthyAI In Law & Medicine @ UNM



- **Analyze accuracy and fairness in risk assessment algorithms used in criminal justice**
- Build on close collaborations with UNM's Institute for Social Research, the Administrative Office of the Courts, the New Mexico Corrections Department, and City and County governments.
- **Identify sources of implicit bias, anomalies and uncertainties in medical records**
- **Provide trustworthy interpretations of medical reports** in close collaboration with stakeholders ensuring that AI is transparent, fair, and inclusive.



Apply trustworthy AI to  
protect rights & improve health

# AI @UNM

---

## **Presenters**

### **Kent Kiehl**

Professor

Department of Psychology

Presentation Title: "Machine learning of brain science in legal contexts"

### **Sarah Dreier**

Assistant Professor

Department of Political Science

Presentation Title: "AI and text quality: Identifying sources of training biases and imperfect classification outcomes"

### **Jessica Feezell**

Associate Professor

Department of Political Science

Presentation Title: "Algorithm-driven News and Political Polarization in the U.S."

### **Cristyn Elder**

Associate Professor

Department of English Language & Literature

### **Iain Thompson**

Professor

Department of Philosophy

Presentation Title: "What Is Called Thinking in the Age of AI?"

### **Brandon Schmandt**

Professor

Department of Earth & Planetary Sciences

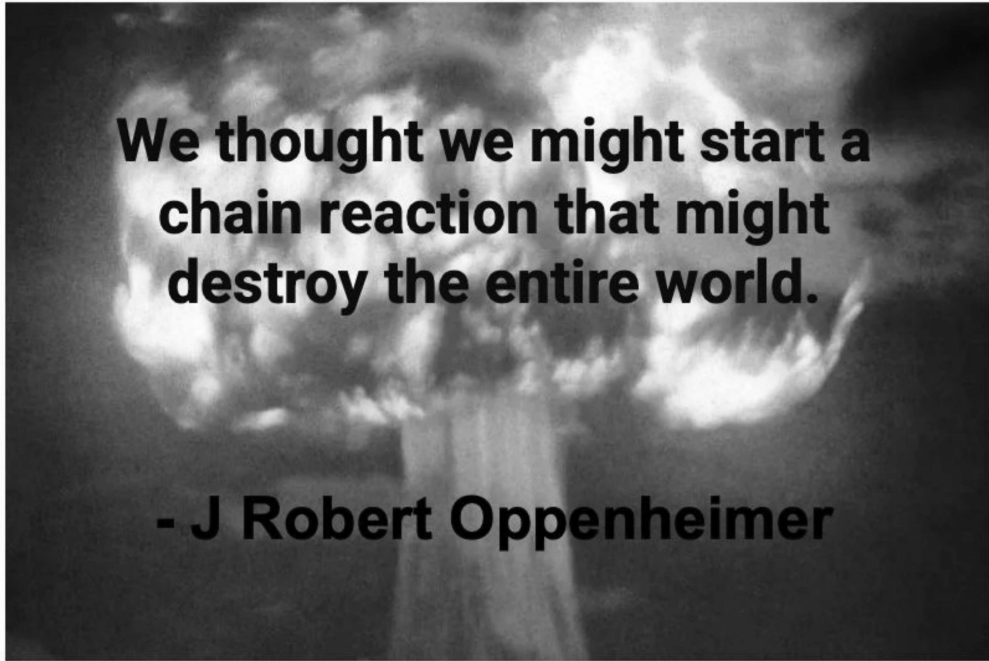
Presentation Title: "Classification of seismic signals with machine learning: applications to environmental change, earthquakes in New Mexico, and national security"

### **Joseph Galewsky**

Professor

Department of Earth & Planetary Sciences

Presentation Title: "Cloudy with a Chance of Bits: Using AI to improve our understanding of water vapor's role in the climate system"



**We thought we might start a  
chain reaction that might  
destroy the entire world.**

**- J Robert Oppenheimer**

Atomic bomb Trinity test.

UNIVERSAL HISTORY ARCHIVE/UIG VIA GETTY IMAGES



**Powerful new technologies  
reshape the world...**

**but not in the ways we initially fear  
or expect.**

# Even deep fakes have positive uses

## An AI Salvador Dalí will answer any question when called on his famous 'lobster phone'

UPDATED APRIL 21, 2024 · 1:08 PM ET

 Chloe Veltman



Ask Dalí at the Dalí Museum in St. Petersburg, Fla., allows visitors talk to the famous surrealist artist via an AI-generated version of his voice.

Martin Pagh Ludvigsen/Goody: Silverstein & Partners

<https://www.npr.org/2024/04/21/1246117954/salvador-dali-artificial-intelligence-lobster-telephone-museum>





John Powell  
UC Berkeley Law Professor

Kennedy told us that we  
ALL had to contribute for  
America to reach the moon

Now we all need to shape our future  
here on earth



# Bottom Line Up Front

## Regulatory priorities

- Deep fakes, beyond elections
- Transparency & equity in high stakes decision-making
- AI/algorithms in healthcare
- Protect creative works
- Establish study group to examine impact and equity of AI/algorithms used by NM state government...and beyond

## Investment priorities

- New Mexicans need
  - AI literacy
  - To be engaged in shaping the future of AI
- Trustworthy AI for science and society
- Partnerships among NM Colleges, Universities & National Labs; catalyze innovation, education & economic growth

Major investments from industry & governments; uncertainty about what's next

# Deep Fake Legislation

NM passed the first of many deep fakes in elections laws in 2024 with HB182 that requires disclaimers for deep fakes in campaign ads

(several passed last year)

Michigan, Washington & California require disclosures like NM

Texas and Minnesota ban deep fakes

**Need to expand beyond elections: scams, pornography, slander, fraud**

**Outlaw “counterfeit humans” because counterfeits destroy trust**

# Transparency, audits & fairness in AI

POLITICO

<https://politico.com/news/magazine/2024/04/30/ai-legislation-states-mistake-00155006>

MAGAZINE

HIGH TECH

## Opinion | Are These States About to Make a Big Mistake on AI?

AI is potentially transformative. Whether that's a good or bad thing depends on whether we set the right rules.



The Connecticut Senate pressed ahead on April 24, with one of the first major legislative proposals in the U.S. to rein in bias in AI decision-making and protect people from harm, including manufactured videos or deepfakes. | Richard Drew/AP

Devil is in the details

Define consequential decision-making

Require transparency, explainability

Eliminate loopholes for trade secrets

Foundations in civil rights

Strong Enforcement



[nature](#) > [nature reviews electrical engineering](#) > [comment](#) > [article](#)

Comment | Published: 12 April 2024

# Medical artificial intelligence should do no harm

[Melanie E. Moses](#)  & [Sonia M. Gipson Rankin](#)

*Nature Reviews Electrical Engineering* (2024) | [Cite this article](#)

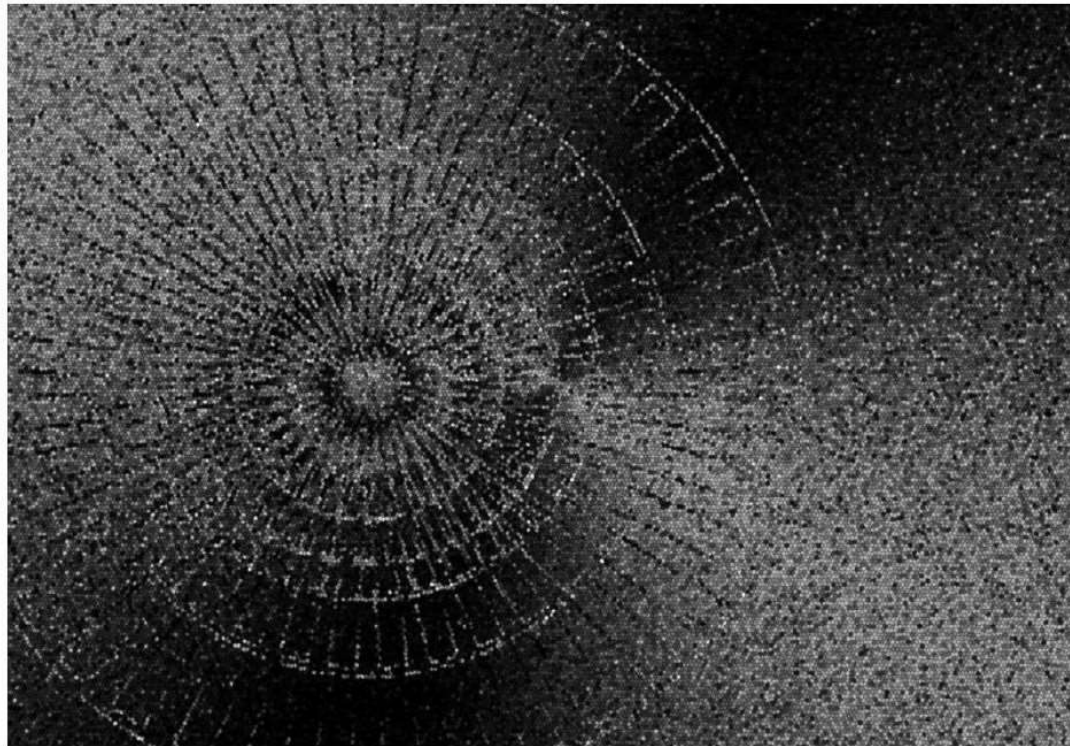
**40** Accesses | **4** Altmetric | [Metrics](#)

**Bias and distrust in medicine have been perpetuated by the misuse of medical equations, algorithms and devices. Artificial intelligence (AI) can exacerbate these problems. However, AI also has potential to detect, mitigate and remedy the harmful effects of bias to build trust and improve healthcare for everyone.**

# How Complexity Thinking Can Help the World Navigate AI

A Paper from the Study Group on Complexity and International Relations

POLICY PAPER



Olena Pylypuha / Shutterstock.com

<https://www.newamerica.org/planetary-politics/policy-papers/how-complexity-thinking-can-help-the-world-navigate-ai/>

# Resources

How AI works

<https://arstechnica.com/science/2023/07/a-jargon-free-explanation-of-how-ai-large-language-models-work/>

Politico: legislation for transparency and audits

<https://www.politico.com/news/magazine/2024/04/30/ai-legislation-states-mistake-00155006>

Complexity & AI

<https://www.newamerica.org/planetary-politics/policy-papers/how-complexity-thinking-can-help-the-world-navigate-ai/>

Moses & Gipson Rankin article on Trustworthy AI in Medicine

<https://www.nature.com/articles/s44287-024-00049-2>

AI and Algorithmic Justice Resources @UNM & SFI

<https://airesources.unm.edu/students.html>

<https://algorithmicjustice.cs.unm.edu/index.html>

<https://www.santafe.edu/research/projects/algorithmic-justice>

[melaniem@unm.edu](mailto:melaniem@unm.edu) <https://moseslab.cs.unm.edu>

## Questions?