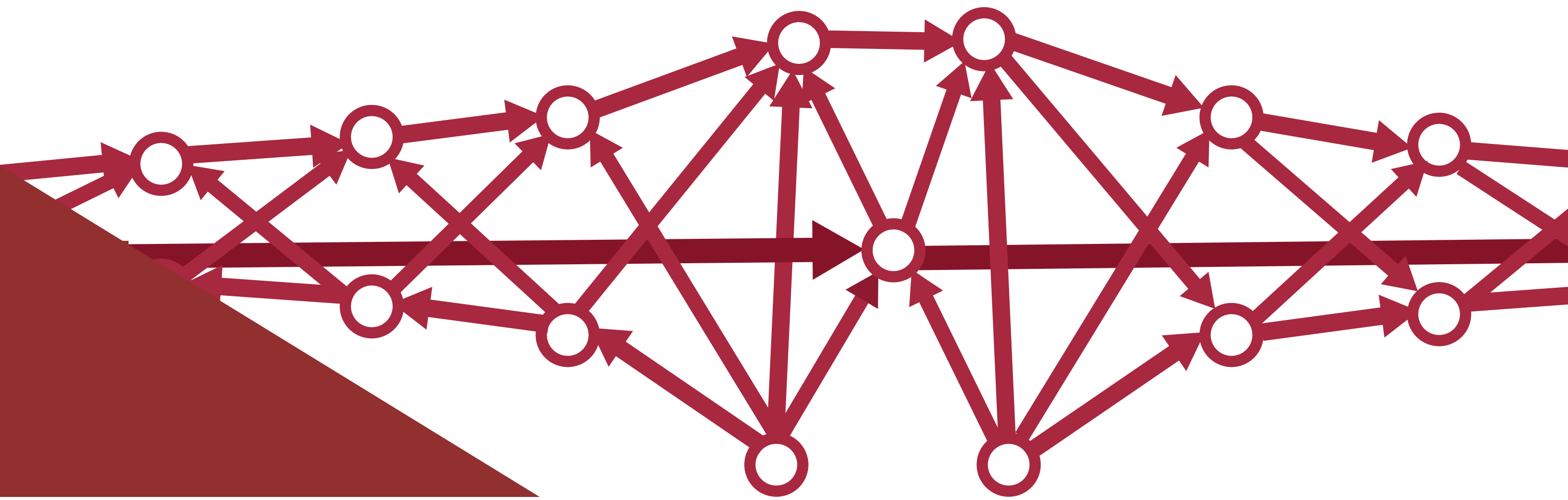


# Machine Learning & Data Analysis



## Week 6: Machine Learning in society

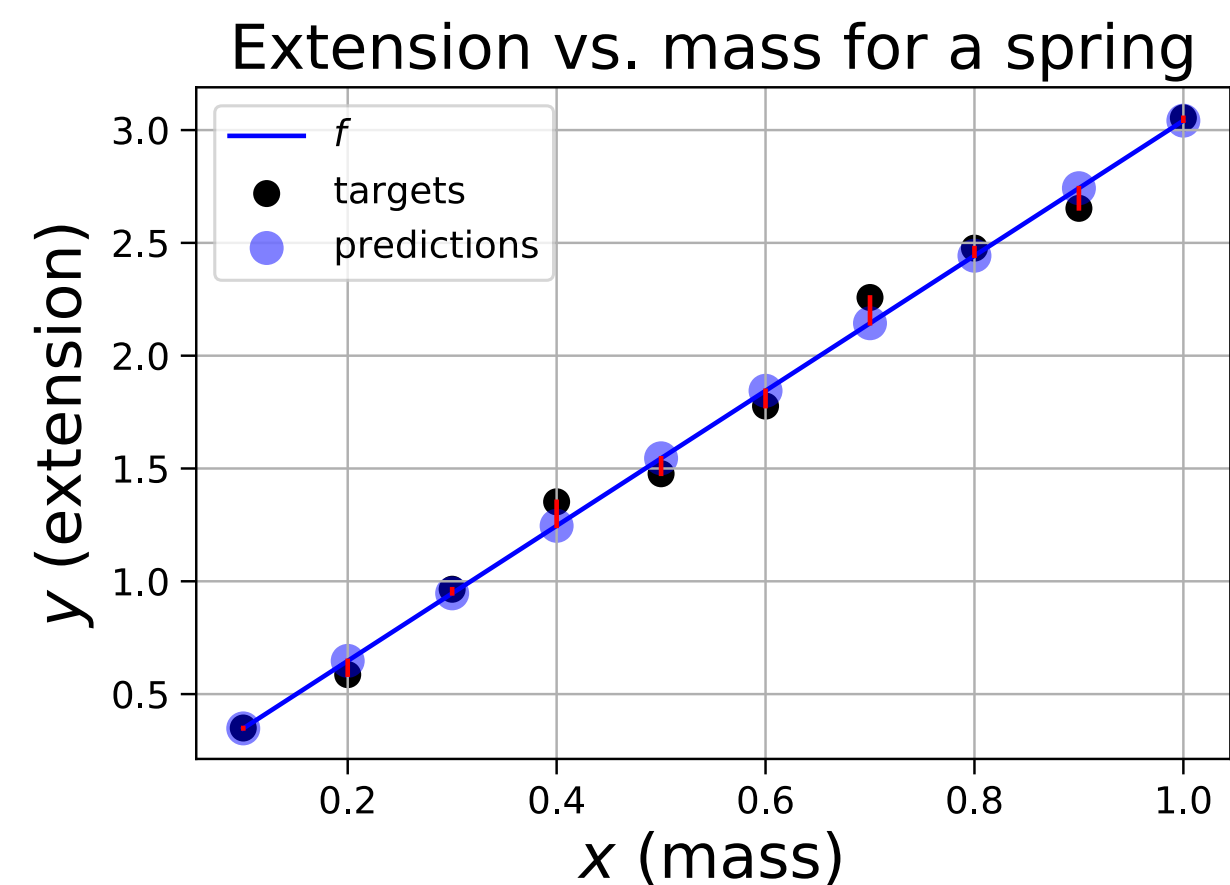
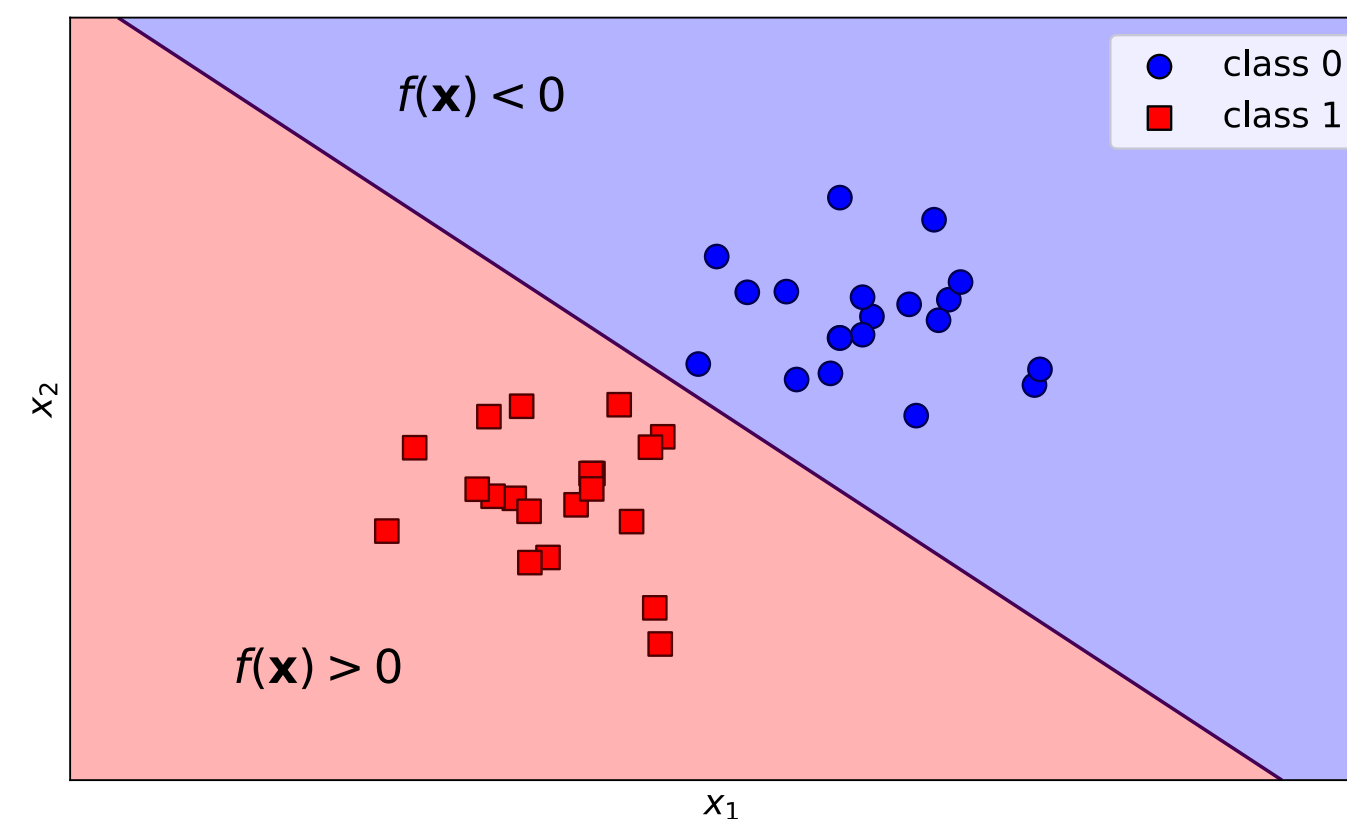
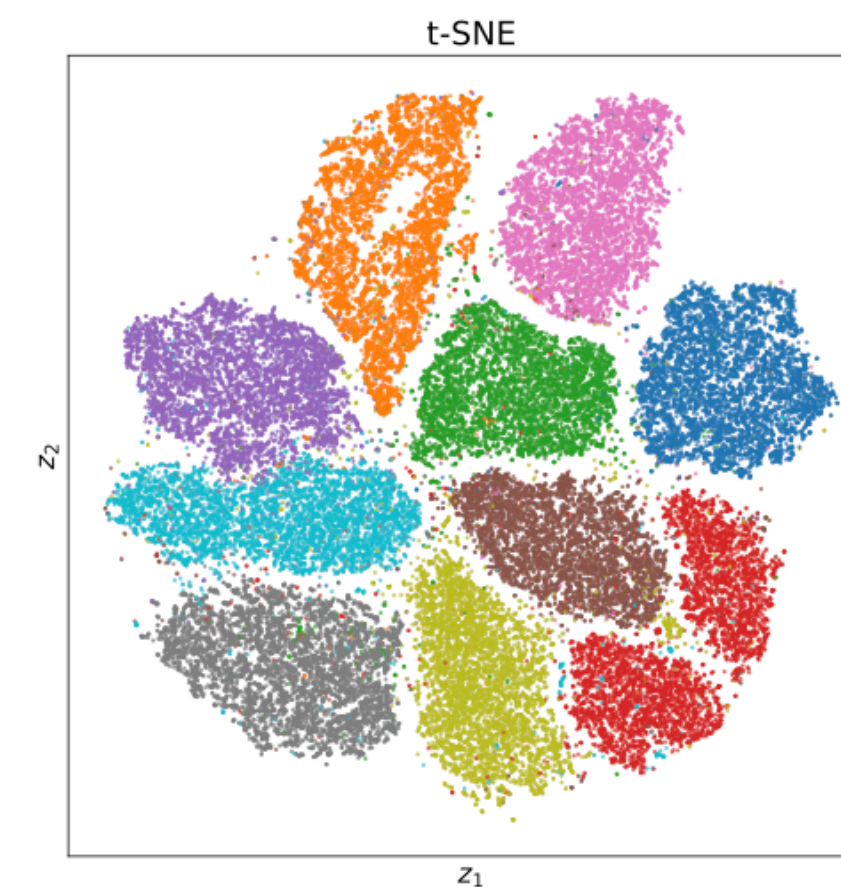
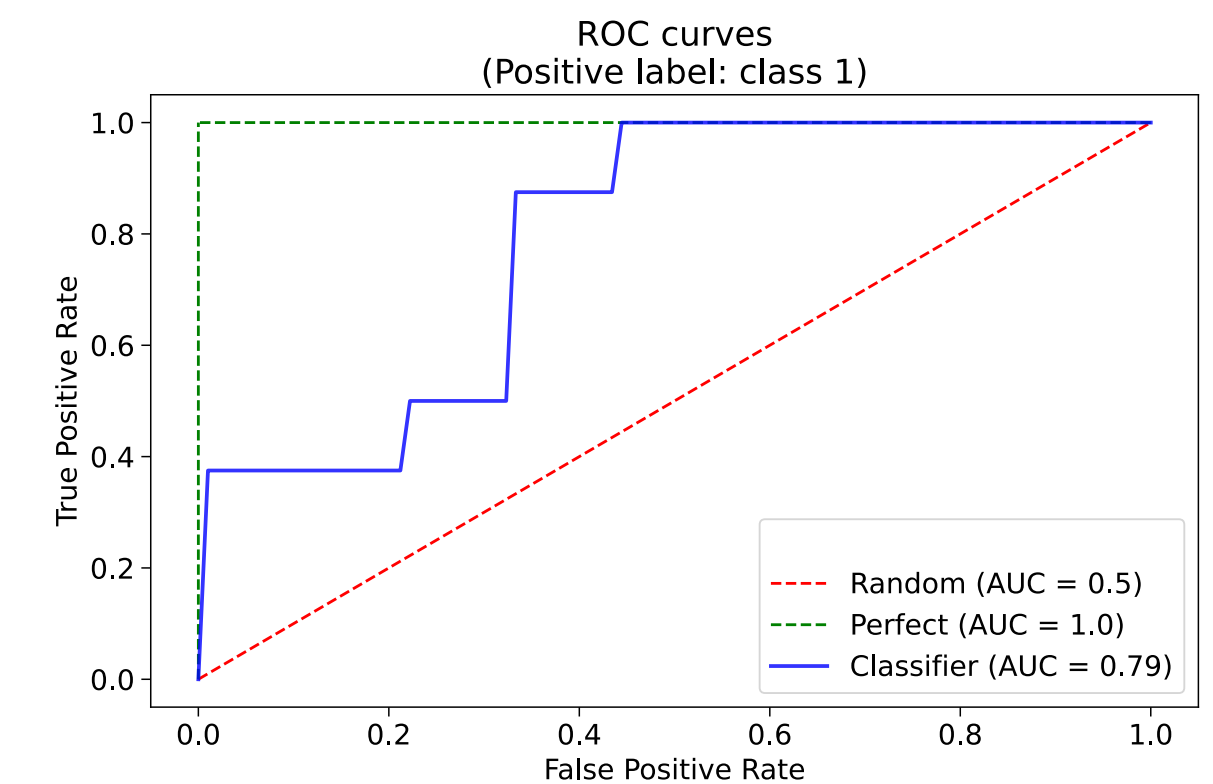
Elliot J. Crowley, 26/10/26



THE UNIVERSITY  
*of* EDINBURGH

# The half-way point. What have we done?

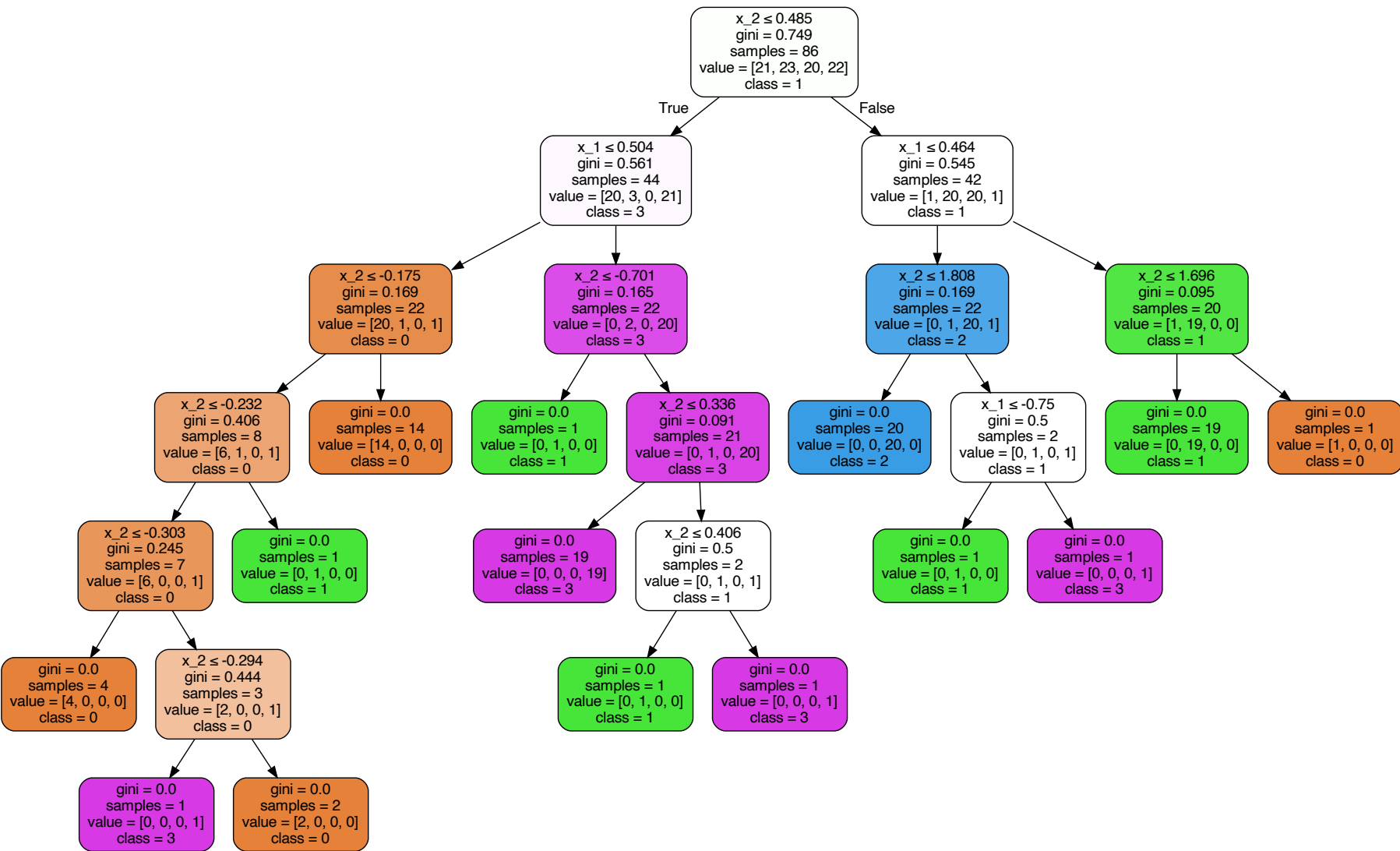
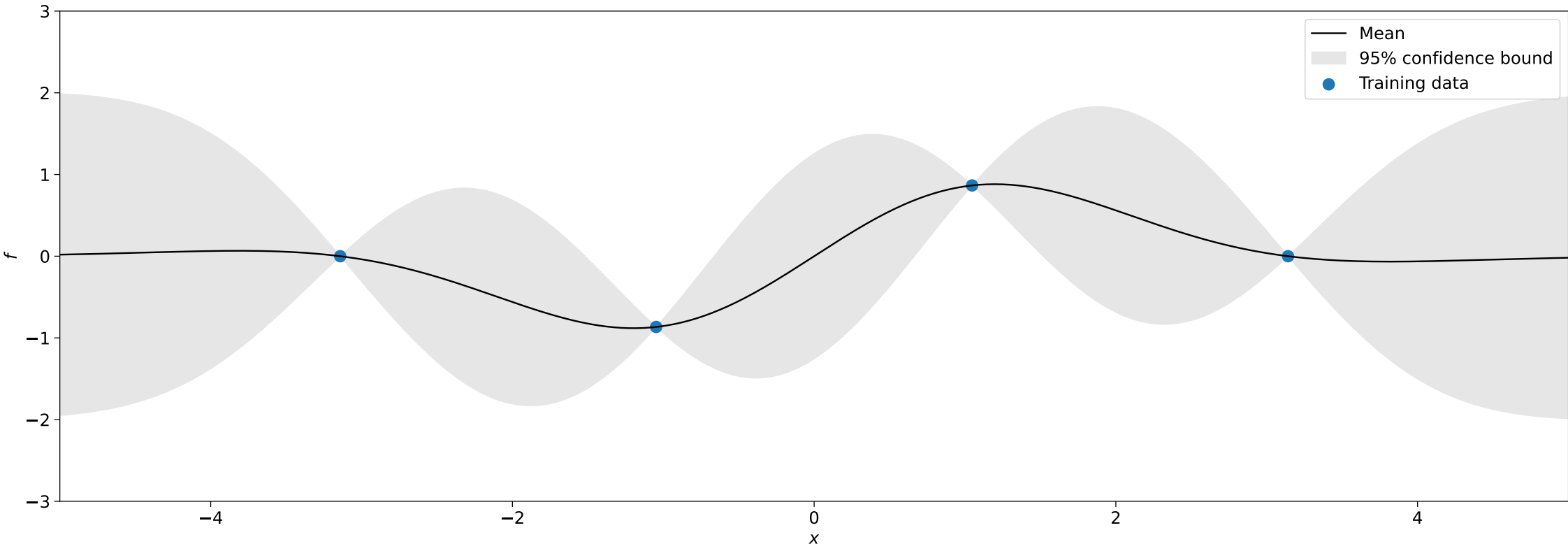
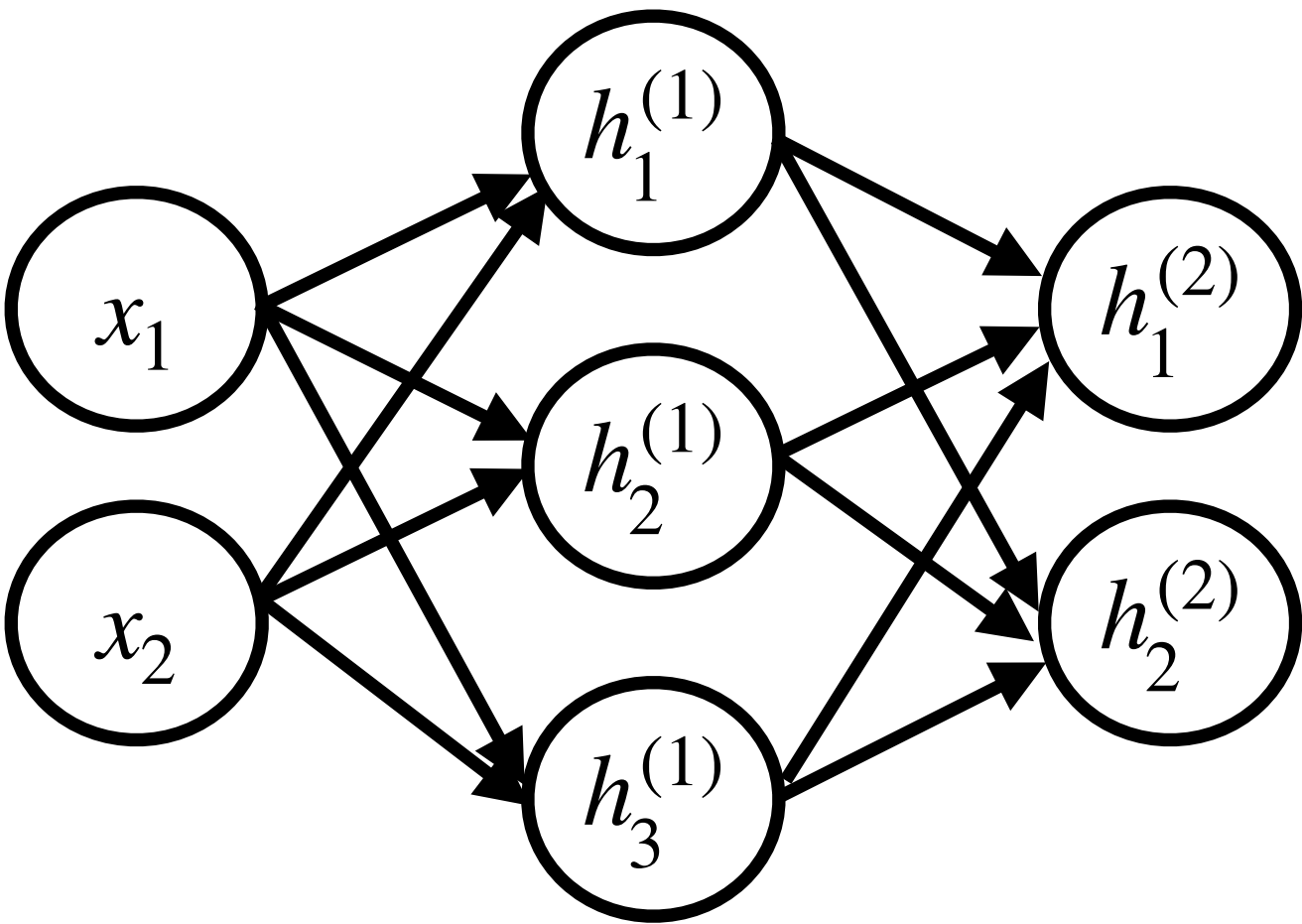
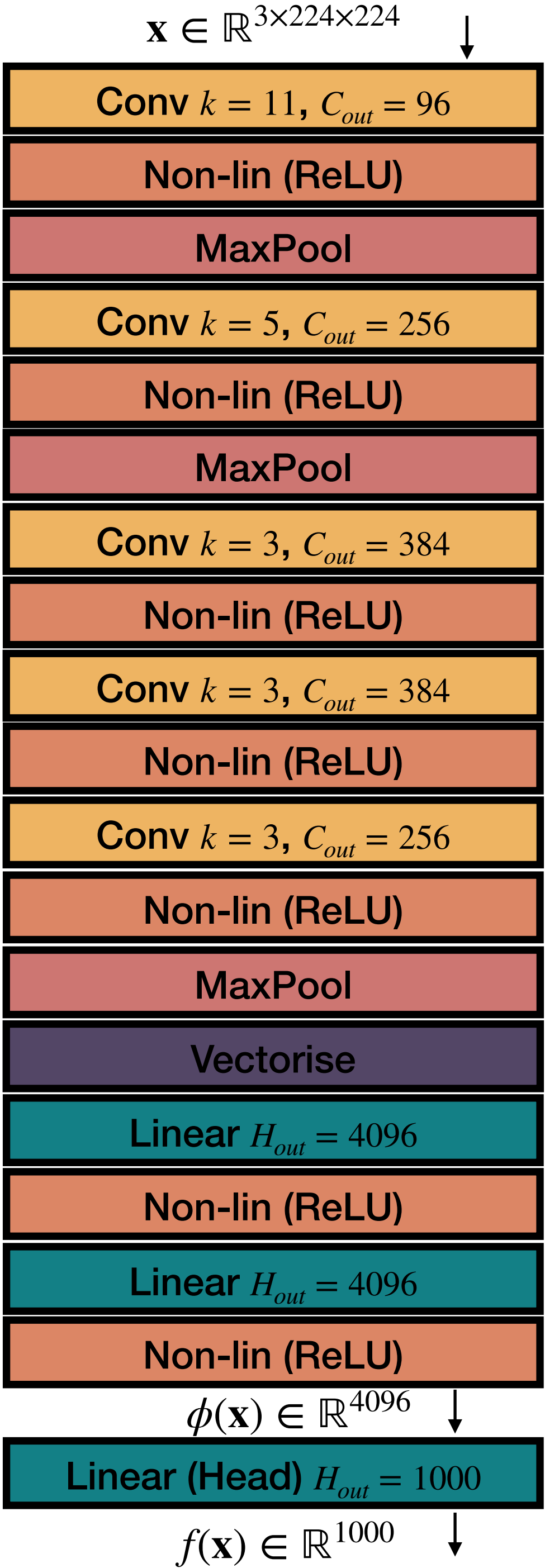
- Linear models for regression
- Linear models for classification
- Data visualisation and dimensionality reduction
- Model selection and evaluation





# Where are we going?

- Decision trees
- Gaussian processes
- Neural networks



# But first, an intermission lecture

- One of the learning objectives on this course is “...**describe ethical issues that can arise when deploying machine learning algorithms in society**”
- While I am no ethicist, I do develop ML algorithms, and people use these
- I do not want to espouse any particular views or opinions....
- I just want you to reflect, as you too may develop or deploy ML in the future



**Let's address “AI”**



# People tend to call machine learning models “AI”





# Be wary of hype (slide created when this course started a few years ago)

## Robots could soon think like HUMANS: Scientists develop AI that can learn the basic common sense rules of the physical world – just like a baby

- AI can be taught 'intuitive physics' - common sense rules of how the world works
- Researchers at DeepMind trained an AI called PLATO with slides of a ball moving
- PLATO demonstrated learning and 'surprise' if a ball moved in an impossible way

By JONATHAN CHADWICK FOR MAILONLINE   
PUBLISHED: 16:11, 11 July 2022 | UPDATED: 17:05, 11 July 2022

Community

Debate over AI sentience marks a watershed moment



**Elon Musk**  @elonmusk · Sep 4, 2017  
It begins ...

**The Verge**  @verge  
Putin says the nation that leads in AI 'will be the ruler of the world' theverge.com/2017/9/4/16251..  


**Elon Musk**  @elonmusk · [Follow](#)

China, Russia, soon all countries w strong computer science. Competition for AI superiority at national level most likely cause of WW3 imo.  
10:33 AM · Sep 4, 2017 

 39.9K  Reply  Share

[Read 3.5K replies](#)

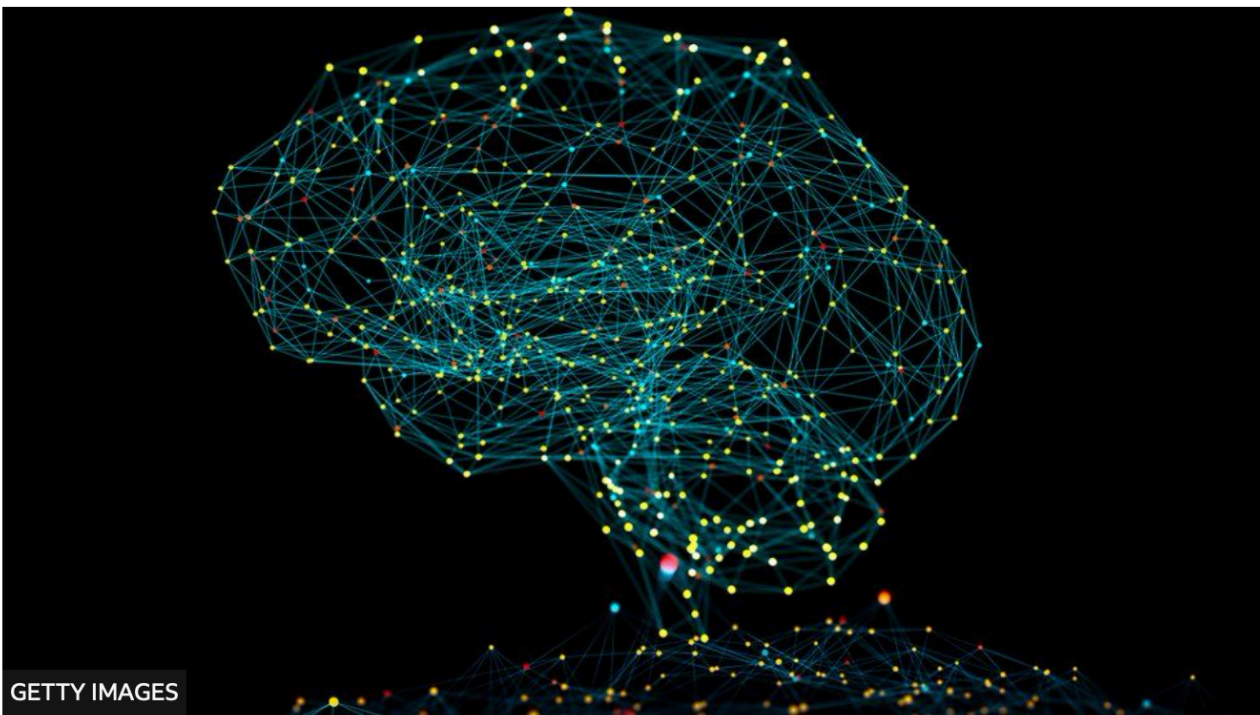
## Artificial Intelligence Market Hits USD 35,870 Million By 2025: By Grand View Research, Inc.

The global [artificial intelligence market](#) is expected to reach USD 35,870.0 million by 2025 from its direct revenue sources, growing at a CAGR of 57.2% from 2017 to 2025, whereas it is expected to garner around USD 58,975.4 million by 2025 from its enabled revenue arenas, according to a new report by Grand View Research, Inc.

Artificial Intelligence (AI) is considered to be the next stupendous technological development, alike past developments such as the revolution of industries, the computer era, and the emergence of smartphone technology. The North American region is

## Google engineer says Lamda AI system may have its own feelings

By Chris Vallance  
Technology reporter  
13 June



A Google engineer says one of the firm's artificial intelligence (AI) systems might have its own feelings and says its "wants" should be respected.



# Be wary of hype?? (slide created a few weeks ago)

GOV.UK

[Home](#) > [Business and industry](#) > [Science and innovation](#) > [Artificial intelligence](#)

Collection

AI Growth Zones

AI Growth Zones will unlock investment in AI-enabled data centres and support infrastructure by improving access to power and providing planning support.

From: [Department for Science, Innovation and Technology](#)

Published 13 November 2025

🔔 Get emails about this page

Contents

— [AI Growth Zones application process](#)

— [AI Growth Zones announcements \(press releases\)](#)

— [AI Growth Zones background](#)

The UK government is accelerating efforts to build the infrastructure needed to support AI growth. AI Growth Zones will help drive innovation, create high-skilled jobs, and strengthen the UK's position as a leader in AI.

INDEPENDENT

PodcastsNewslettersENHANCED BY

NEWSPREMIUMSPORTVOICESCULTURELIFESTYLEINDYBESTTRAVELMORE

Tech

Mark Zuckerberg to invest hundreds of billions into AI that can outperform humans

Jaspreet Singh & Aditya Soni  
Tuesday 15 July 2025 10:22 BST

🔗🔖💬 3 Comments

INDEPENDENT tv

Mark Zuckerberg announces Meta AI upgrade

the new version of Meta AI,

Mark Zuckerberg announces Meta AI upgrade

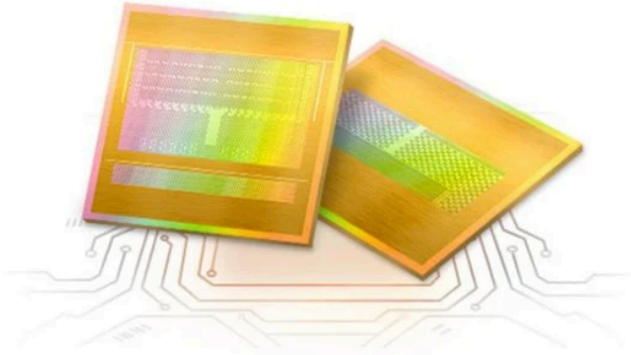
Business Standard

HOME TECHNOLOGYTECH NEWS TECH REVIEWS GADGETS E-PAPERDECODEDOPINION

[Home](#) / [Technology](#) / [Tech News](#) / AI-led memory shortage threatens to make phones, TVs, PCs costlier in 2026

AI-led memory shortage threatens to make phones, TVs, PCs costlier in 2026

A global supply squeeze driven by AI servers is raising costs for phones, PCs and TVs, while some brands consider Chinese memory makers as a fallback



High-bandwidth memory (HBM) chips from SK Hynix

Harsh Shrivam | New Delhi  
7 min read Last Updated : Feb 09 2026 | 3:36 PM IST

Interview

Demis Hassabis on our AI future: ‘It’ll be 10 times bigger than the Industrial Revolution – and maybe 10 times faster’

Steve Rose



📷 'I think about the unknowns and how we can make them less unknown' ... Demis Hassabis at Google's London offices. Photograph: Antonio Olmos/The Guardian

Search

FORTUNE

Subscribe for \$1Sign in

HomeLatestFortune 500FinanceTechLeadershipLifestyleRankingsMultimedia

AI • SAM ALTMAN

Sam Altman thinks AI will surpass human intelligence by 2030. His rival AI billionaires say it’ll be even sooner

By Marco Quiroz-Gutierrez  
Reporter

September 26, 2025, 2:00 PM ET

Add us on

🔗



Sam Altman, CEO of OpenAI

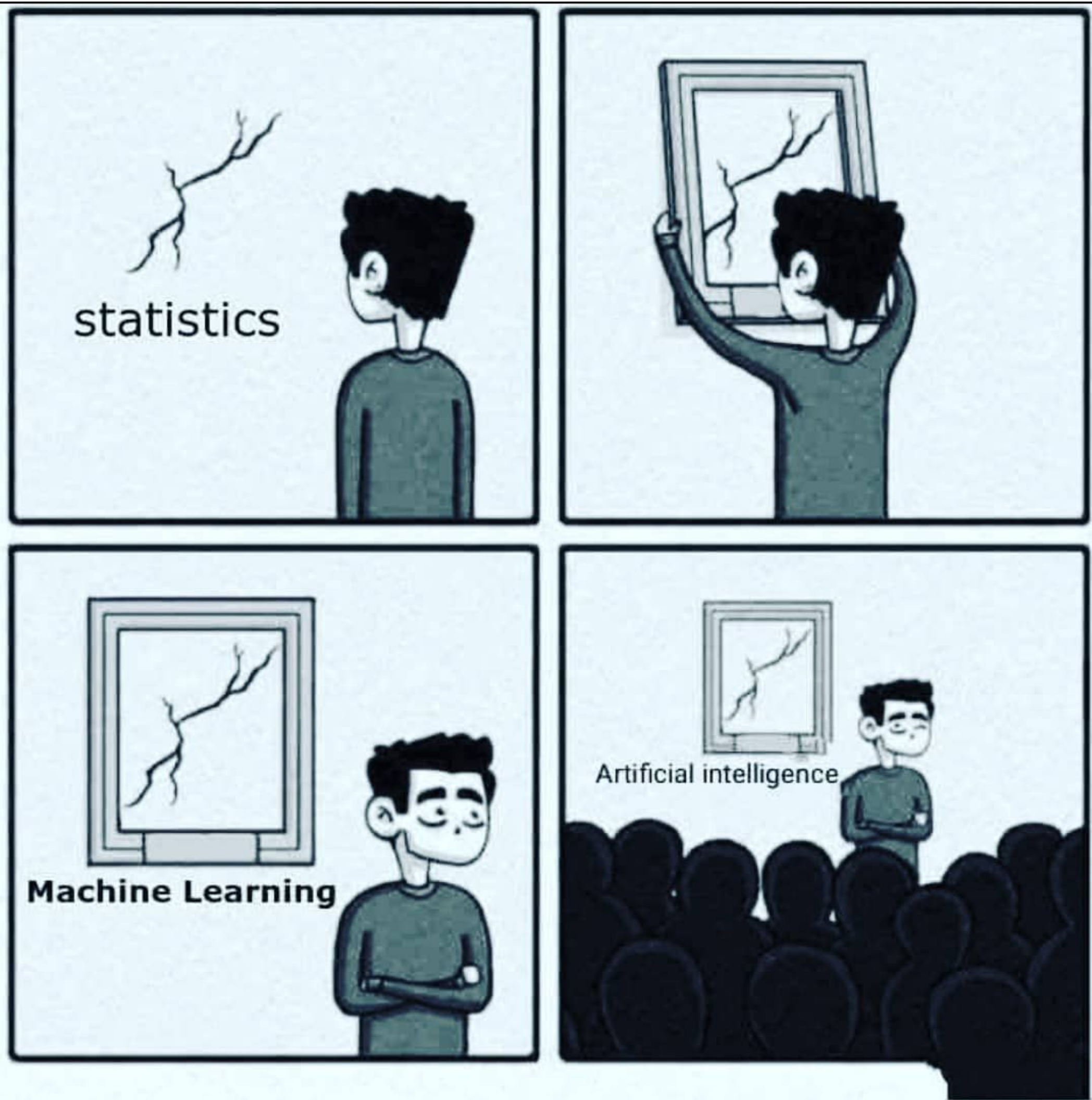
KYLE COHEN / GETTY IMAGES



# Reality



<https://xkcd.com/1838/>



<https://www.pinterest.co.uk/pin/591449363544632526/>



# Reality

Analysis | [Open Access](#) | [Published: 15 March 2021](#)

## Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans

[Michael Roberts](#) , [Derek Driggs](#), [Matthew Thorpe](#), [Julian Gilbey](#), [Michael Yeung](#), [Stephan Ursprung](#), [Angelica I. Aviles-Rivero](#), [Christian Etmann](#), [Cathal McCague](#), [Lucian Beer](#), [Jonathan R. Weir-McCall](#), [Zhongzhao Teng](#), [Effrossyni Gkrania-Klotsas](#), [AIX-COVNET](#), [James H. F. Rudd](#), [Evis Sala](#) & [Carola-Bibiane Schönlieb](#)

[Nature Machine Intelligence](#) **3**, 199–217 (2021) | [Cite this article](#)

in EMBASE and MEDLINE in this timeframe are considered. Our search identified 2,212 studies, of which 415 were included after initial screening and, after quality screening, 62 studies were included in this systematic review. Our review finds that none of the models identified are of potential clinical use due to methodological flaws and/or underlying biases.

<https://www.nature.com/articles/s42256-021-00307-0>

EDITORS' PICK | INNOVATION > AI

Companies Are Pouring Billions Into AI. Here's Why They're Not Seeing Returns

By [Kolawole Samuel Adebayo](#), Contributor. © I write about the economics of AI. 

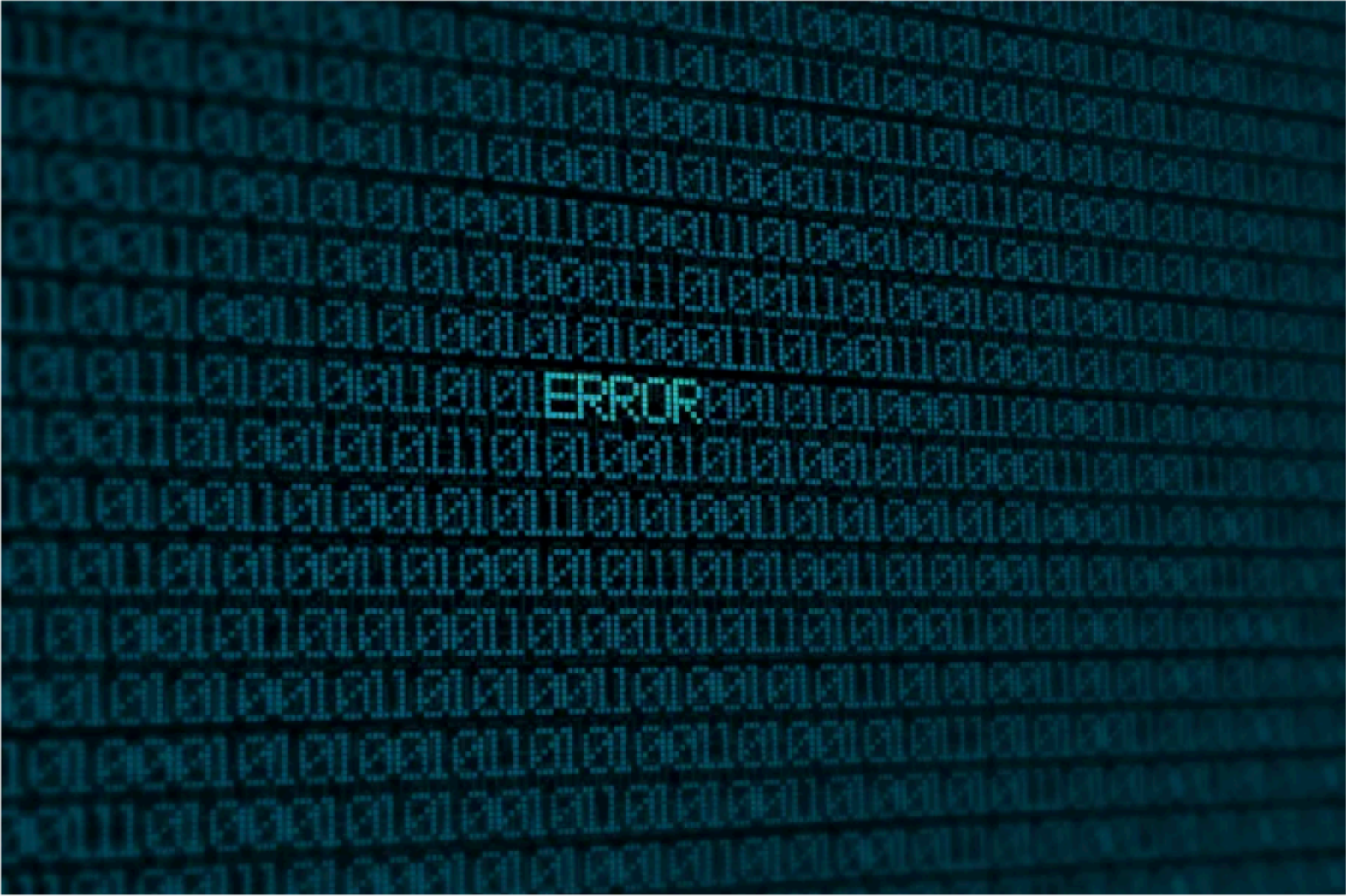
Follow Author

Published Aug 18, 2025 at 08:18am EDT, Updated Aug 18, 2025 at 03:00pm EDT

<

🔖

💬



Companies are pouring billions into enterprise AI, but poor data, weak infrastructure and missing strategy keep killing results. GETTY

FORBES' FEATURED VIDEO



# Reality



**elliot** 10:32

The email you sent Chris clearly looks like it came out of an LLM

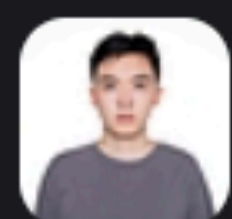
Dear Chris,

I hope this message finds you well. I am writing to follow up on the email I sent last Thursday (1-25) regarding technical support for my PGR project plan. I understand that you have a busy schedule, and I appreciate the time and attention you give to this matter. If you require any further information or clarification from my side, please let me know.

Thank you for your time and support, I am looking forward to hearing from you soon.

Best regards,  
Shiwen Qin

don't send stuff that looks like it came out of an LLM 😂



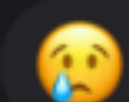
**Shiwen** 10:38

Sorry about that, will not use it to write emails anymore. Didn't know it can be spotted so easily 😂



**elliot** 10:38

It is painfully obvious.: )

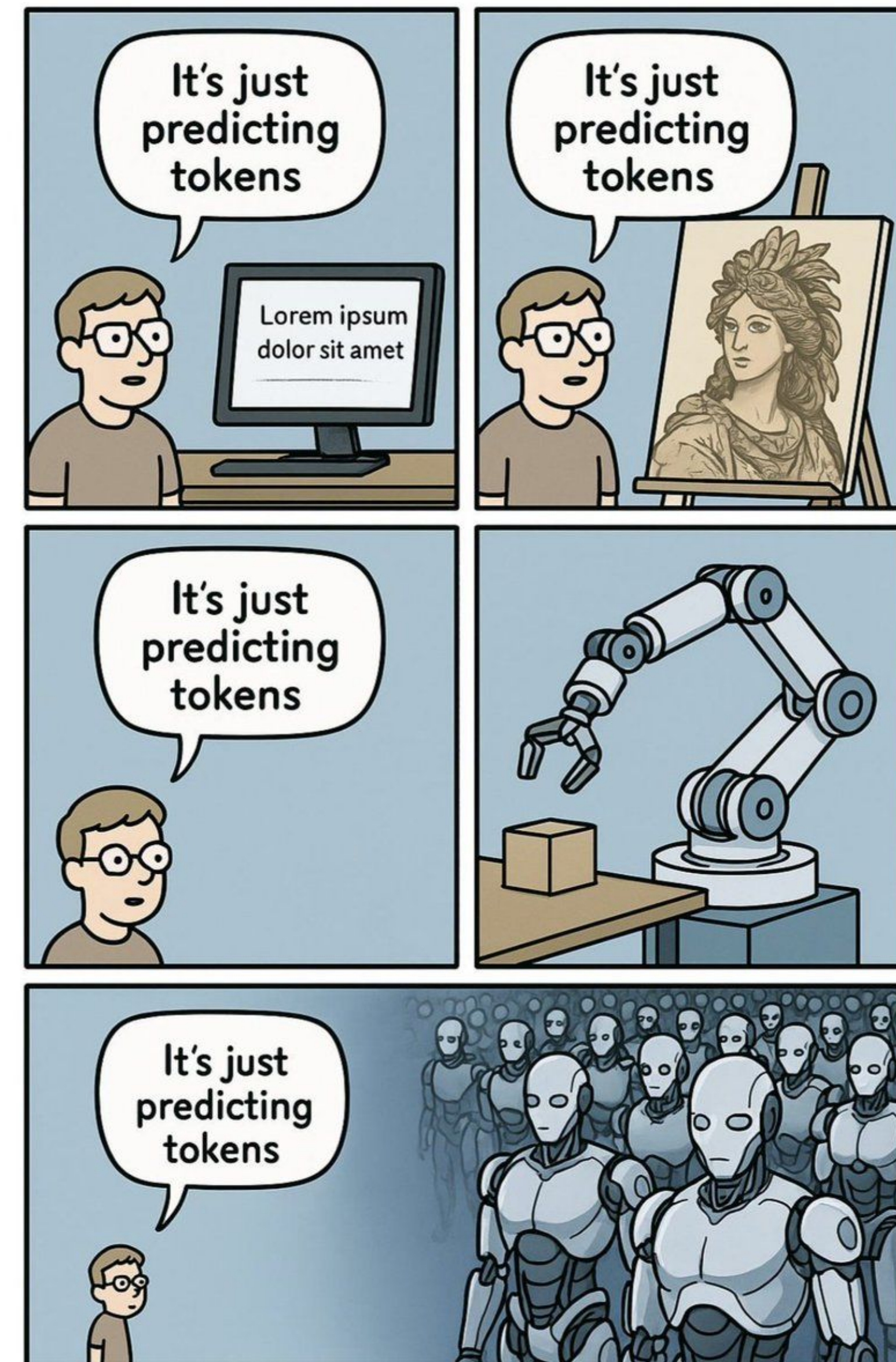


1





# Or...





# ML can do amazing things



A photo of a Corgi dog riding a bike in Times Square. It is wearing sunglasses and a beach hat.

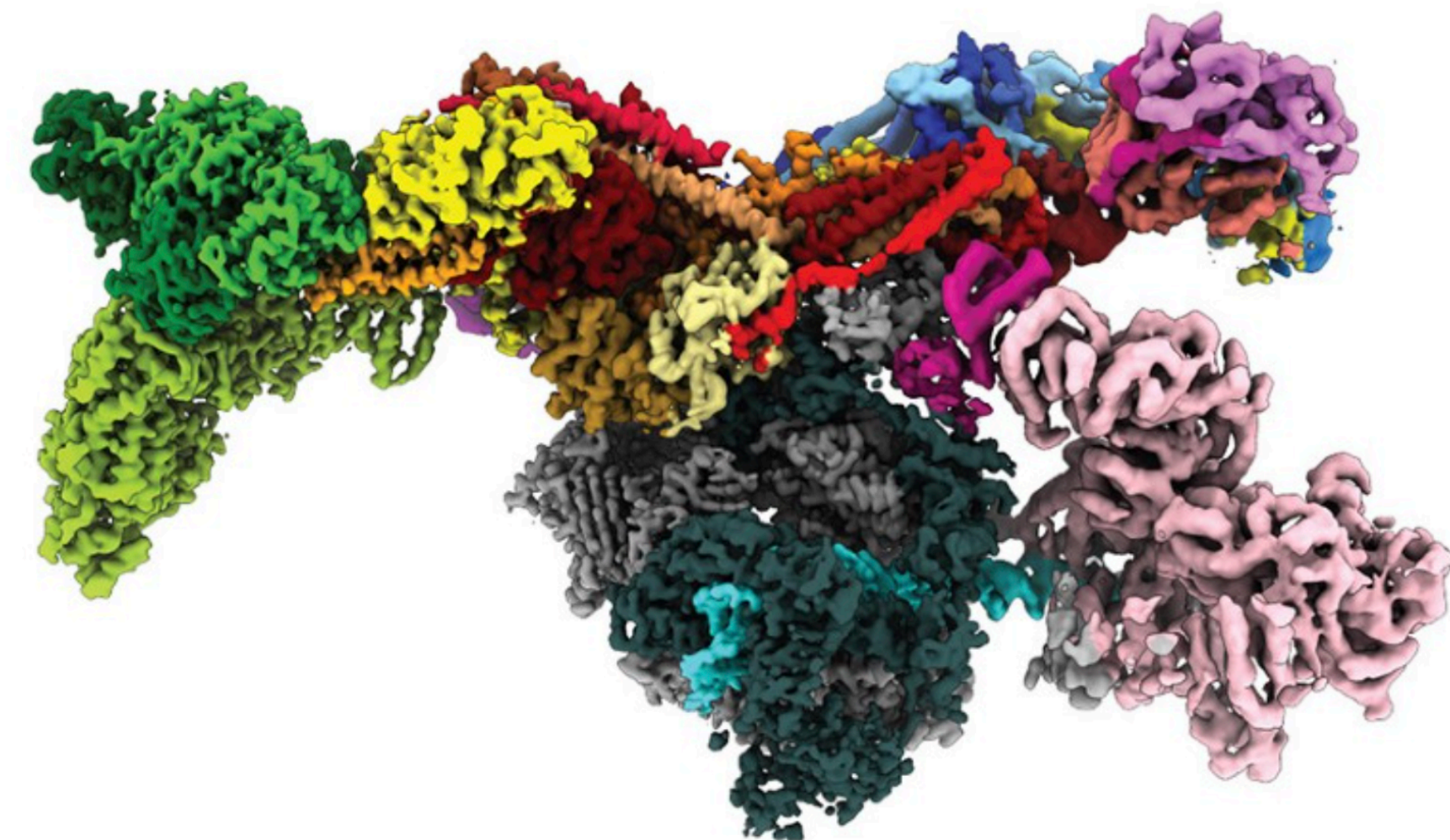
<https://imagen.research.google>

NEWS | 22 July 2021

## DeepMind's AI predicts structures for a vast trove of proteins

AlphaFold neural network produced a 'totally transformative' database of more than 350,000 structures from *Homo sapiens* and 20 model organisms.

[Ewen Callaway](#)



<https://www.nature.com/articles/d41586-021-02025-4>



# ML can do truly amazing things



Source: Someone in my D&D group via an “AI”



# But things can go wrong

- Machine learning models are **dumb**
- **They are not human!**
- They just learn to map inputs to outputs
- They don't care what the data is
- They don't care where you got the data
- They don't care what the task is

I am incredibly stupid.





# Wait a second... the models I've seen so far can't do that!

- Recent breakthroughs are largely driven by **deep neural networks**
- With modern compute, we can train very large, highly expressive models (often billions of parameters)
- When trained on massive datasets, they can learn very powerful feature representations that boost performance across many tasks
- That said, in many deployed applications (especially with tabular data), linear models and tree-based models work very well

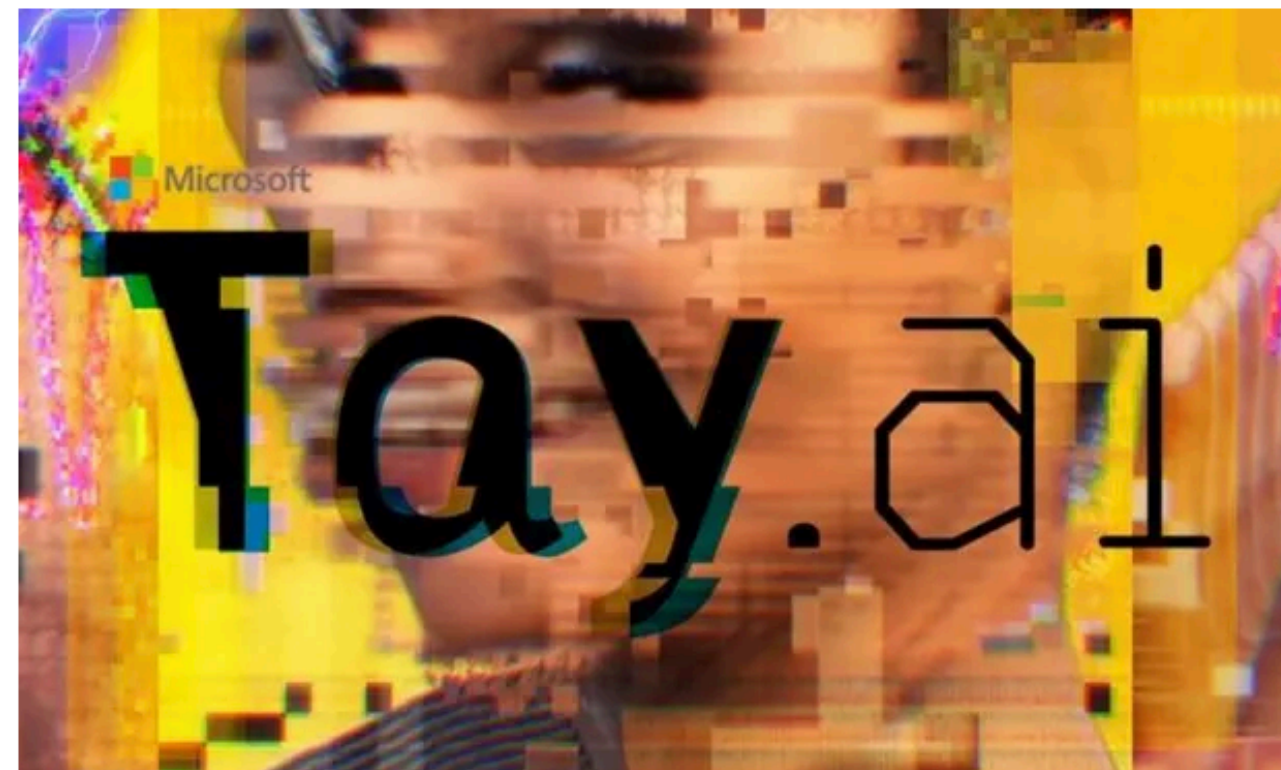


# Data bias

- The data in that training set may not be representative of the world (or the world one wants)
- It may contain biases - biased data gives a biased model

Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter

Attempt to engage millennials with artificial intelligence backfires hours after launch, with TayTweets account citing Hitler and supporting Donald Trump



Tay uses a combination of artificial intelligence and editorial written by a team including improvisational comedians. Photograph: Twitter

<https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter>



Sam Biddle

December 8 2022, 1:44 p.m.

## The Internet's New Favorite AI Proposes Torturing Iranians and Surveilling Mosques

ChatGPT, the latest novelty from OpenAI, replicates the ugliest war on terror-style racism.

<https://theintercept.com/2022/12/08/openai-chatgpt-ai-bias-ethics/>



# Data privacy

## Facebook sued for 'losing control' of users' data

🕒 9 February 2021



Facebook-Cambridge Analytica scandal

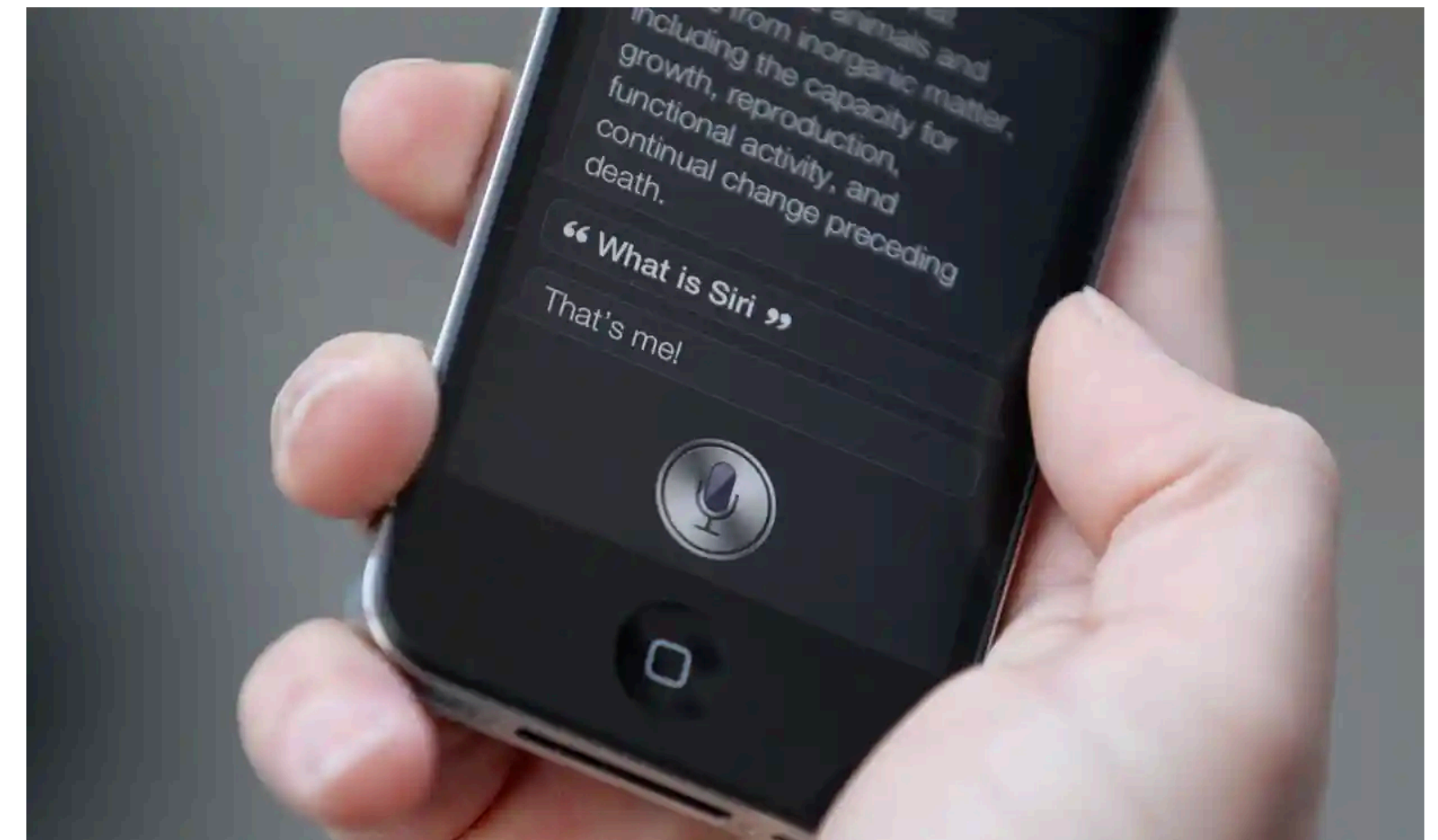


Facebook is being sued for "losing control" of the data of about a million users in England and Wales.

<https://www.bbc.co.uk/news/technology-55998588>

## Apple contractors 'regularly hear confidential details' on Siri recordings

Workers hear drug deals, medical details and people having sex, says whistleblower



📷 Workers heard the information when or providing quality control for Apple's Siri voice assistant. Photograph: Oli Scarff/Getty Images

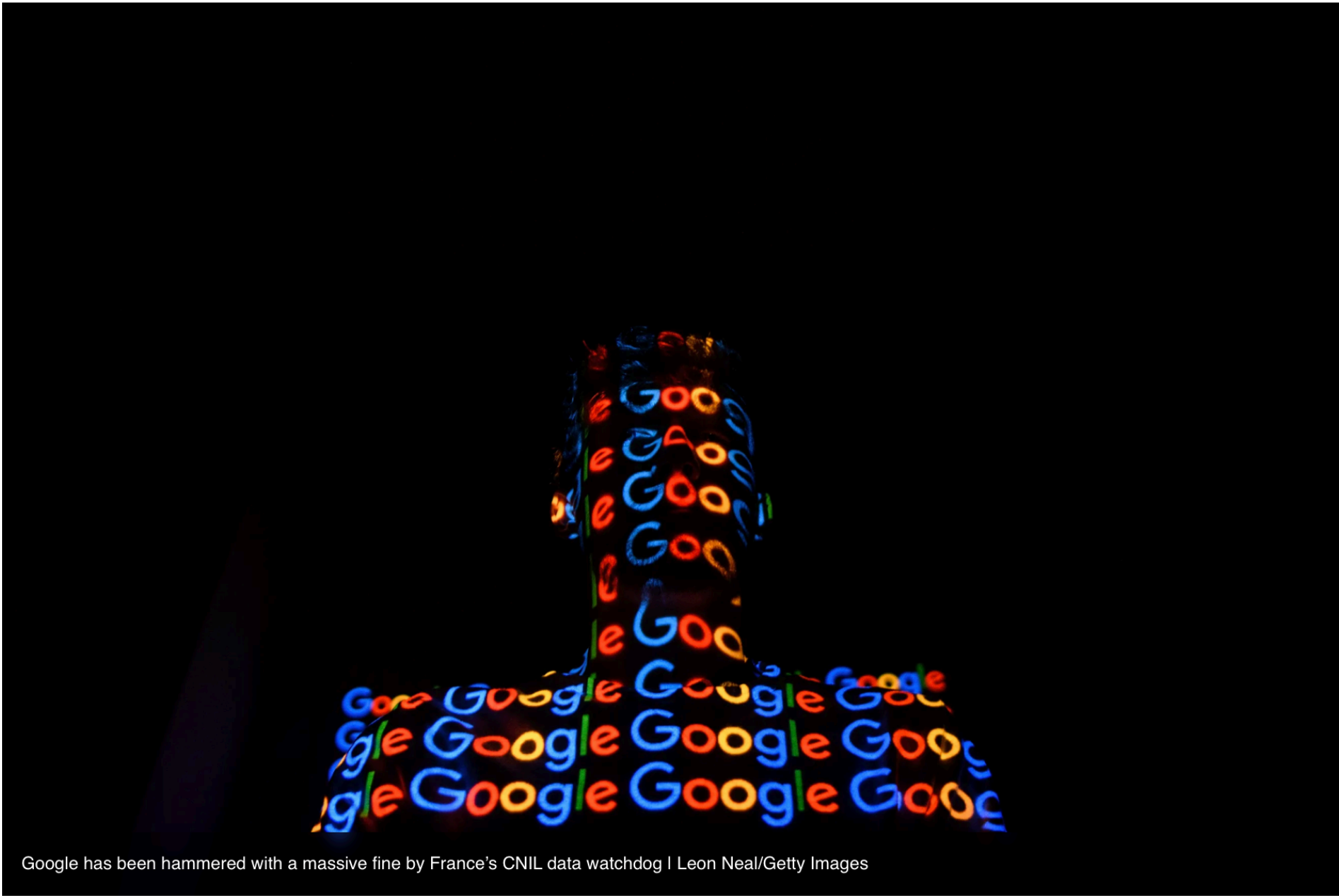
<https://www.theguardian.com/technology/2019/jul/26/apple-contractors-regularly-hear-confidential-details-on-siri-recordings>



# GDPR

## Google fine launches new era in privacy enforcement

The search giant is the first big tech company to feel the full brunt of GDPR enforcement. It won't be the last.



Google has been hammered with a massive fine by France's CNIL data watchdog | Leon Neal/Getty Images

### Fines and notices [\[ edit \]](#)

| Date ↕     | Organisation ↕                         | Amount ▼     | Issued by ↕       | Reason(s)                                                                                                                                                                                           |
|------------|----------------------------------------|--------------|-------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2023-05-12 | <a href="#">Meta Platforms</a>         | €1.2 billion | Ireland           | Transferring data from the European Union to the United States without adequate privacy protections <sup>[77][78]</sup>                                                                             |
| 2021-06-16 | <a href="#">Amazon Europe Core Srl</a> | €746,000,000 | Luxembourg (CNPD) | The largest fine for violating GDPR at the time. Related to <a href="#">targeted advertising</a> . <sup>[72][73]</sup>                                                                              |
| 2021-09-02 | <a href="#">WhatsApp Ireland Ltd</a>   | €225 M       | Ireland           | <sup>[74]</sup>                                                                                                                                                                                     |
| 2019-07-08 | <a href="#">British Airways</a>        | £183,000,000 | UK (ICO)          | Use of poor security arrangements that resulted in a 2018 <a href="#">web skimming</a> attack affecting 500,000 consumers. <sup>[24][25][26]</sup> Was later reduced to £20 million <sup>[27]</sup> |
| 2020-12-10 | <a href="#">Google LLC</a>             | €60,000,000  | France (CNIL)     | Deposit of cookies without obtaining consent, lack of information provided to users and defective "opposition" mechanism <sup>[48]</sup>                                                            |
| 2019-01-21 | <a href="#">Google LLC</a>             | €50,000,000  | France (CNIL)     | Insufficient transparency, control, and consent over the processing of personal data for the purposes of <a href="#">behavioural advertising</a> . <sup>[4][5]</sup>                                |
| 2020-12-10 | <a href="#">Google Ireland Limited</a> | €40,000,000  | France (CNIL)     | Deposit of cookies without obtaining consent, lack of information provided to users and defective "opposition" mechanism <sup>[48]</sup>                                                            |
| 2020-10-01 | <a href="#">H&amp;M</a>                | €35,300,000  | Germany (HmbBfDI) | Illegal surveillance of several hundred employees <sup>[46]</sup>                                                                                                                                   |

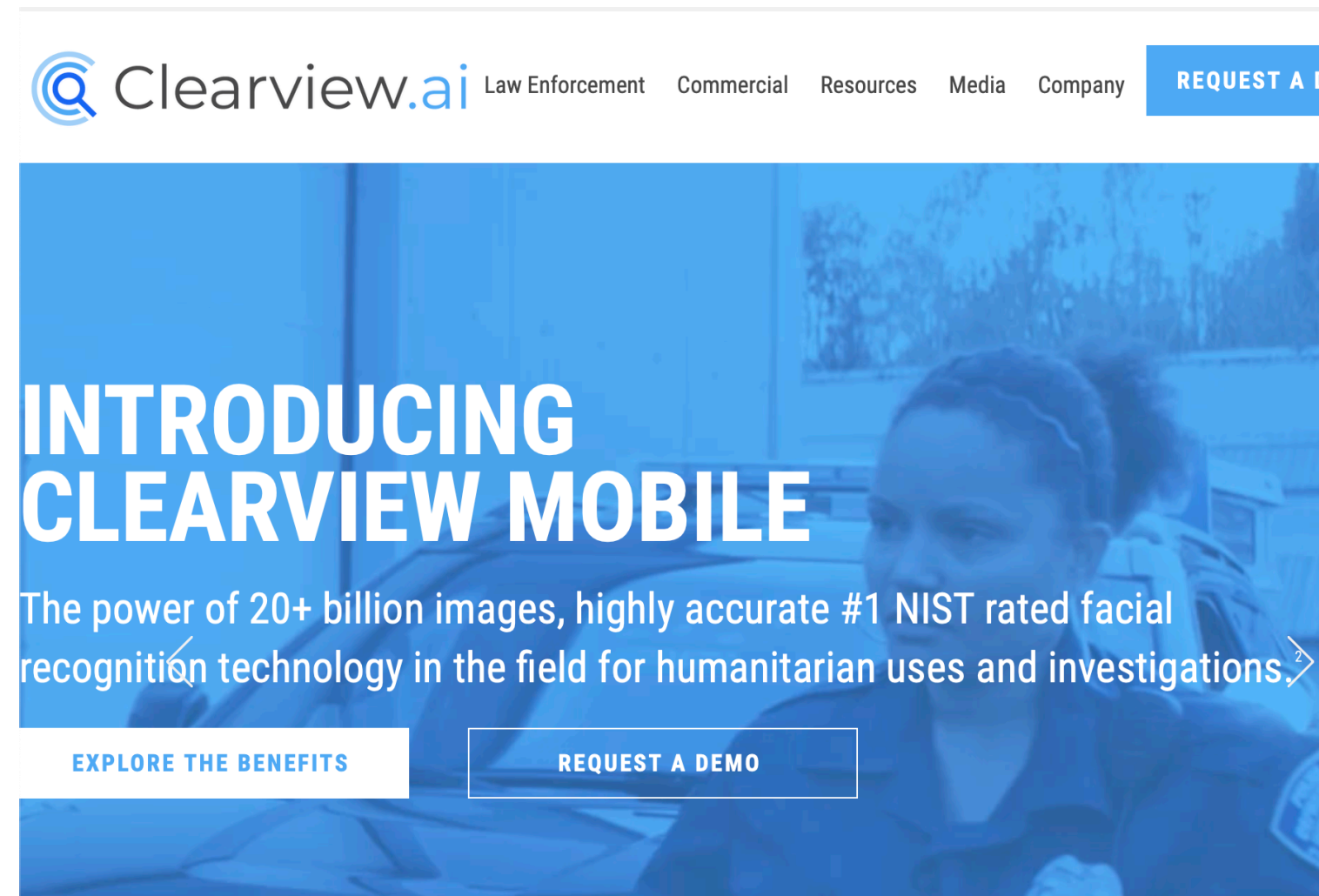


# Applications

- ML models can be used to spread misinformation
- ML models can be used for surveillance
- ML models can be used to kill people



<https://www.bbc.co.uk/news/articles/clykpwg8lpeo>



## Mohsen Fakhrizadeh: 'Machine-gun with AI' used to kill Iran scientist

7 December 2020

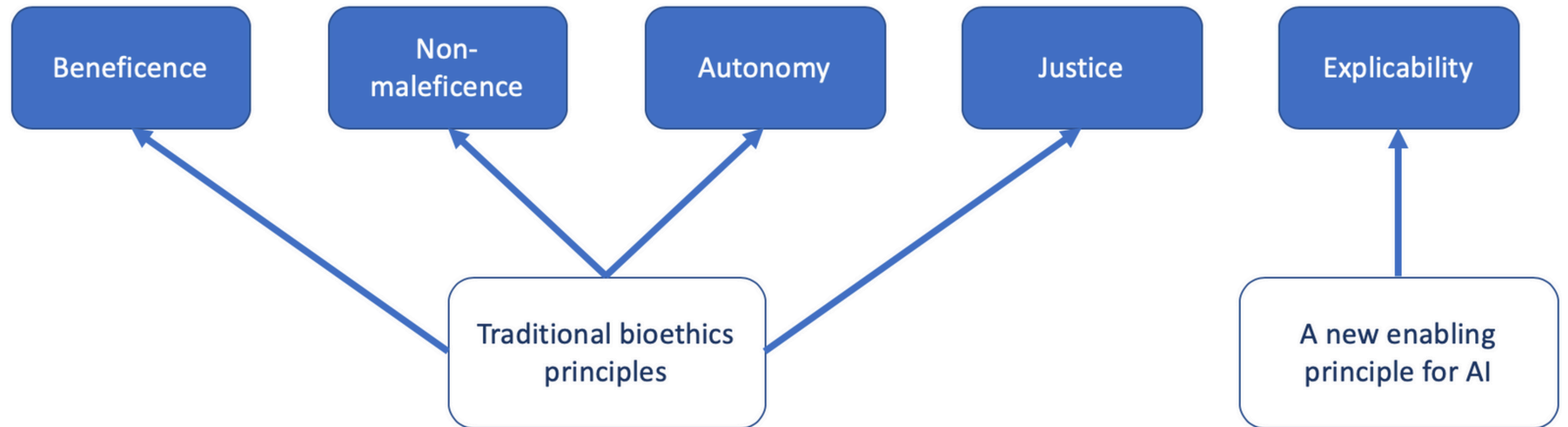


REUTERS

The Iranian authorities have put out conflicting accounts of how the scientist was killed



# Machine learning ethics



**A Unified Framework of Five Principles  
for AI in Society**

*by Luciano Floridi and Josh Cowls*

Published on Jul 01, 2019

<https://hdsr.mitpress.mit.edu/pub/l0jsh9d1/release/8>



do you know the 5 principles for AI in society by Floridi and Cowls?



Yes. Luciano Floridi and Josh Cowls (2019) propose **five ethical principles for AI in society**, adapted from bioethics:



1. **Beneficence**

AI should **promote well-being**—support human flourishing, social good, and beneficial outcomes.

2. **Non-maleficence**

AI should **avoid causing harm**—including safety risks, discrimination, privacy harms, manipulation, and other negative impacts.

3. **Autonomy**

AI should **respect and support human agency**—people should be able to make informed choices, retain meaningful control, and not be unduly coerced or nudged.

4. **Justice**

AI should be **fair and inclusive**—benefits and burdens should be distributed fairly, avoiding bias and unjust exclusion, and supporting due process where relevant.

5. **Explicability**

AI should be **understandable and accountable**—there should be intelligibility (some level of explanation/interpretability) and responsibility (clear accountability for outcomes).

If you want, I can also summarise how they connect these principles to policy/"AI4People" recommendations, or give examples of what each looks like in practice (e.g., in healthcare or recruitment).



## My attempt (no LM used!) - a quick ethical analysis of Language models in society

- **Beneficence** - It can increase productivity for some tasks. It allows people to ask questions they may otherwise be embarrassed to ask
- **Non-maleficence** - Companies may collect (and then train on) every single thing you type which, to me, feels like a violation of privacy. Usage may cause cognitive decline in offloading “simple” tasks
- **Autonomy** - I’m not sure. LLMs are being integrated into e.g. web search and are quite hard to avoid.
- **Justice** - Requires reliable internet access, is likely best for English speakers because there is the most training data
- **Explainability** - Not a chance in hell. There are trillions of parameters



# My attempt after I gave it to GPT 5.2

- **Beneficence** – Can increase productivity for some tasks (drafting, summarising, coding help). Can also lower barriers to asking “embarrassing” or sensitive questions by providing low-stigma support.
- **Non-maleficence** – Providers may log, retain, and potentially train on user inputs, creating privacy risks (secondary use, breaches, loss of control over sensitive data). Overuse may also lead to de-skilling/over-reliance by offloading routine cognitive work, and can cause harm when users trust plausible-but-wrong outputs.
- **Autonomy** – Hard to avoid as LLMs are being integrated into core services (e.g., web search, office suites). This can reduce meaningful choice/consent and may subtly steer user decisions through framing and confident-sounding answers.
- **Justice** – Benefits are unevenly distributed: requires reliable internet (and often paid access / newer devices) and tends to work best for English and other high-resource languages, disadvantaging speakers of low-resource languages/dialects and widening existing inequalities.
- **Explainability** – Strong “why this output?” explanations are not realistically available: models are extremely complex (billions–trillions of parameters) and outputs aren’t produced via human-interpretable reasoning, limiting transparency and contestability (even if some audit/documentation measures exist).



# Recruitment tools

- A machine learning model trained on predominantly male applicants
- Penalised CVs that included the word “women”

About 55% of US human resources managers said that AI would play a role in recruitment within the next five years, according to a survey by software firm CareerBuilder.

## Amazon scrapped 'sexist AI' tool

🕒 10 October 2018



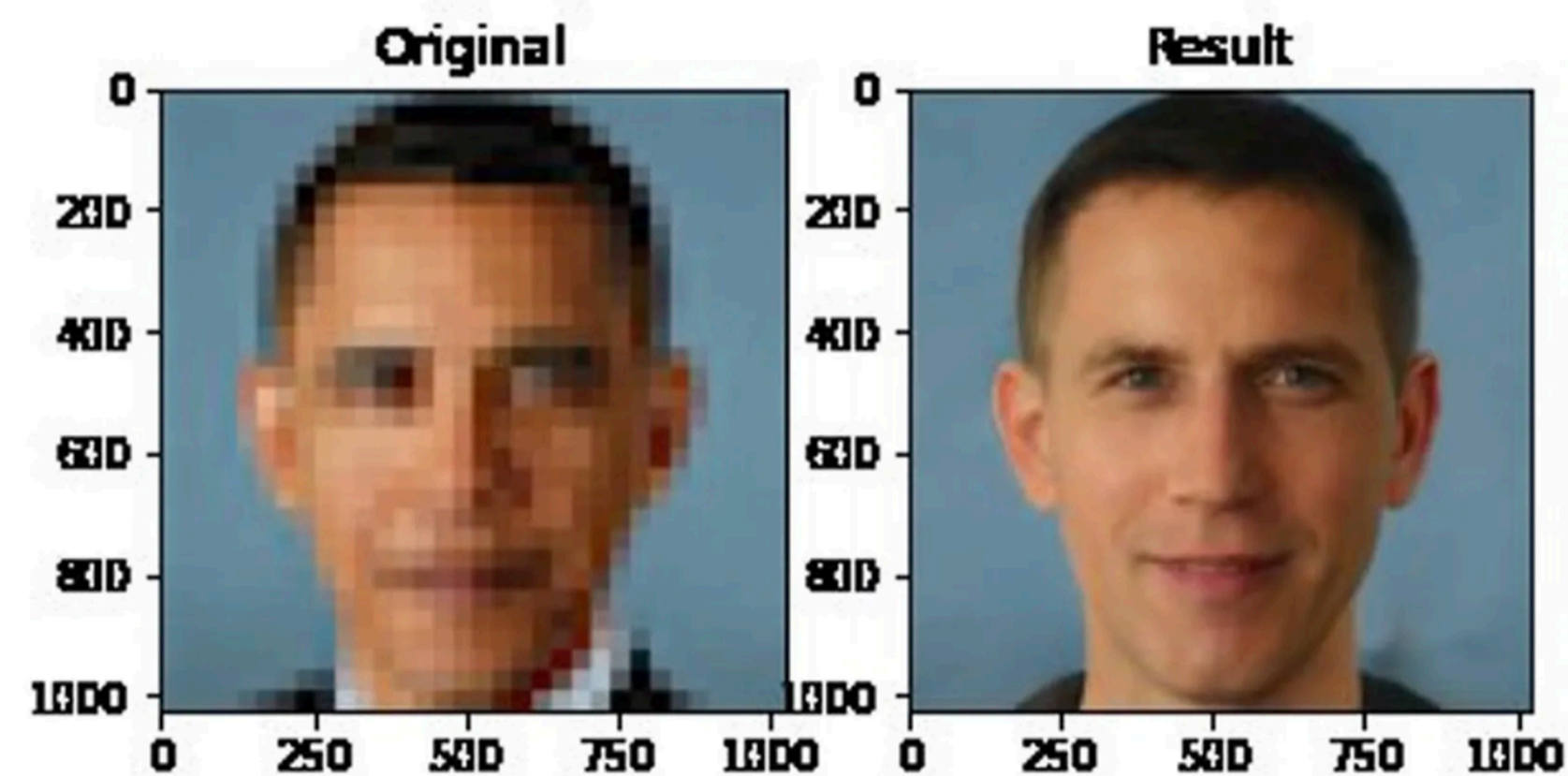
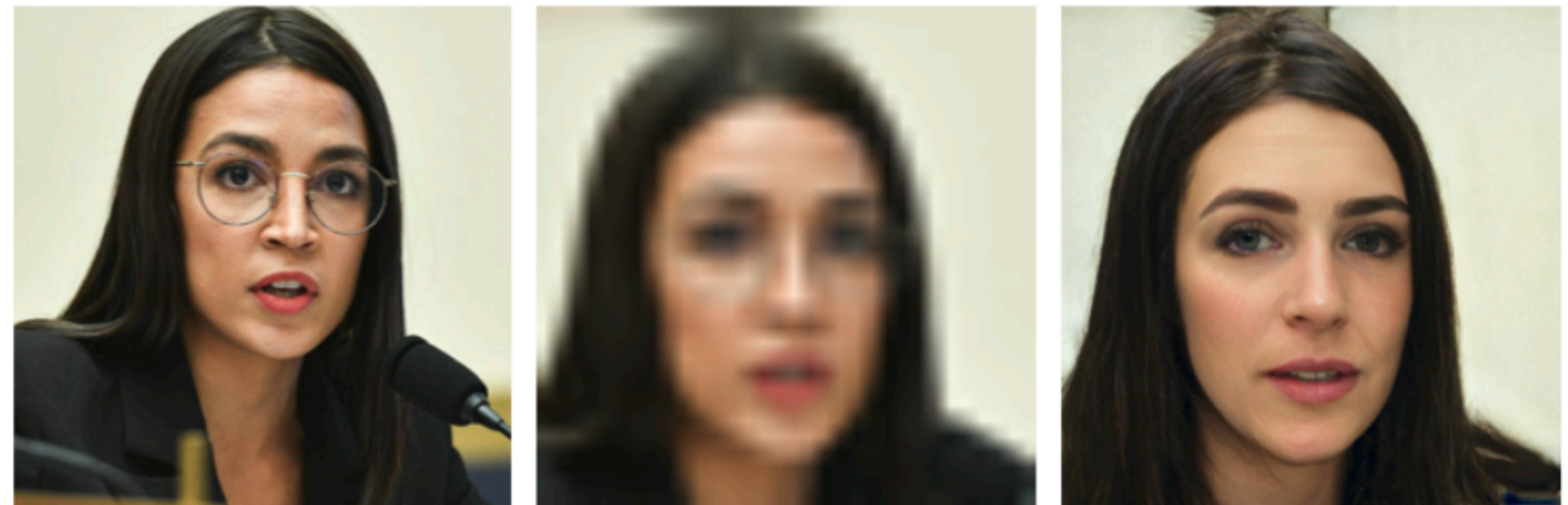
| The algorithm repeated bias towards men, reflected in the technology industry

<https://www.bbc.co.uk/news/technology-45809919>



# Generative models

- PULSE depixelizer generates high-res images of faces from low-res inputs
- It tends to generate white faces
- This could be because its training data is dominated by white faces





# Generative models

- But...





# Auto-grading

- A level results in England + Wales were (initially) algorithmically generated in 2020
- This disproportionately benefited privately educated students
- High-performing students at underperforming schools lost out
- Weight was placed on a school's historical performance





# Deepfakes

- ML models that generate realistic faces in photos and videos
- They have been used to create hoaxes and pornography



<https://thispersondoesnotexist.com>



<https://www.theguardian.com/music/2024/jan/26/taylor-swift-deepfake-pornography-sparks-renewed-calls-for-us-legislation>



# Advertising

- Google and Meta control a massive chunk of the world's digital advertising
- They have both invested heavily in machine learning
- An ML model can use **your** internet activity to profile **you!**
- This allows better targeting of adverts

GOOGLE ADS

Putting machine learning into the hands of every advertiser

Jul 10, 2018 · 4 min read



Jerry Dischler  
Vice President / General  
Manager, Ads

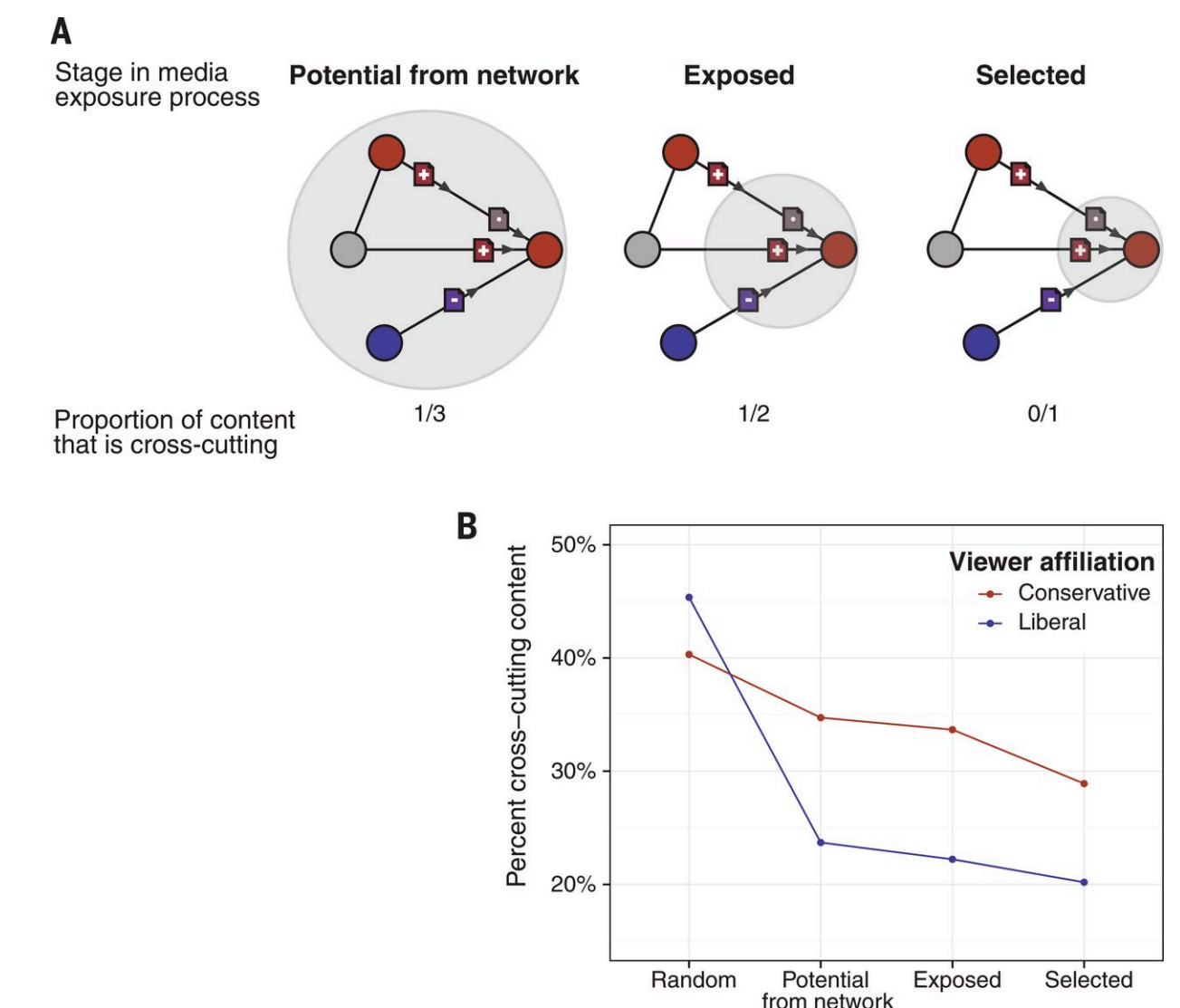
Share





# Recommender systems

- Lots of people get their news from social media
- ML models can decide which content to present to users
- Users are presented with content that aligns with their views
- Is this ideal?
- Could this amplify radical views? (...Yes)





# Cambridge Analytica

## The Cambridge Analytica files: the story so far

**What is the company accused of, how is Facebook involved and what is the Brexit link?**



Facebook ran adverts on Sunday in several UK and US newspapers apologising for the data breach. Photograph: Dominic Lipinski/PA

### **What are the allegations against Cambridge Analytica?**

The data analytics firm used personal information harvested from more than 50 million Facebook profiles without permission to build a system that could target US voters with personalised political advertisements based on their psychological profile, according to Christopher Wylie, a former Cambridge Analytica contractor who helped build the algorithm. Employees of Cambridge Analytica, including the suspended CEO Alexander Nix, were also filmed boasting of using manufactured sex scandals, fake news and dirty tricks to swing elections around the world.



# Lighting the fuse?

## ‘Storm the Capitol’: Violence organised on social media as warnings of far-right post-election went unheard

Federal law enforcement has warned for years about far-right threats. Trump supporters openly planned armed insurrection for weeks, writes **Alex Woodward**

Friday 08 January 2021 21:31 • [Comments](#)



**Dan Milmo** *Global technology correspondent*

Mon 6 Dec 2021 17.03 GMT



## Rohingya sue Facebook for £150bn over Myanmar genocide

**Victims in US and UK legal action accuse social media firm of failing to prevent incitement of violence**



Residents of the Rohingya refugee camp in Cox's Bazar, Bangladesh. Photograph: Tanbirul Miraj Ripon/EPA

Facebook’s negligence facilitated the **genocide of Rohingya Muslims in Myanmar** after the social media network’s algorithms amplified hate speech and the platform failed to take down inflammatory posts, according to legal action launched in the US and the UK.

<https://www.theguardian.com/technology/2021/dec/06/rohingya-sue-facebook-myanmar-genocide-us-uk-legal-action-social-media-violence>

<https://www.independent.co.uk/news/world/americas/us-politics/capitol-riot-was-openly-organized-on-mainstream-social-media-b1784703.html>

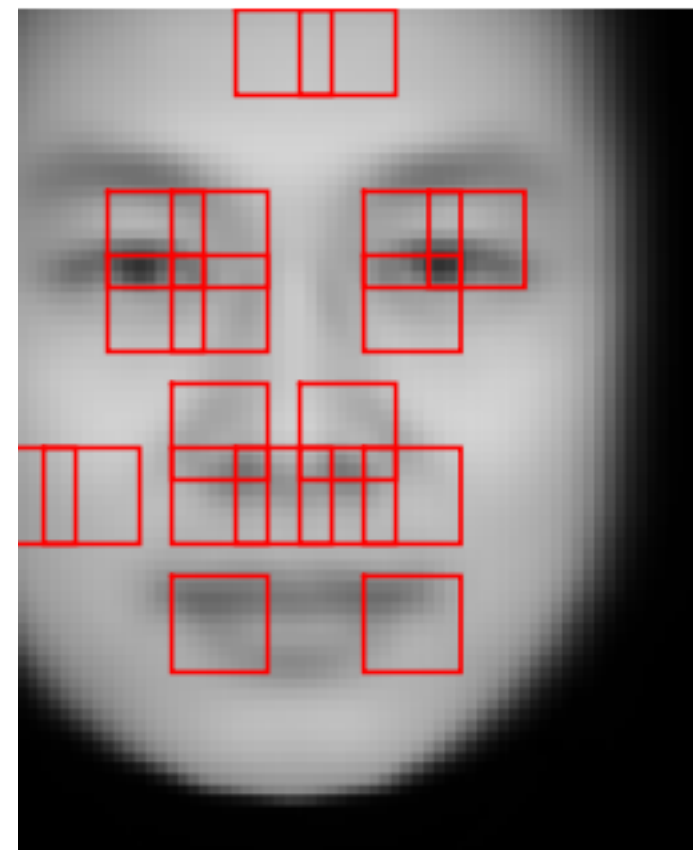
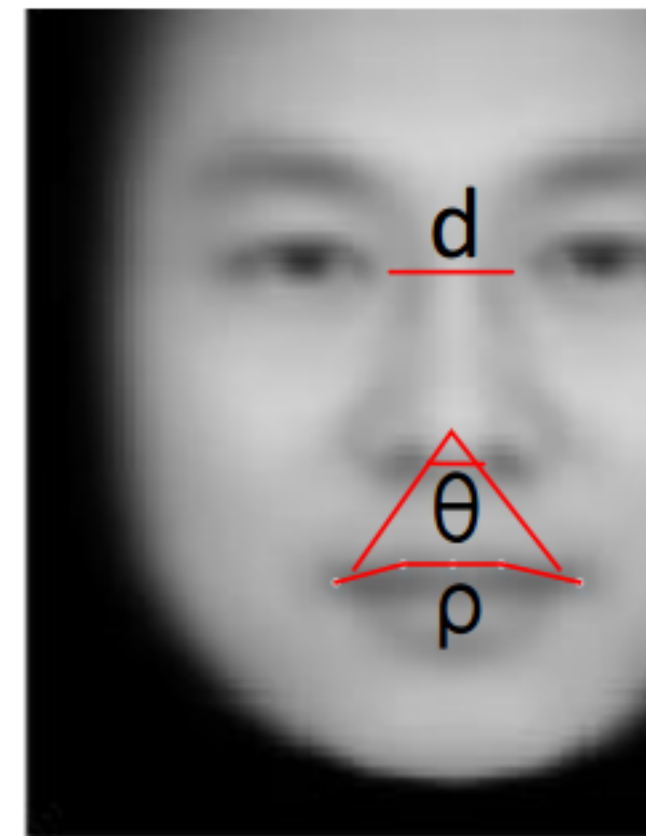


# Contentious applications

## ARTIFICIAL INTELLIGENCE

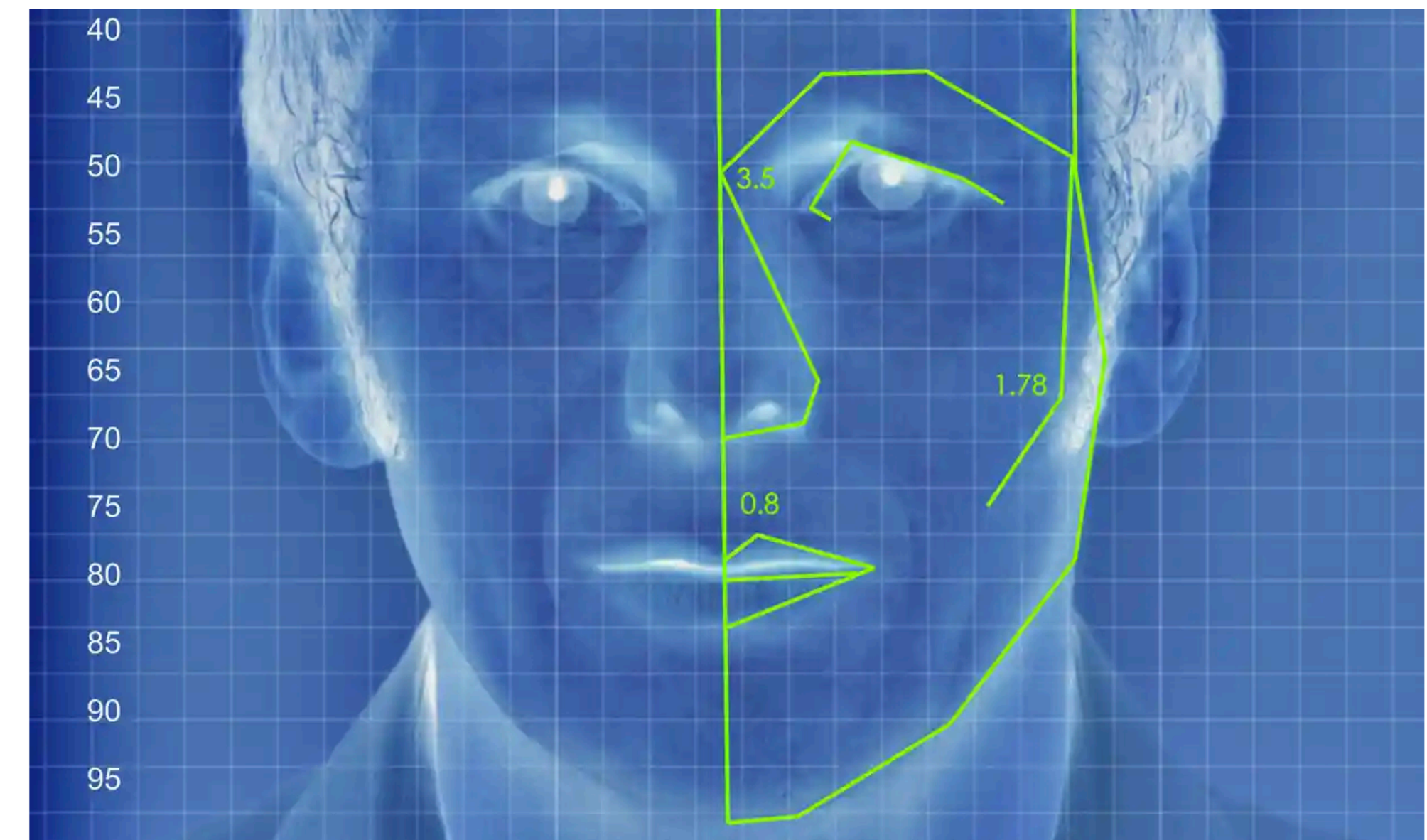
### Neural Network Learns to Identify Criminals by Their Faces

The effort aimed at identifying criminals from their mugshots raises serious ethical issues about how we should use artificial intelligence.



### New AI can guess whether you're gay or straight from a photograph

An algorithm deduced the sexuality of people on a dating site with up to 91% accuracy, raising tricky ethical questions



📷 An illustrated depiction of facial analysis technology similar to that used in the experiment.  
Illustration: Alamy

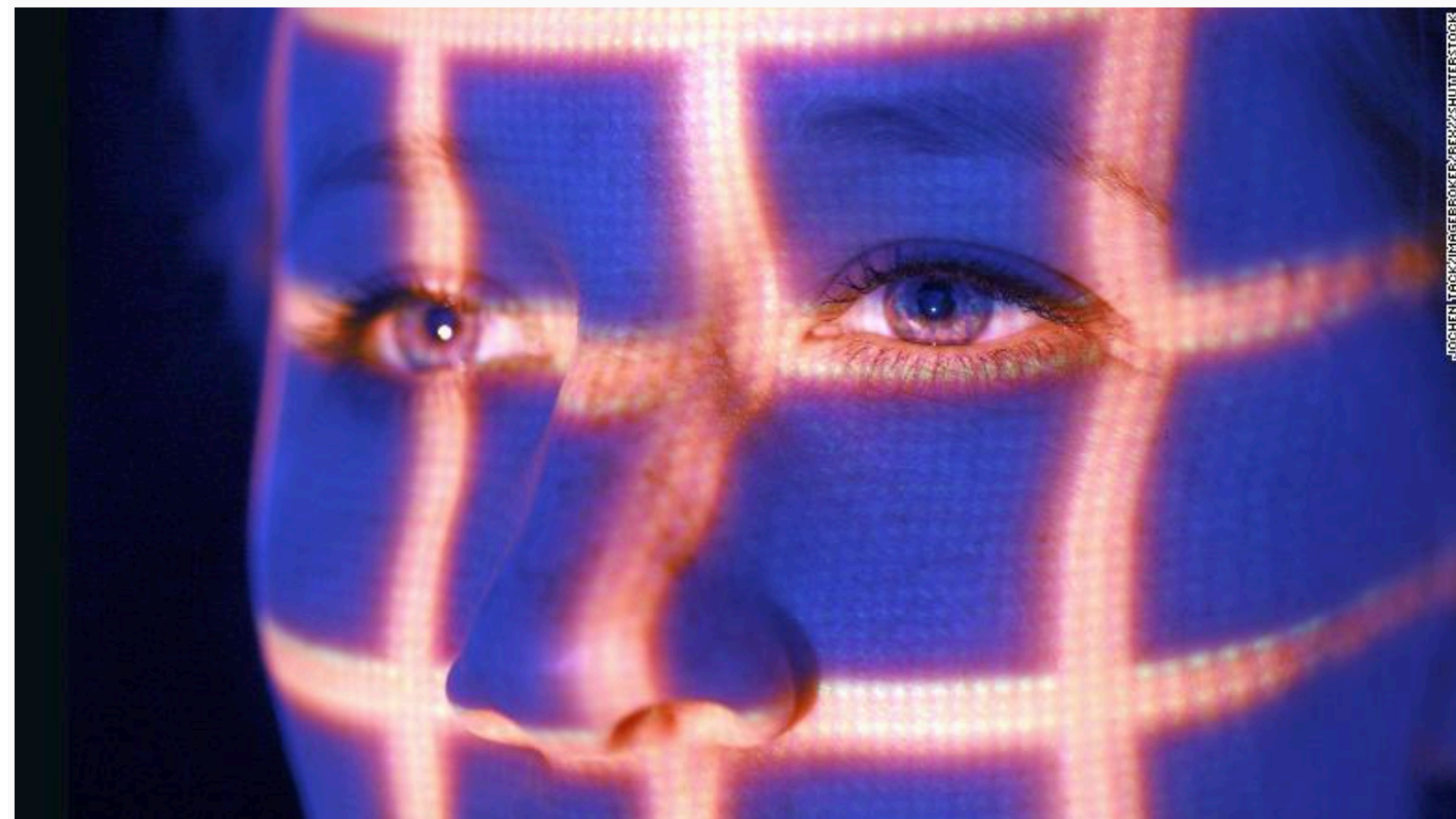


# Contentious applications

## AI technology can identify genetic diseases by looking at your face, study says

By Nina Avramova, CNN

🕒 Updated 2117 GMT (0517 HKT) January 8, 2019



New AI technology could identify rare genetic diseases from patients' facial images.

**(CNN)** — A new artificial intelligence technology can accurately identify some rare genetic disorders using a photograph of a patient's face, according to a new study.

The AI technology, called DeepGestalt, outperformed clinicians in identifying a range of syndromes in three trials and could add significant value in personalized care, according to the [study](#) published Monday in the journal Nature Medicine.



**FACEPTION IS A FACIAL PERSONALITY ANALYTICS TECHNOLOGY COMPANY**

We reveal personality from facial images at scale to revolutionize how companies, organizations and even robots understand people and dramatically improve public safety, communications, decision-making, and experiences.



# Plagiarism?

Artificial  
intelligence (AI)

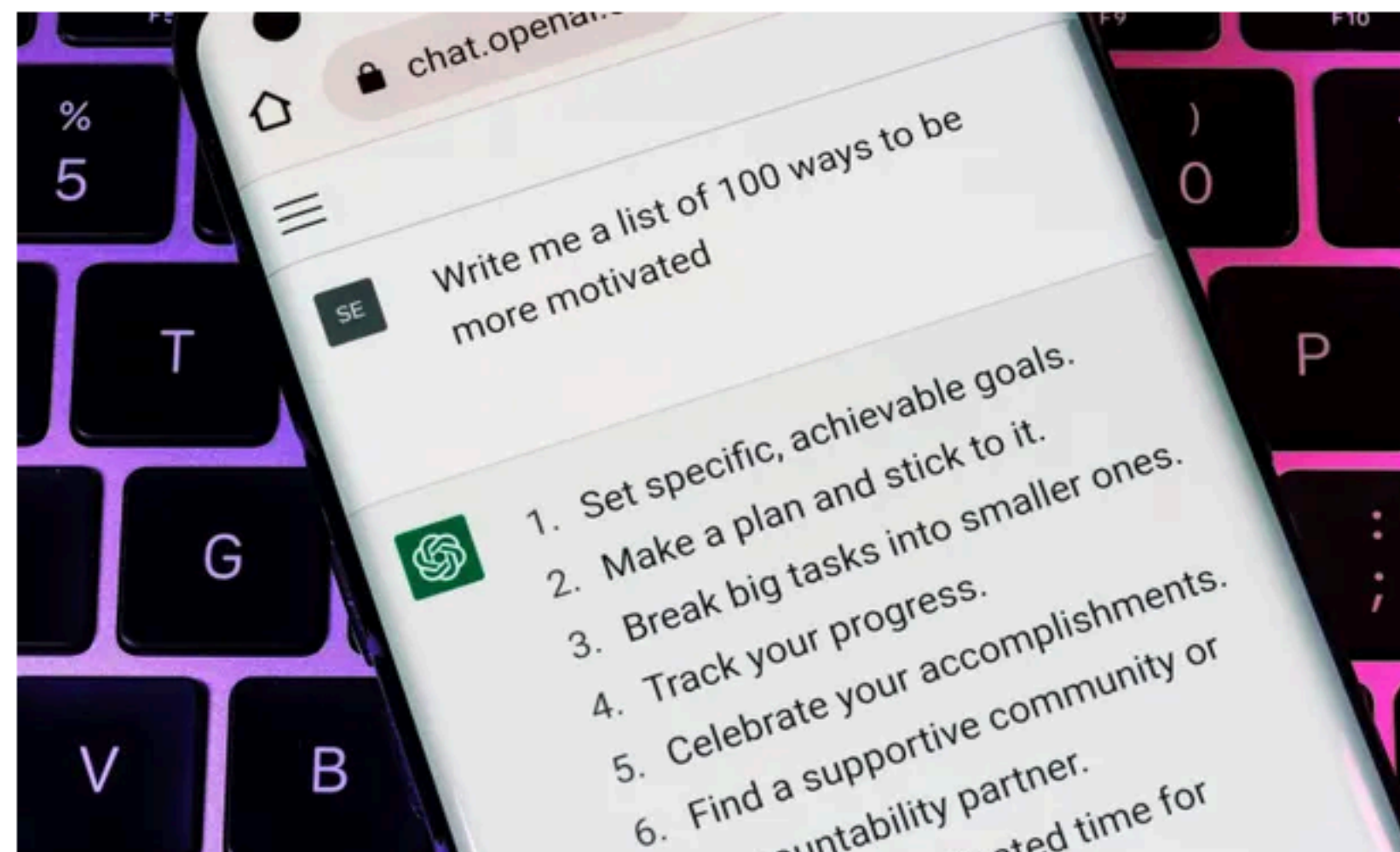
## Lecturers urged to review assessments in UK amid concerns over new AI tool

ChatGPT is capable of producing high-quality essays with minimal human input

- **ChatGPT: what can the extraordinary artificial intelligence chatbot do?**

Sally Weale  
Education  
correspondent

Fri 13 Jan 2023 16.23 GMT



ChatGPT has already triggered concerns about the potential for hard-to-detect plagiarism and questions about the validity of the essay as a future form of assessment. Photograph: Ascannio/Alamy

Lecturers at UK universities have been urged to review the way in which their courses are assessed amid concerns that students are already using a potent new AI tool capable of **producing high-quality essays** with minimal human input.

Did you write that yourself?  
Student using ChatGPT:





# Cognitive decline? :(

## MIT Study Finds Artificial Intelligence Use Reprograms the Brain, Leading to Cognitive Decline

FEED

By Nicolas Hulscher, MPH

A new MIT study titled, Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task, has found that using ChatGPT to help write essays leads to long-term cognitive harm—measurable through EEG brain scans. Students who repeatedly relied on ChatGPT showed **weakened neural connectivity, impaired memory recall, and diminished sense of ownership over their own writing**. While the AI-generated content often scored well, the brains behind it were shutting down.

### Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task<sup>Δ</sup>

|                                                                          |                                                  |                                                    |                                               |
|--------------------------------------------------------------------------|--------------------------------------------------|----------------------------------------------------|-----------------------------------------------|
| Nataliya Kosmyna <sup>1</sup><br>MIT Media Lab<br>Cambridge, MA          | Eugene Hauptmann<br>MIT<br>Cambridge, MA         | Ye Tong Yuan<br>Wellesley College<br>Wellesley, MA | Jessica Situ<br>MIT<br>Cambridge, MA          |
| Xian-Hao Liao<br>Mass. College of Art and Design (MassArt)<br>Boston, MA | Ashly Vivian Beresnitzky<br>MIT<br>Cambridge, MA | Iris Braunstein<br>MIT<br>Cambridge, MA            | Pattie Maes<br>MIT Media Lab<br>Cambridge, MA |

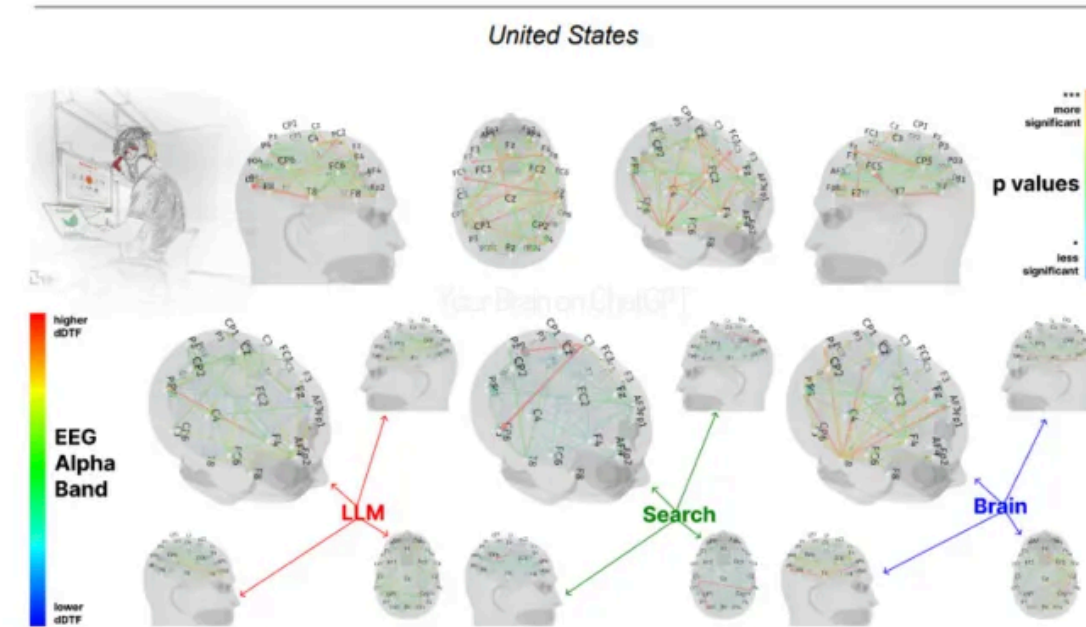
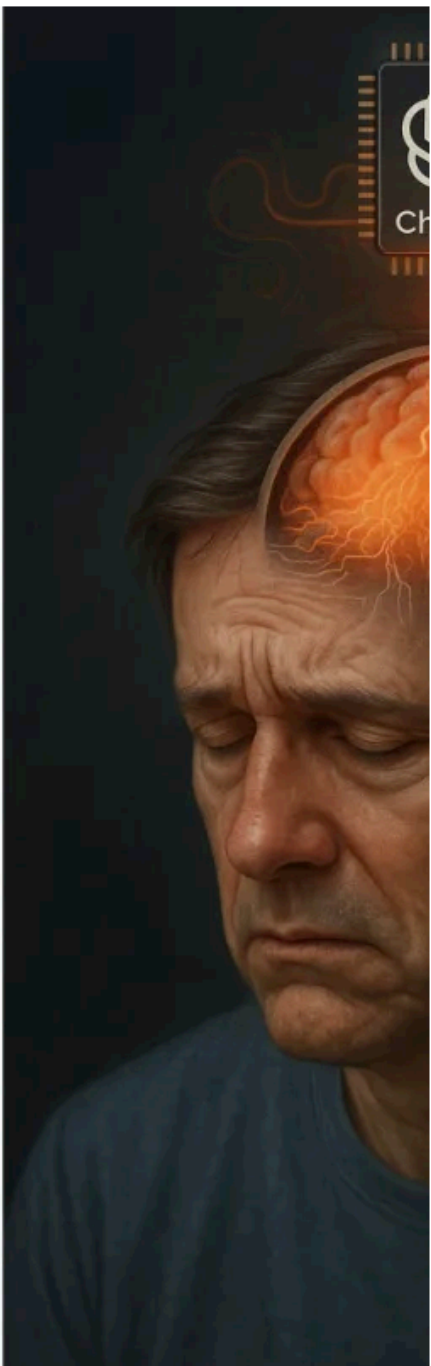


Figure 1. The dynamic Direct Transfer Function (dDTF) EEG analysis of Alpha Band for groups: LLM, Search Engine, Brain-only, including p-values to show significance from moderately significant (\*) to highly significant (\*\*\*).



<https://publichealthpolicyjournal.com/mit-study-finds-artificial-intelligence-use-reprograms-the-brain-leading-to-cognitive-decline/>



# Censorship and surveillance

COMPUTING

## How WeChat censors private conversations, automatically in real time

The super app instantly blocks even the images for over 1 billion users and growing.

By Patrick Howell O'Neill

July 15, 2019

GLOBAL

## China's New Frontiers in Dystopian Tech

Facial-recognition technologies are proliferating, from airports to bathrooms.

By Rene Chun

<http://tinyurl.com/2s3jffuz>

<https://tinyurl.com/4hf77rdt>

SURVEILLANCE

## Miami Police Used Clearview AI Facial Recognition in Arrest of Homeless Man

Facial recognition technology is increasingly being deployed by police officers across the country, but the scope of its use has been hard to pin down.

C.J. CIARAMELLA | 1.19.2024 2:37 PM

**MOTHERBOARD**  
TECH BY VICE

## Pentagon Wants to Predict Anti-Trump Protests Using Social Media Surveillance

A series of research projects, patent filings, and policy changes indicate that the Pentagon wants to use social media surveillance to quell domestic insurrection and rebellion.

<http://tinyurl.com/yeyvs6dz>

<https://tinyurl.com/wtzt7me3>



# Weaponry

## SLAUGHTERBOTS ARE HERE.

The era in which algorithms decide who lives and who dies is upon us.  
We must act now to prohibit these weapons.

↓ LEARN MORE

<https://autonomousweapons.org>



# Accelerating the climate crisis

ENERGY, FUTURE, POLLUTION

## The Green Dilemma: Can AI Fulfil Its Potential Without Harming the Environment?

CRISIS - ATMOSPHERIC CO2 LEVELS

CRISIS - POLLUTION CRISES

BY ALOKYA KANUNGO

GLOBAL COMMONS

JUL 18TH 2023

5 MINS



EARTH.ORG IS POWERED BY OVER 150 CONTRIBUTING WRITERS



<https://earth.org/the-green-dilemma-can-ai-fulfil-its-potential-without-harming-the-environment/>



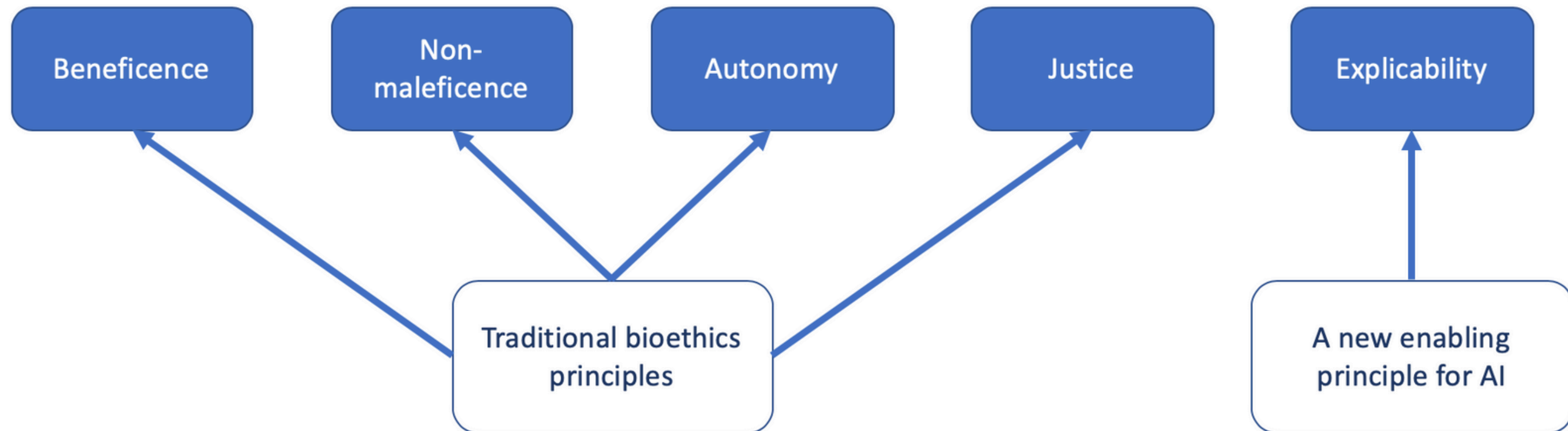
# Summary

- We looked at some “AI” hype and reality
- We considered data bias and privacy
- We learnt about an ethical framework that can be used for ML
- We considered some example applications and issues of using ML in society



Q.

- Let's find / propose an ML scenario and critique it according to Floridi and Cowls



A Unified Framework of Five Principles  
for AI in Society

by Luciano Floridi and Josh Cowls

Published on Jul 01, 2019

<https://hdsr.mitpress.mit.edu/pub/l0jsh9d1/release/8>