



**Trinity College Dublin**

Coláiste na Tríonóide, Baile Átha Cliath

The University of Dublin

THE LIBRARY OF TRINITY COLLEGE DUBLIN, THE  
UNIVERSITY OF DUBLIN  
*LEABHARLANN CHOLÁISTE NA TRÍONÓIDE,  
OLLSCOIL ÁTHA CLIATH*

# Demystifying AI

Jessica Leonard

2024

# Workshop Overview

The Berkeley



Trinity  
College  
Dublin

The University of Dublin

- Intro to GenAI
- An overview of how it works
- Ethical concerns with using AI
- Workshop Question: Can we use AI ethically, and if so, what would ethical use of AI look like?

## What are some of your concerns about AI?

- Write for 30 seconds
- Talk to the person next to you
- Share with the group

# Traditional AI vs. GenAI



The University of Dublin

Traditional AI - if a trigger is met, the algorithm will execute a set of instructions.

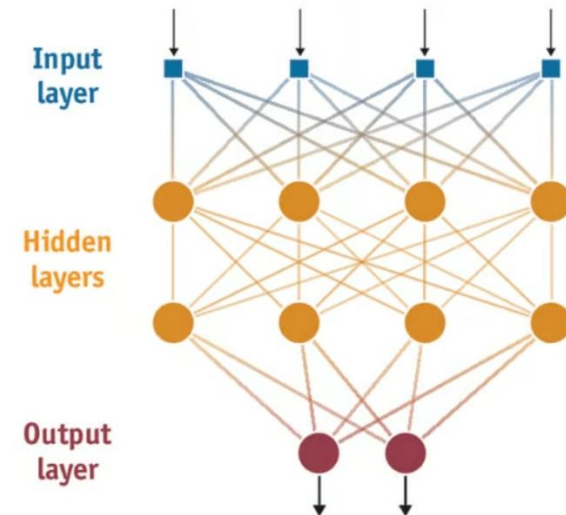
Examples: Siri, Alexa, Netflix recommendations, Google search, predictive text, self driving vehicles

GenAI - sets of instruction that allow machines to “analyse” data, “make decisions”, and generate something new.

Examples: ChatGPT, DALL-E, Bard, Microsoft Co-pilot, image and text generation

- Machine Learning
  - Defined in the 1950s by AI pioneer Arthur Samuel as “the field of study that gives computers the ability to learn without explicitly being programmed.”
- Neural Networks
  - First proposed in 1944, now Deep Learning
  - Algorithms that train on large data sets over and over in order to detect patterns and relationships and continually improve\*

How an artificial neural network processes data



[From not working to neural networking \(economist.com\)](http://economist.com)

# MACHINE LEARNING AND NEURAL NETWORKS



The University of Dublin

- Generative AI uses machine learning to process a massive amount of visual or textual data and then determines what things are most likely to appear near other things.



Katelyn Chedraoui via Leonardo AI/CNET



# LARGE LANGUAGE MODELS (LLMs)

- An LLM is a machine-learning neuro network trained through data input/output sets.
- LLM is a probability engine:
  - Assigns a percentage to each possible answer
  - Does not have an internal state representation of the world
  - Based on probability (% of likelihood that things (letters, words, pixels, etc.) will occur near one another





## Process a massive amount of visual or textual data

Visual and textual data comes from internet users:  
Wikipedia, Facebook, LinkedIn, Reddit, artwork, articles,  
posts, etc.



# DATA PRIVACY: A BLACK BOX UNDER CURRENT REGULATION



The University of Dublin

– Intellectual property and personal information (including medical and financial) are used to train AI

Example: ChatGPT's data collection practices:

- Trained on massive amount of personal data
- No legal basis
- People whose data were used were not notified
- No clear security of data collected and stored by the AI
- No data confidentiality

# PLAGIARISM AND COPYRIGHT

- Plagiarism “**the three ‘C’s**”:
  - 1) **Consent** not given to have copyrighted work included
  - 2) **Compensation** not received
  - 3) **Credit** not given when A.I. text images were produced using their work

“A.I. generators could not operate without the labour of humans who unwittingly provide source material.”

[Kyle Chayka, Is A.I. Art Stealing from Artists? | The New Yorker](#)



Greg Rutkowski's name has been used as a prompt in AI tools more than 400,000 times without his consent

# PLAGIARISM AND COPYRIGHT



- Rejected copyright protection for an award-winning image generated using Midjourney
- It "put me in a terrible position, with no recourse against others who are blatantly and repeatedly stealing my work."

[Blake Brittain, Artist sues after US rejects copyright for AI-generated image | Reuters](#)

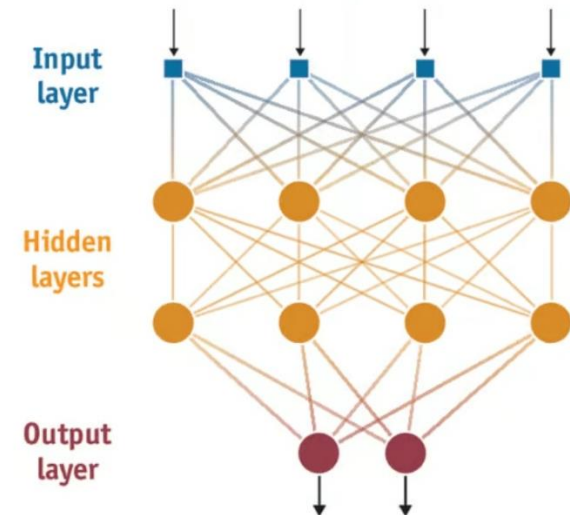


Jason M. Allen won art competition with GenAI image



- Organised into layers
- Information fed into input layer
- Training the network involves adjusting internal weights until particular inputs elicit desired response

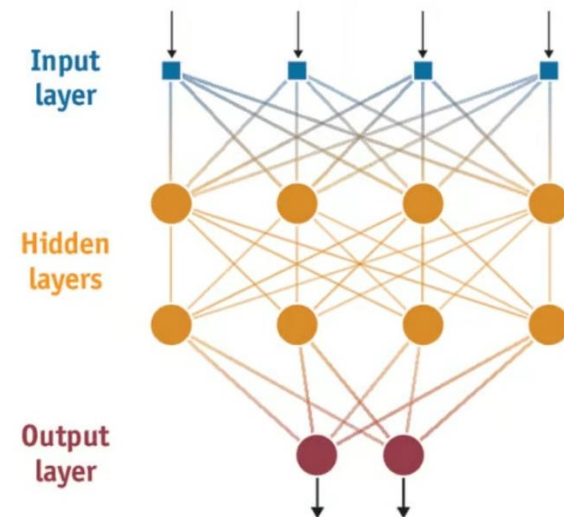
How an artificial neural network processes data



[From not working to neural networking \(economist.com\)](http://economist.com)

- Subset of machine learning based on neural networks.
- Deep refers to layers in the network (can be hundreds or thousands).
- Weights – signify importance of each input.
- Bias – shifts activation on each plane.

How an artificial neural network processes data

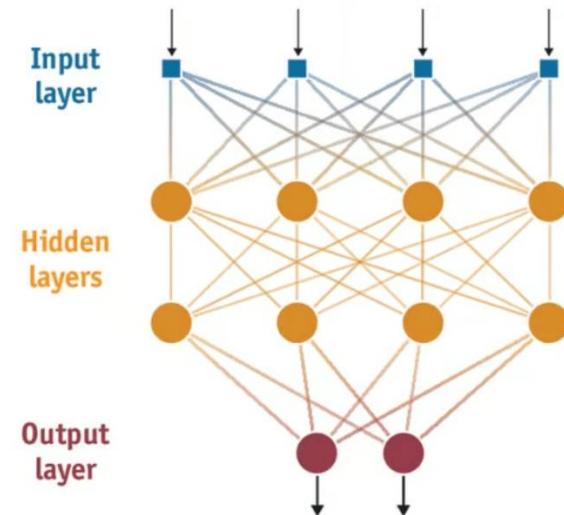


[From not working to neural networking \(economist.com\)](http://economist.com)



- Based on the data ingested from the internet and programmers
- Language models are used in real-world settings:
  - hiring,
  - facial recognition
  - medical diagnoses
  - policing
  - gender classification

How an artificial neural network processes data



[From not working to neural networking \(economist.com\)](http://economist.com)



Nature

<https://www.nature.com> › ... › review articles

Review Article | [Open access](#) | Published: 13 September 2023

## Ethics and discrimination in artificial intelligence-enabled recruitment practices

However, algorithmic bias results in discriminatory hiring practices based on gender, race, color, and personality traits. The study indicates that algorithmic bias stems from limited raw

# BIAS AND HALLUCINATIONS

The Berkeley



Trinity  
College  
Dublin

The University of Dublin

- Very confident and convincing
- May be partly true
- Or may be completely false



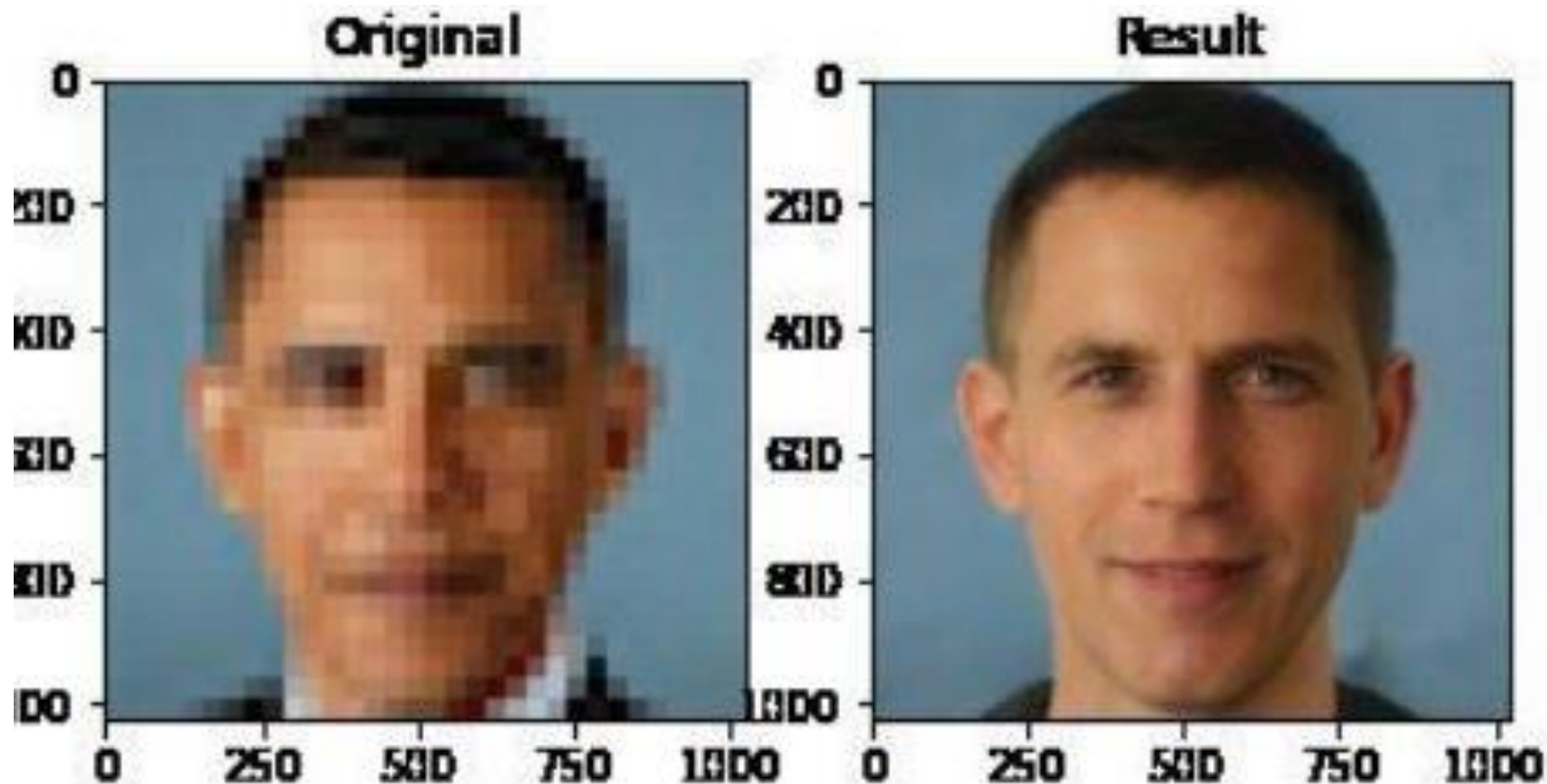
Janelle Shane <https://www.aiweirdness.com/depixelation-or-hallucination-20-06-26/>

# BIAS AND HALLUCINATIONS



Janelle Shane <https://www.aiweirdness.com/depixelation-or-hallucination-20-06-26/>

# Bias and Hallucinations



Janelle Shane <https://www.aiweirdness.com/depixelation-or-hallucination-20-06-26/>



# A USEFUL TOOL? PREDICTING CRIMINALITY



The University of Dublin

## AI experts say research into algorithms that claim to predict criminality must end



/ AI is in danger of revisiting the pseudoscience of physiognomy

**Algorithms trained on racist data produce racist results**

“Let’s be clear: there is no way to develop a system that can predict or identify ‘criminality’ that is not racially biased — because the category of ‘criminality’ itself is racially biased,” write the group.

[James Vincent, AI experts say research into algorithms that claim to predict criminality must end - The Verge](#)

# A USEFUL TOOL? MEDICAL DIAGNOSIS

The Berkeley



Trinity  
College  
Dublin

The University of Dublin

## Worse than Random? An Embarrassingly Simple Probing Evaluation of Large Multimodal Models in Medical VQA

**Qianqi Yan**

University of California, Santa Cruz  
qyan79@ucsc.edu

**Xuehai He**

University of California, Santa Cruz  
xhe89@ucsc.edu

**Xiang Yue**

Carnegie Mellon University  
xyue2@andrew.cmu.edu

**Xin Eric Wang**

University of California, Santa Cruz  
xwang366@ucsc.edu

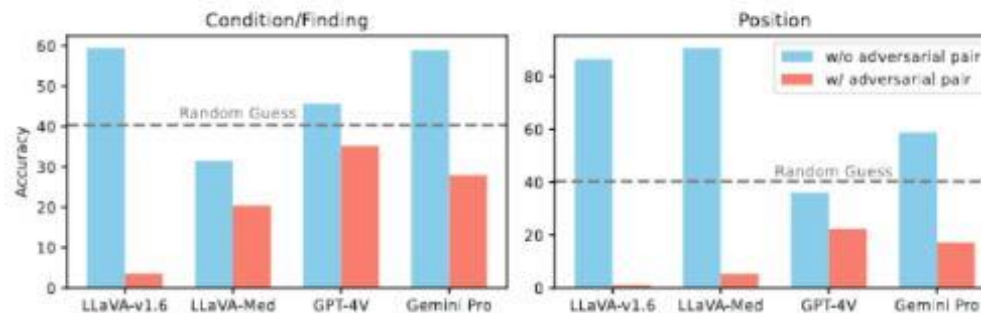


Figure 1: Accuracy of four LMMs on two types of specialized questions in medical diagnoses, with and without adversarial pairs. The significant drop in accuracy with adversarial pairs highlights the models' unreliability in handling medical diagnoses.

Yan et al., 2024

PROOF POINTS

# Kids who use ChatGPT as a study assistant do worse on tests

*Researchers compare math progress of almost 1,000 high school students*

- Students with ChatGPT solved 48% more of the practice problems correctly
- Scored 17 % worse on a test

Basani et al., 2024

## Ireland's datacentres overtake electricity use of all urban homes combined

Training LLMs to use the right data requires the use of massive, expensive server farms that act as supercomputers.

Statistics raise concerns that rise in demand for data processing driven by AI could derail climate targets



[Ireland's datacentres overtake electricity use of all urban homes combined | Ireland | The Guardian](#)





Each time you search for something like “how many rocks should I eat” and Google’s AI “snapshot” tells you “at least one small rock per day,” you’re consuming approximately three watt-hours of electricity, according to Alex de Vries, the founder of Digiconomist, a research company exploring the unintended consequences of digital trends. That’s ten times the power consumption of a traditional Google search, and roughly equivalent to the amount of power used when talking for an hour on a home phone. (Remember those?)



# ENERGY AND EMISSIONS

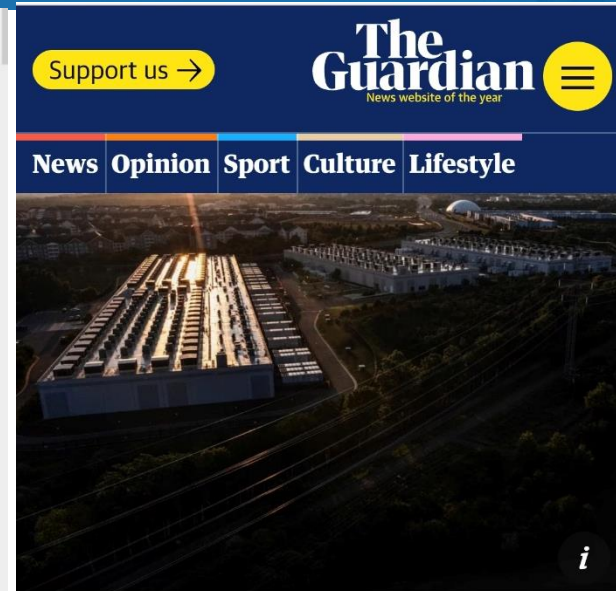


The University of Dublin



[Data](#) [Markets](#) [Technology](#)

## AI computing is on pace to consume more energy than India, Arm predicts



### Technology

## Data center emissions probably 662% higher than big tech claims. Can it keep up the ruse?

Emissions from in-house data centers of Google, Microsoft, Meta and Apple may be 7.62 times higher than official tally

### Google

## Google's emissions climb nearly 50% in five years due to AI energy demand

Bloomberg News

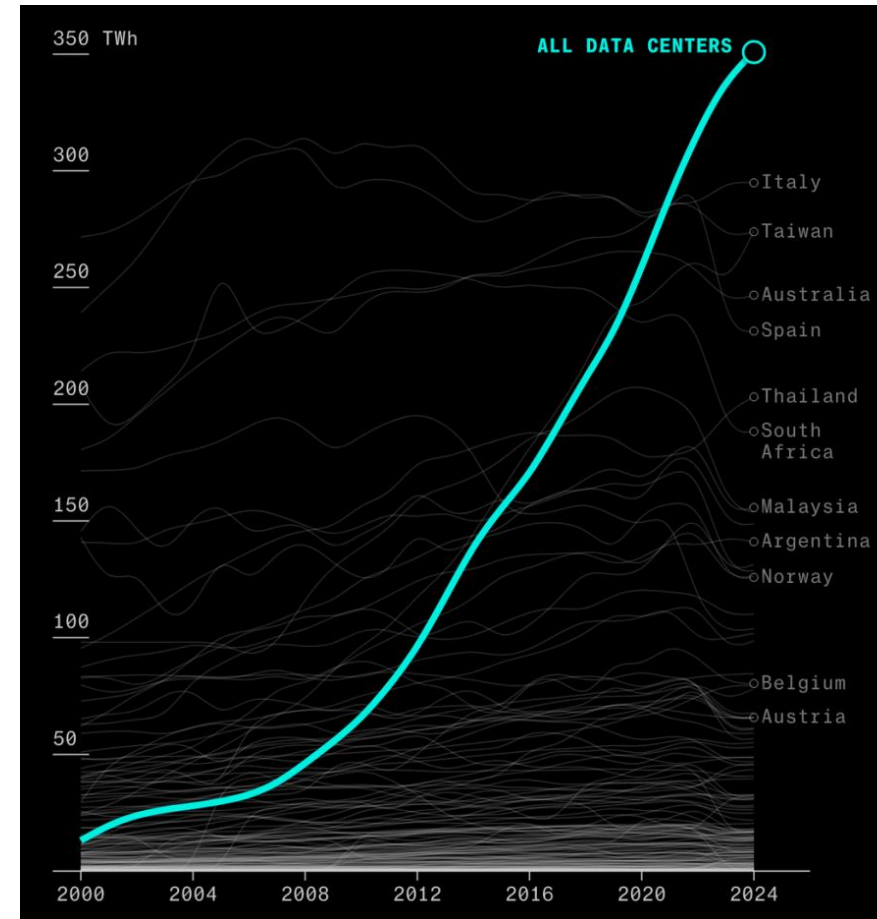
April 30, 2024

<https://www.theguardian.com/technology/article/2024/jul/02/google-ai-emissions>

# ENERGY AND EMISSIONS



- In the UK, AI is expected to suck up 500% more energy over the next decade. “That kind of electricity growth that hasn’t been seen in a generation.”
  - Goldman Sachs, Bloomberg, 21 June 2024



# Top 5 GenAI Corporations



The University of Dublin

---

1. Microsoft Corporation (biggest funder of OpenAI)

---

Market Capitalisation: \$2.442 Trillion

---

2. Alphabet Inc. (Google)

---

\$1.718 Trillion

---

3. Amazon.com, Inc.

---

\$1.425 Trillion

---

4. NVIDIA Corporation

---

\$1.198 Trillion

---

5. Meta Platforms, Inc.

---

\$762.633 Billion

---

[18 Biggest Generative AI Companies In The World \(yahoo.com\)](#)

# Top 5 GenAI Corporations



The University of Dublin

---

1. Microsoft Corporation (biggest funder of OpenAI)

---

Market Capitalisation: \$2.442 Trillion

---

2. Alphabet Inc. (Google)

---

\$1.718 Trillion

---

3. Amazon.com, Inc.

---

\$1.425 Trillion

---

4. NVIDIA Corporation

---

\$1.198 Trillion

---

5. Meta Platforms, Inc.

---

\$762.633 Billion

---

[18 Biggest Generative AI Companies In The World \(yahoo.com\)](#)

## CASE STUDY: MICROSOFT



The University of Dublin

“Microsoft has committed to becoming carbon negative by 2030, achieving zero waste production and positive water usage, and offsetting all greenhouse gas emissions produced since our founding by 2050.”

<https://news.microsoft.com/en-cee/2023/05/18/microsoft-is-committed-to-achieving-zero-carbon-emissions-and-waste-by-2030/>



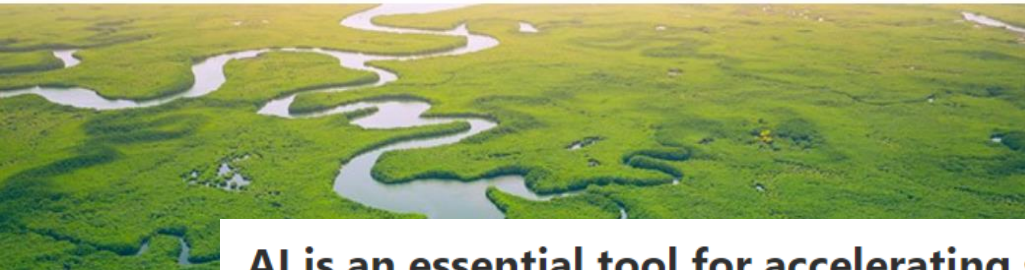
# CASE STUDY: MICROSOFT



The University of Dublin

## Accelerating Sustainability with AI: A Playbook

Nov 16, 2023 | Brad Smith, Vice Chair and President, and Melanie Nakagawa, Chief Sustainability Officer



### **AI is an essential tool for accelerating sustainability**

Given the urgency of the planetary crisis, society needs to push harder on the AI accelerator while establishing guardrails that steer the world safely, securely, and equitably toward net-zero emissions, climate resilience, and a nature-positive future.

[Accelerating Sustainability with AI: A Playbook - Microsoft On the Issues](#)

# AI IS FUELING A RESURGENCE IN OIL AND GAS



The University of Dublin

- Major new coalition between BlackRock Inc. and Microsoft Corp.
  - One of the largest efforts to fund data warehouses and energy infrastructure behind AI
- Partners include:
  - Global Infrastructure Partners, private equity firm
  - Abu Dhabi created an investment vehicle this year for AI
  - United Arab Emirates investment vehicle
  - Nvidia Corp., the and expert in AI data centres

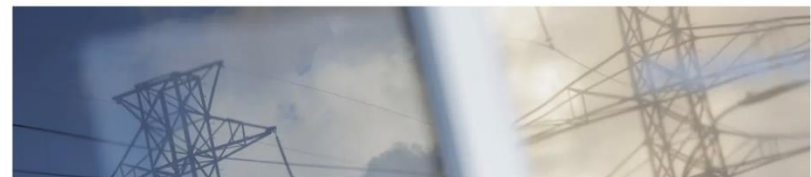
[Microsoft News BlackRock, Global Infrastructure Partners, Microsoft and MGX launch new AI partnership to invest in data centers and supporting power infrastructure - Stories](#)

**The AI Race:** AI Startups to Watch | AI's Real Carbon Footprint

Green | New Energy

## AI Boom Is Driving a Surprise Resurgence of US Gas-Fired Power

New gas plants just keep on coming, defying expectations that their rapid growth was nearing an end.



## CASE STUDY: MICROSOFT



The University of Dublin

To help make her case, Alpine presented a memo to Smith in December 2021 that calculated the effects of the company's oil and gas deals. She pointed, for example, to a single 2019 deal with ExxonMobil that could purportedly "expand production by as much as 50,000 oil-equivalent barrels a day by 2025," according to a Microsoft press release. Those extra barrels would produce an estimated 6.4 million metric tons of emissions, drastically outweighing a carbon-removal pledge that Microsoft made in 2020, she wrote. (I verified her estimate with multiple independent carbon analysts. ExxonMobil declined to comment.)

Karen Hao, The Atlantic, September 13, 2024

## CASE STUDY: MICROSOFT



Trinity  
College  
Dublin

The University of Dublin

production. Over the past year, it has leaned into the generative-AI rush in an effort to clinch more deals—each of which can be worth more than hundreds of millions of dollars. Microsoft employees have noted that the oil and gas industries could represent a market opportunity of \$35 billion to \$75 billion annually, according to documents I viewed.

Karen Hao, The Atlantic, September 13, 2024

# WATER CONSUMPTION



The University of Dublin



THE IRISH TIMES



Subscriber Only Business

## AI boom sparks concern over Big Tech's water consumption

Microsoft, Google and Meta are using more water to cool down data centres that power artificial intelligence products



TECHNOLOGY

## AI IS TAKING WATER FROM THE DESERT

New data centers are springing up every week. Can the Earth sustain them?

By Karen Hao





# SUPERVISED TRAINING AND THE MYTH OF AUTOMATION

- What is supervised training, who does the supervising, and how much do they get paid?
  - Despite the automation mythology, AI is heavily dependant on human input
  - AI business model: users are unpaid or low wage human testers/trainers
  - Trained on work made by human labour without their consent
  - Depends on low-cost and emotionally demanding labour outsourced to the global south



Photo by [Kelly Sikkema](#) on [Unsplash](#)

# SUPERVISED TRAINING AND THE MYTH OF AUTOMATION



The University of Dublin

BUSINESS • TECHNOLOGY

## Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic

15 MINUTE READ



“Making (ChatGPT) safe destroyed my family. It destroyed my mental health. As we speak, I’m still struggling with trauma.”

<https://slate.com/technology/2023/05/openai-chatgpt-training-kenya-traumatic.html>

# GUIDANCE FROM TCD ACADEMIC INTEGRITY



The University of Dublin

Some uses of GenAI are illegal and break GDPR law:

- Using private student or staff information as part of your queries or instructions when using GenAI tools.
- Sharing private information with a third party (the GenAI provider).
  - GenAI tools harvest this information and depend on interactions with users to train their systems.
- Content which is not public or **for which you do not own the copyright** should NOT be used in creating prompts or providing context information for GenAI to generate content.

## Can we use AI ethically, and if so, what would ethical use of AI look like?

- Group 1: Plagiarism/ copyright
- Group 2: Problems with bias
- Group 3: Use as a tool - inaccuracy, racial and gender bias,
- Group 4: Energy usage
- Group 5: Water usage
- Group 6: Human labour - unpaid/low wage



**Trinity College Dublin**

Coláiste na Tríonóide, Baile Átha Cliath

The University of Dublin

THE LIBRARY OF TRINITY COLLEGE DUBLIN, THE  
UNIVERSITY OF DUBLIN  
*LEABHARLANN CHOLÁISTE NA TRÍONÓIDE,  
OLLSCOIL ÁTHA CLIATH*

**Thank You!**

The Berkeley Library



## SOME USES OF GENAI ARE ILLEGAL AND BREAK GDPR LAW



The University of Dublin

- Using private student or staff information as part of your queries or instructions when using GenAI tools.
- Sharing private information with a third party (the GenAI provider).
  - GenAI tools harvest this information and depend on interactions with users to train their systems.
- Content which is not public or **for which you do not own the copyright** should NOT be used in creating prompts or providing context information for GenAI to generate content.

GenAI can generate erroneous content and information generated from GenAI cannot be considered correct all of the time. Therefore, students are expected to fact-check information that they generate from GenAI when using it to learn or explain subjects.

If a student **generates content from a GenAI tool and submits it as their own work, it is considered plagiarism**, which is defined as academic misconduct in accordance with the College Academic Integrity Policy. If a sentence or quotation is being used by a student in their answer, then it must be referenced.

Cases of plagiarism are considered under College Academic Misconduct Procedures.



## SCHOOL POLICIES OR APPROACHES TO THE USE OF GENAI

### EXAMPLE: FACULTY OF SCIENCE, TECHNOLOGY, ENGINEERING & MATHEMATICS

Our (current) policy towards Gen AI could perhaps best be described as follows:

- Use is encouraged within modules under the specific conditions prescribed by the module co-ordinator. All use should be acknowledged according to standard citing conventions
- In the absence of specific conditions, the default will be that use of Gen AI is not permitted



## SCHOOL POLICIES OR APPROACHES TO THE USE OF GENAI

### EXAMPLE: FACULTY OF SCIENCE, TECHNOLOGY, ENGINEERING & MATHEMATICS

Should I use AI in my academic work?

There is no right or wrong answer to this question. Writing your entire assessment using an LLM will most certainly fall under the College Academic Integrity policy. Using machine learning as a statistical method for your research project will most certainly not and might create a very exciting and state of the art project. Beyond that, use your judgment, keeping in mind the caveats above. Some modules will have a specific AI policy, and you should follow that. If in doubt, do not hesitate to ask the module coordinator

## SCHOOL POLICIES OR APPROACHES TO THE USE OF GENAI

### EXAMPLE 1: FACULTY OF ARTS, HUMANITIES & SOCIAL SCIENCES

This policy n  
help you to r  
your notes in  
mind the lim  
more concis  
are students  
such tools to  
skills.

Beyond such  
**own work, v**  
**misconduct.**  
assignment

Generative AI Declaration	
	<p>Please delete the statement <b>that does not apply</b> to this work:</p> <p><b>(A) Nothing to declare. I did not use generative AI software as part of the work for this assessment.</b></p> <p><b>(B) I used generative AI as part of the work for this assessment.</b></p> <p><b>NOTE:</b> Where (B) is selected, students are required to include an appendix containing a &lt;350-word explanation of how generative AI was used (which tool; how the output was generated; how this output was used in the assignment) and how that use conforms with the principles of academic integrity. Questions and/or prompts and the generative AI output should be included as part of the appendix. These are not included in the word count for the assignment. Where these are not included, <b>they may be requested as part of the marking process.</b></p>

ning and to  
ummarise  
(bearing in  
r to make it  
ve AI, nor  
ying on  
academic

as your  
mic  
xam or





## SCHOOL POLICIES OR APPROACHES TO THE USE OF GENAI

### EXAMPLE 2: FACULTY OF ARTS, HUMANITIES & SOCIAL SCIENCES

The School is committed to upholding the highest standards of academic integrity and intellectual rigour. For this reason, **the School will treat the presence of any text which has been generated by artificial intelligence (AI) tools in any work submitted for assessment as a breach of College's guidelines on Academic Integrity.** Any use of AI-generated material in an essay or exam answer will render that work inadmissible for assessment and will be subject to the sanctions outlined in the College Calendar