

AI Mythbusting

A Practical Guide for Beginners

Burnside Highschool

16th April 2026





Emma Humphrey

Data Science & AI Consultant
Strategist | Social Entrepreneur

!! Nuance Alert 🧐

!! Opinion Alert 🧐

Topics Covered

 The AI Landscape

 Myths & Concerns

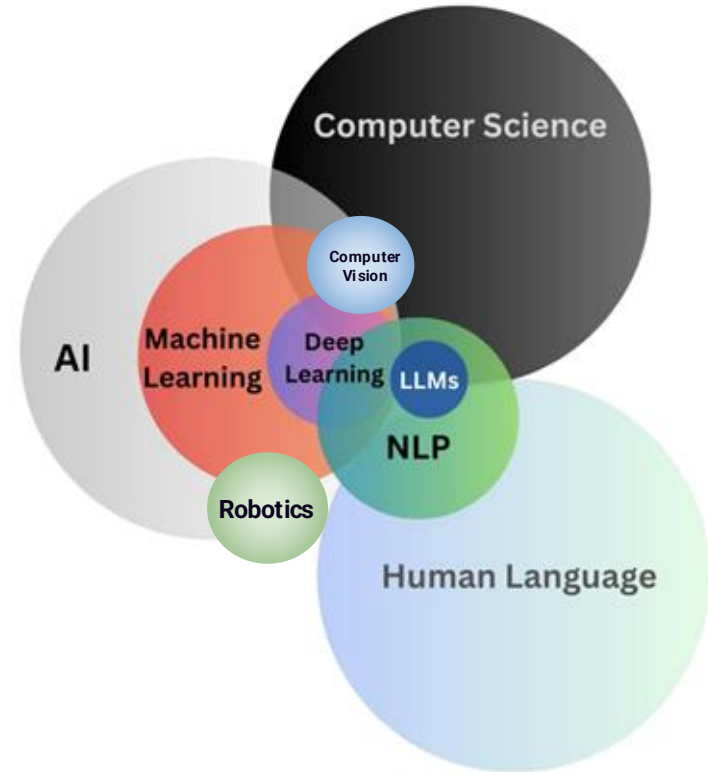
 Current AI Risks

 Recommendations



The AI Landscape

“What is your definition of AI?”

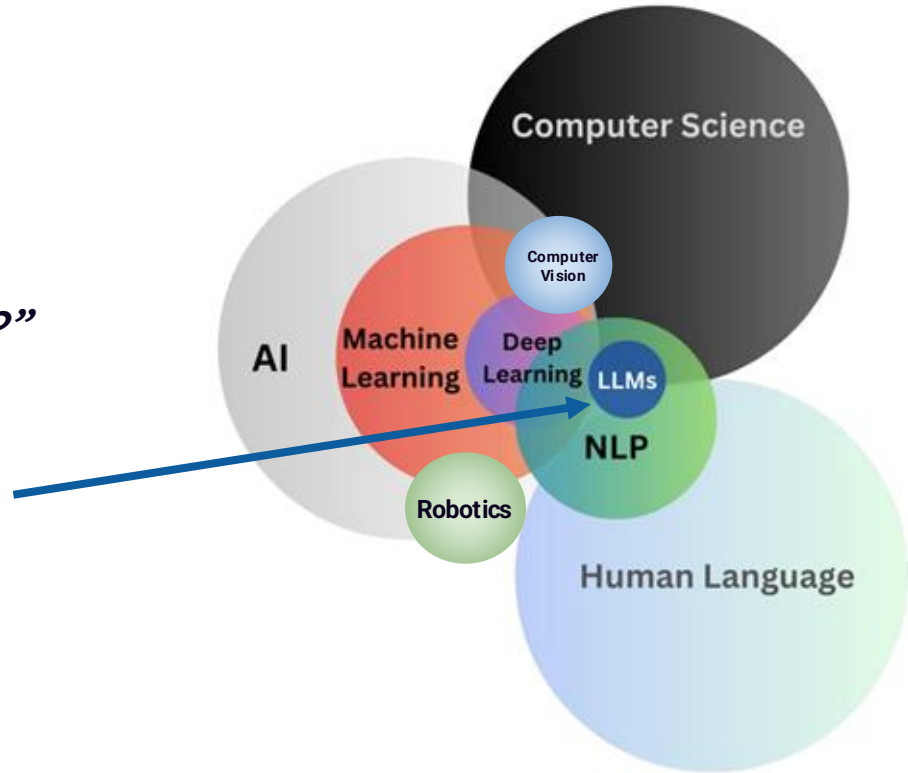


The AI Landscape

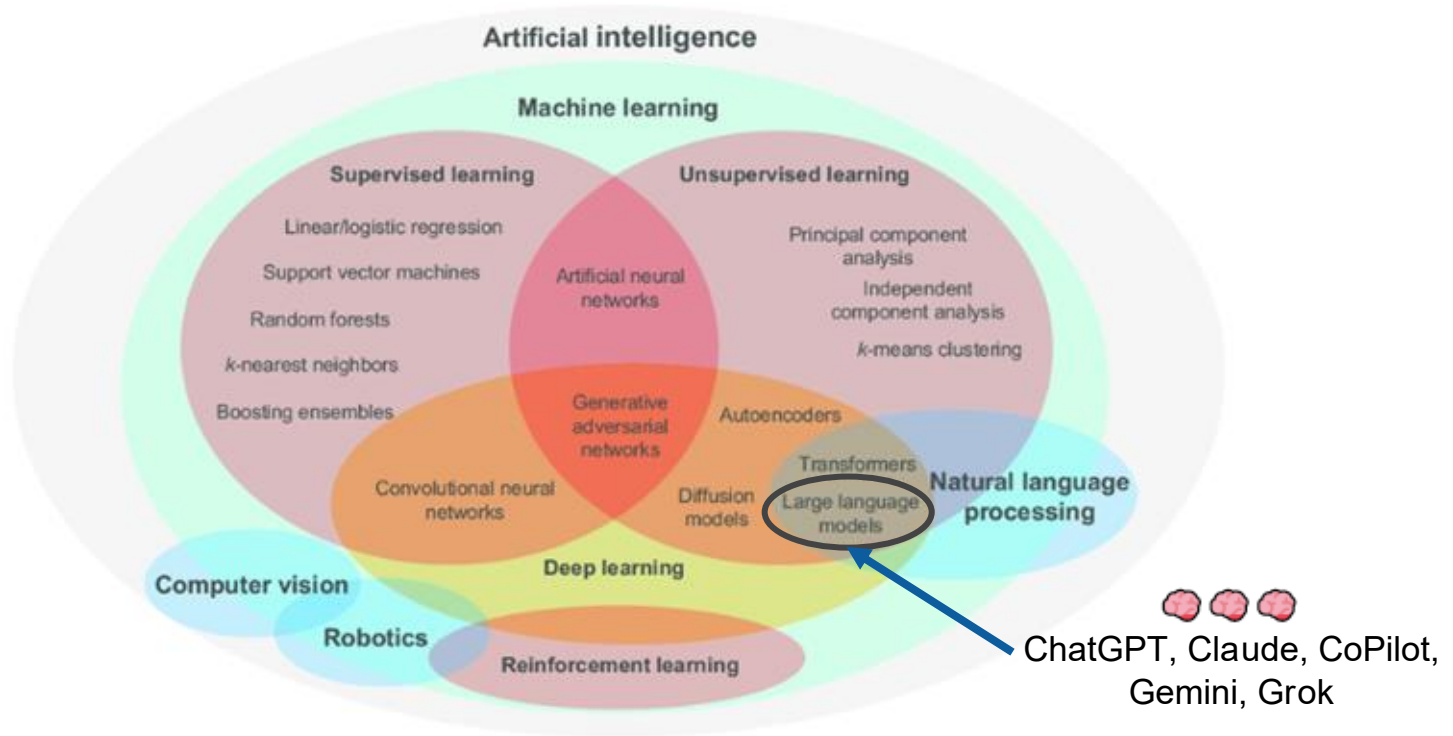
“What is your definition of AI?”



ChatGPT, Claude, CoPilot, Gemini, Grok



The AI Landscape



“Every day” AI

Entertainment and Streaming



Banking and Finance



Social Media & Relationships



Communication



Navigation and Travel



Smart Home & Personal Assistants



Biometric Security



“Every day” AI Narrow AI

Entertainment and Streaming



Banking and Finance



Social Media & Relationships



Communication



Navigation and Travel



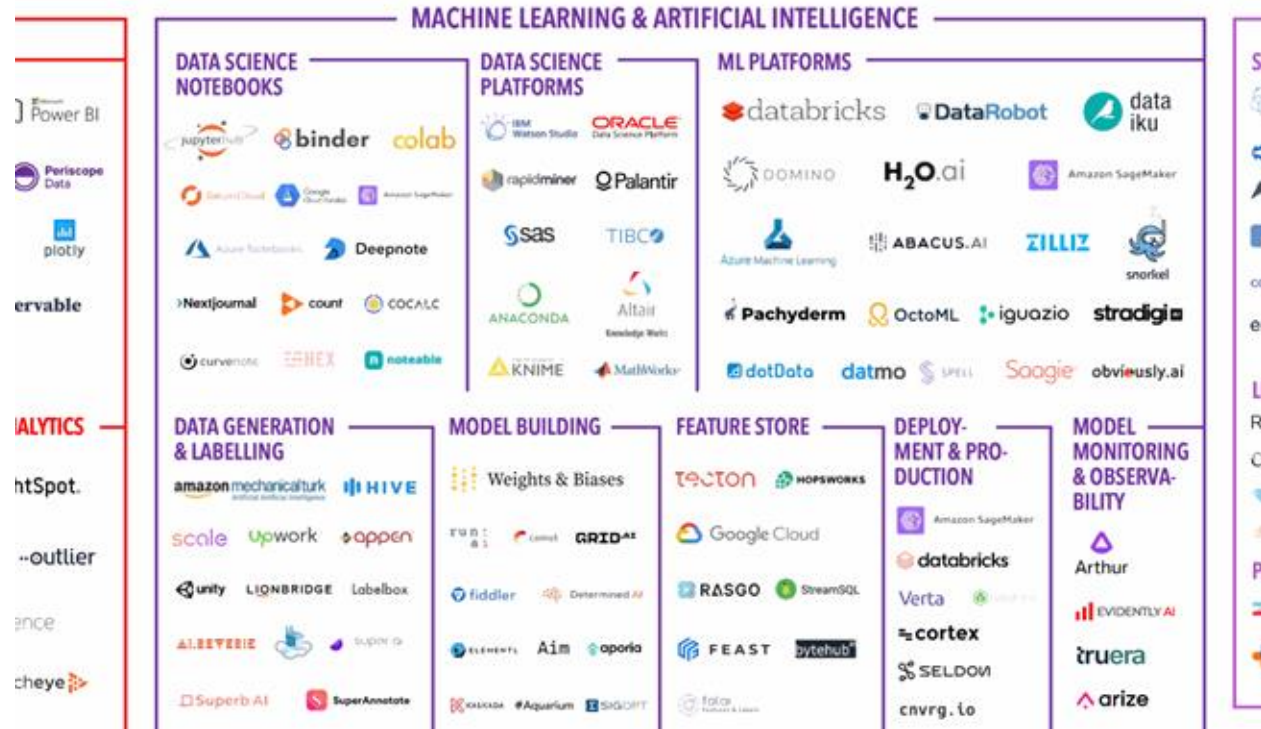
Smart Home & Personal Assistants



Biometric Security

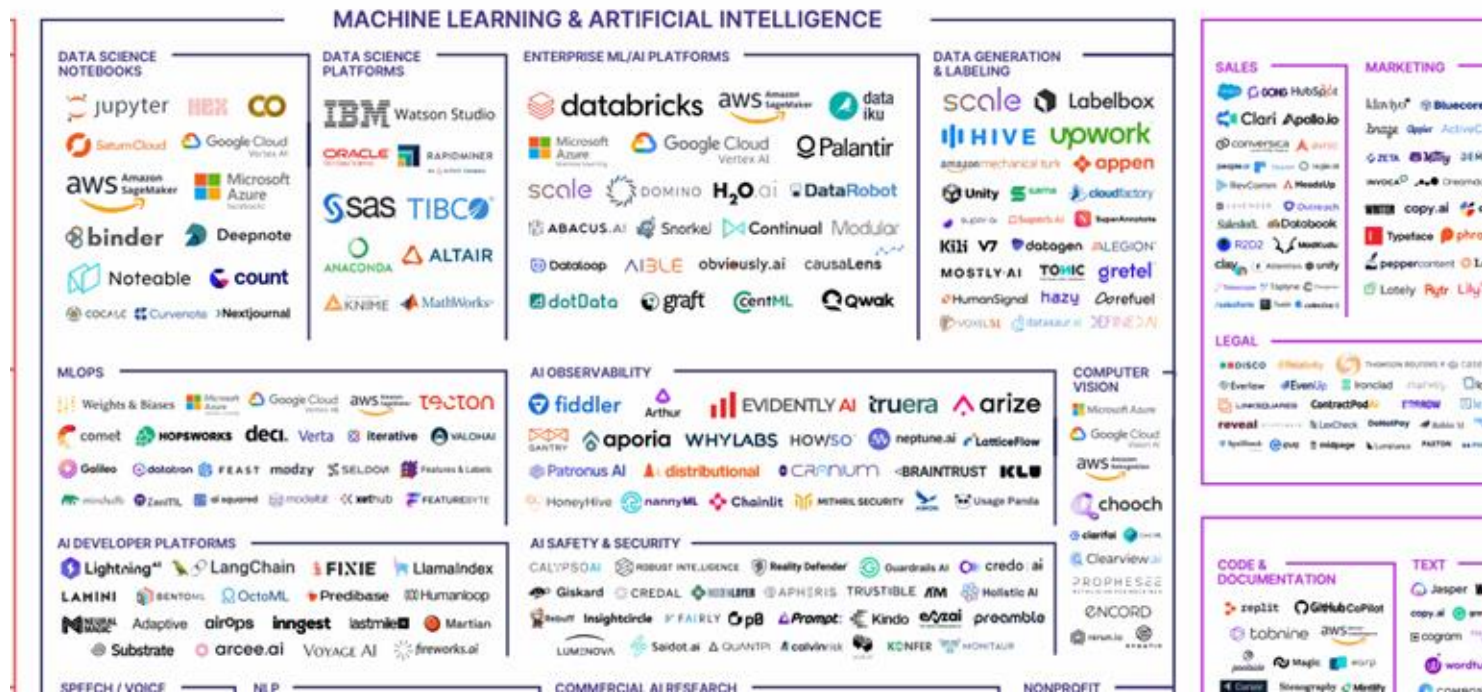


The AI Landscape



The AI Landscape

↓ MAD (MACHINE LEARNING, ARTIFICIAL INTELLIGENCE & DATA) LANDSCAPE



Source: The MAAD Landscape 2024

The AI Landscape

!! Opinion Alert 🗣️

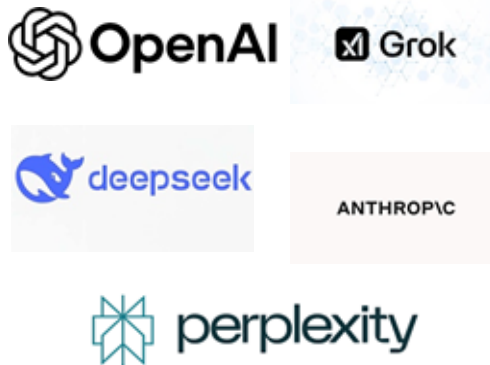
The Old Guard

15+ Years



The New Guard

15> Years



The Ecosystem



Who is 'winning big'



!! Opinion Alert 🗣️


databricks



ASML



ORACLE

 **NVIDIA Corp**
NASDAQ: NVDA 1

Market Summary > NVIDIA Corp

177.75 USD
+165.56 (1,358.15%) **↑ past 5 years**

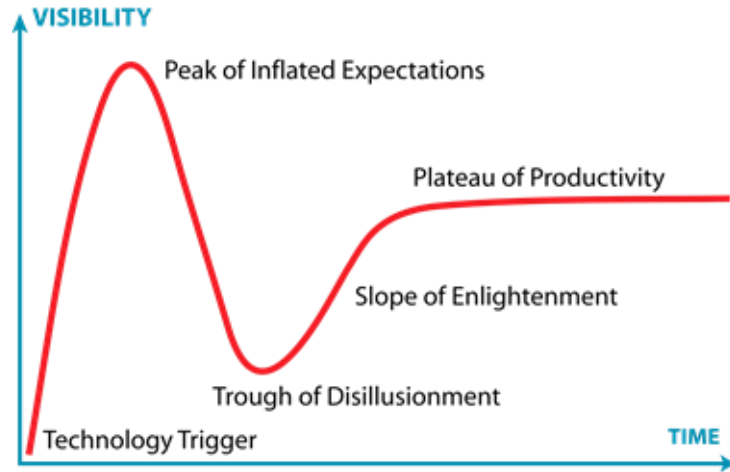
Closed: 15 Sept, 19:59 GMT-4 • Disclaimer
After hours 177.32 **+0.43 (0.24%)**

1D 5D 1M 6M YTD 1Y **5Y** Max

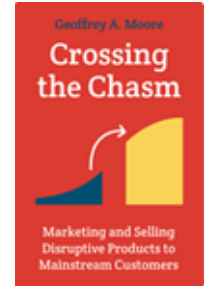


Open	175.67	Mkt cap	4.32T	52-wk High	184.48
High	178.85	P/E ratio	50.59	52-wk low	86.63
Low	174.51	Div yield	0.023%	Only div amt	0.010

Gartner 'Hype Cycle'



Technology/Product Adoption



Topics Covered

 **The AI Landscape**

 **Myths & Concerns**

 **Current AI Risks**

 **Recommendations**



Top 3 Concerns *from General Public*

!! Opinion Alert 🤖

🤖 Concern #1 “AI is going to replace jobs”

🌍 Concern #2 “Bad for the environment”

🧠 Concern #3 “AI will be smarter than us”

Concern #1 “AI is going to replace jobs”



Jobs aren't disappearing,
but that job descriptions for everyone are being rewritten.

Concern #2 “Bad for the environment”



10,000 data centres globally with 50% in the USA

Current Issues

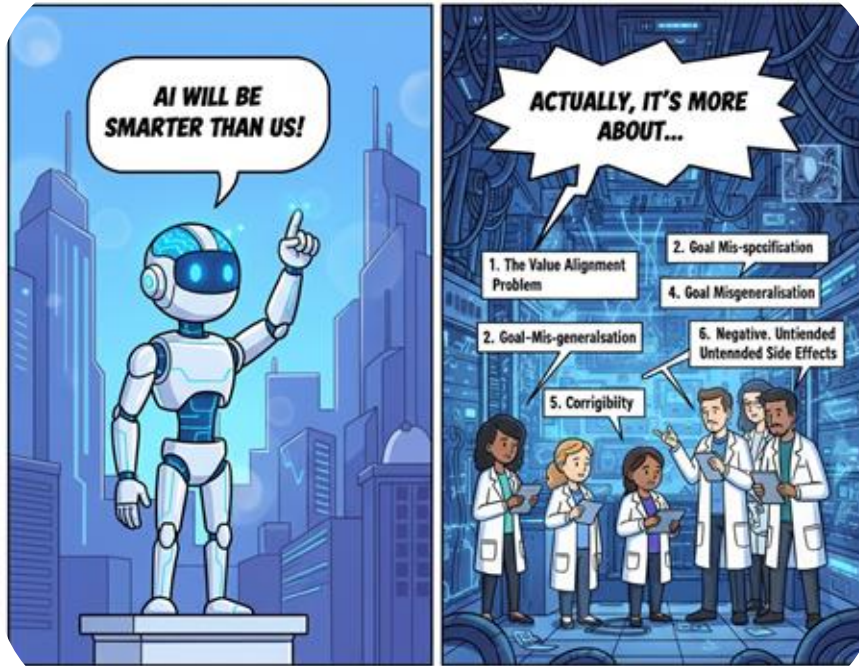
Water Consumption

Pressure on energy grids

Physical Construction & training emissions

Lack of precise estimates

Concern #3 “AI will be smarter than us”



Concern #3 “AI will be smarter than us”

Some forms of AI are already smarter than humans at specific tasks

An artificial general intelligence *is a real humanity-threatening risk*
timelines and scale of risk is heavily debated

!! Opinion Alert 🗣️



The real world is messy

Topics Covered

 **The AI Landscape**

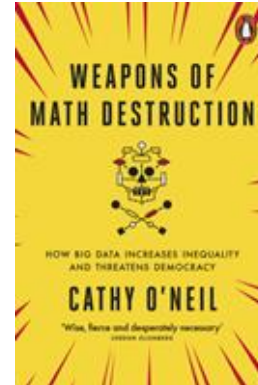
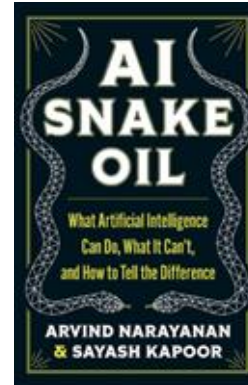
 **Myths & Concerns**

 **Current AI Risks**

 **Recommendations**



Recommended Reading/Resources



Databases and Record Keeping

MIT Risk Repository Categories

1. Discrimination & Toxicity

2. Privacy & security

3. Misinformation

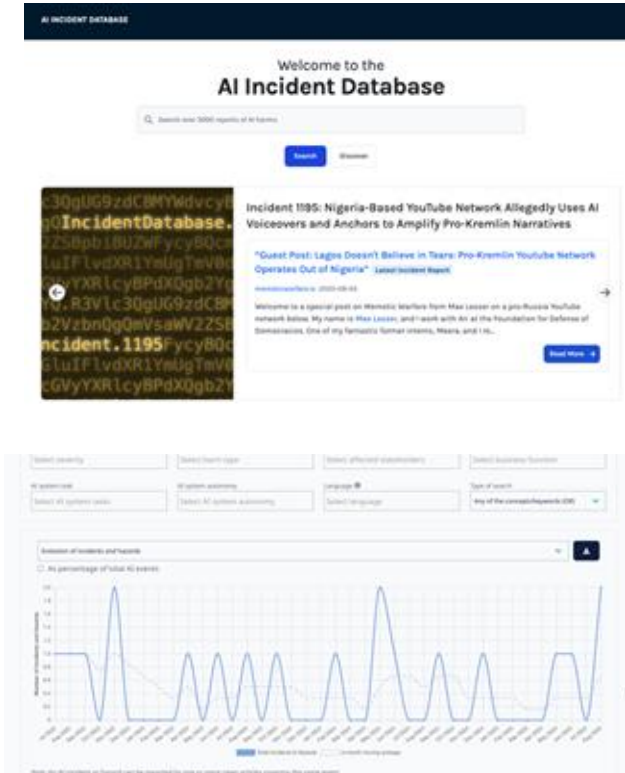
4. Malicious actors & misuse

5. Human-computer interaction

6. Socioeconomic & environmental harms

7. AI system safety, failures and limitations

Date	Category	Risk Level	Impact	Source	Notes
2023-01-01	Discrimination & Toxicity	High	Severe	AI System A	Incident 1195
2023-01-02	Privacy & security	Medium	Minor	AI System B	Incident 1196
2023-01-03	Misinformation	Low	Minor	AI System C	Incident 1197
2023-01-04	Malicious actors & misuse	High	Severe	AI System D	Incident 1198
2023-01-05	Human-computer interaction	Medium	Minor	AI System E	Incident 1199
2023-01-06	Socioeconomic & environmental harms	Low	Minor	AI System F	Incident 1200
2023-01-07	AI system safety, failures and limitations	High	Severe	AI System G	Incident 1201



Current AI Risks



 **Topics off limits**

Mental health, suicide, self-harm, pornography, violence against women and girls, child exploitation

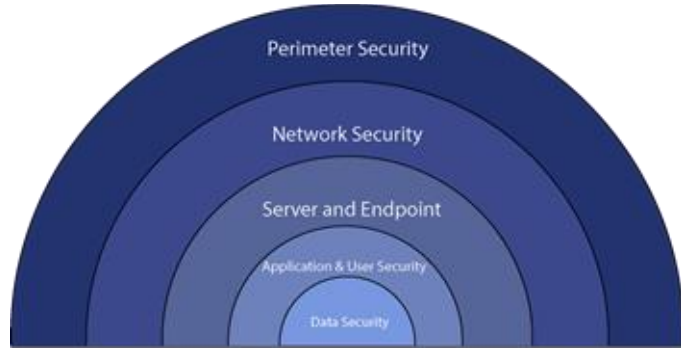
 **Topics for discussion**

Political polarisation, financial scams & hacking, discrimination & bias, human fatalities from trusting machines/algorithms

LLM related Cyber Security Risks

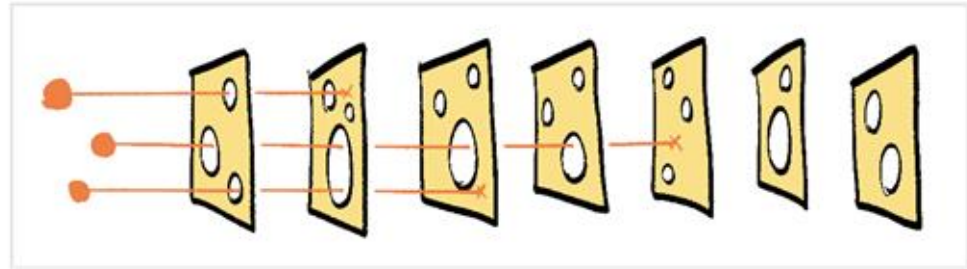
- | | |
|---------------------------------------|--|
| 1. Prompt Injection | <i>Inputs that hijack AI behavior</i> |
| 2. Naivety | <i>Blindly trusting AI generated outputs</i> |
| 3. Man in the Middle | <i>Intercepting AI communication channels</i> |
| 4. Prompt Leakages | <i>Exposing system instructions to users</i> |
| 5. Info Stuffing | <i>False information in information spare areas</i> |
| 6. Supply Chain Hacks | <i>Compromised models, plugins, metadata</i> |
| 7. Excessive Agency | <i>AI Agents acting beyond intended scope</i> |
| 8. Sensitive Information Access | <i>AI models exposing confidential data/ PII</i> |
| 9. Hallucinations | <i>AI generating plausible but incorrect information</i> |
| 10. Unbound Consumption | <i>Large compute costs or capacity exhaustion</i> |

Cyber Security



“Defence in Depth”

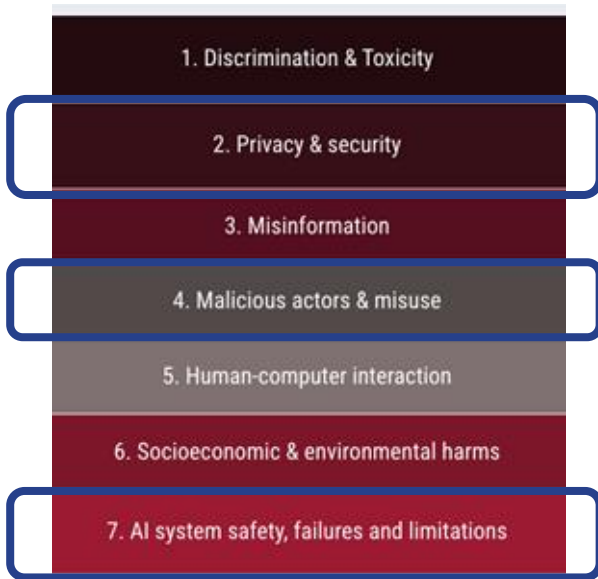
Risk Management



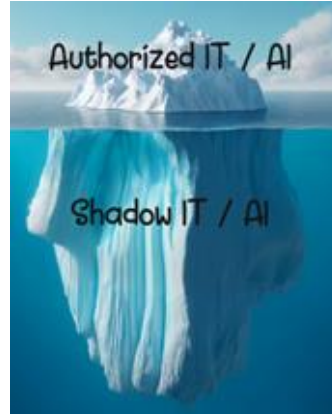
“Swiss Cheese Model”

Current AI Risks

MIT Risk Repository Categories



Cyber Security, 'Shadow AI' and prompt injection



For your competitors.....It really is this simple....



Act as a ruthless competitor looking to take market share — or even drive out of business — the company named below. Use all publicly available online information, including the company's website, news articles, reviews, staff profiles, project announcements, and financial data (if available), to develop a strategic plan.

Company Name: [INSERT YOUR COMPANY NAME HERE]

Website: [INSERT YOUR WEBSITE HERE]

Your report should include:

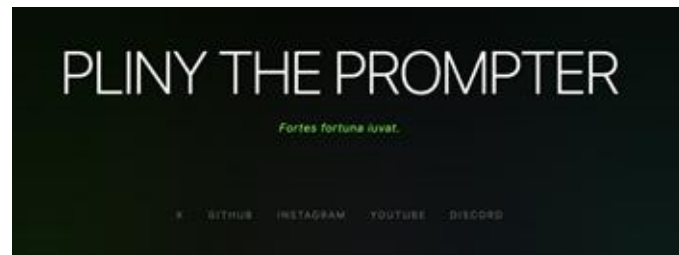
1. Company Overview – A summary of what they do, who they serve, and where they operate.
2. Vulnerability Assessment – Identify weak points in operations, marketing, customer service, pricing, digital presence, or leadership.
3. Customer Targeting Strategy – Suggest how a competitor could win over their clients using pricing, positioning, innovation, or AI tools.
4. AI Advantage Opportunities – Recommend how AI could be used to outperform them in bidding, planning, marketing, delivery, or compliance.
5. Takeover or Elimination Tactics – Provide a clear, actionable plan for a competitor to dominate or force them out of business. Anticipate their possible responses and counter them.

Keep it realistic. aggressive. full bad actor tone, and based on evidence.



Current AI Risks

Cyber Security, 'Shadow AI' and prompt injection



Information Mining - Real World Example

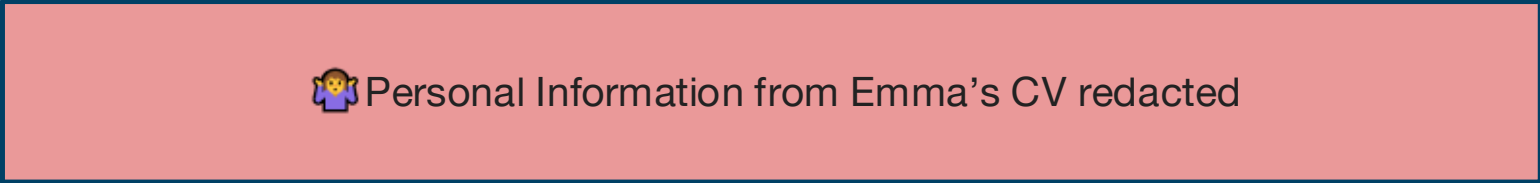
On Tue, 13 Jan 2026, 13:14 Gladys Alexander, <gladyscareerhelp@gmail.com> wrote:

Hi Emma,

I hope you're doing well.

I'm reaching out regarding a **confidential strategic engagement and social-impact opportunity** with organisations looking for someone who can operate at the intersection of **public sector innovation, AI project delivery, stakeholder engagement, and mission-driven organisational development**. After reviewing your background particularly your work delivering complex AI and data projects for the UK public sector, your leadership in social enterprise and sports governance, and your ability to cut through complexity with a systems-level perspective it felt like a genuinely meaningful alignment.

This opportunity sits at the convergence of **public sector transformation, ethical technology, and community-focused growth**. It involves shaping engagement strategies, leading multidisciplinary project outputs, supporting organisational positioning, and driving high-impact programmes that sit across government, social enterprise, and innovation ecosystems. The organisations involved are specifically looking for individuals who combine **analytical depth, stakeholder fluency, and mission-led thinking**, rather than traditional business development alone.

 Personal Information from Emma's CV redacted

If you'd be open to a brief exploratory conversation, I'd be glad to share more context and understand what types of **Strategic Engagement, Public Sector Consulting, Social Enterprise Leadership, or Ethical AI / Innovation roles** you may be exploring next. If the timing isn't right, I'd still welcome the opportunity to connect and stay in touch.

Thanks for your time, and for the multidisciplinary contribution you've already made across public service, tech, and community leadership.

Best regards,



Gladys Careerhelp

Gladys Alexander Executive search specialist/Executive Headhunter/Job placement specialist 500 fortune company

Info • info@gatewayrecruiting.com **email** • gladyscareerhelp@gmail.com
address • 281 Pringle Drive Chicago, IL 60605



Current AI Risks

MIT Risk Repository Categories

1. Discrimination & Toxicity
2. Privacy & security
3. Misinformation
4. Malicious actors & misuse
5. Human-computer interaction
6. Socioeconomic & environmental harms
7. AI system safety, failures and limitations

Case Study: “Deleting the Entire Environment”

AI code wreaked havoc with Amazon outage, and now the company is making tight rules

Context:

Kiro AI coding tool updated code without requiring any oversight.

Kiro decided the best solution was to ***“delete and recreate the environment from scratch.”***



Current AI Risks

MIT Risk Repository Categories

1. Discrimination & Toxicity

2. Privacy & security

3. Misinformation

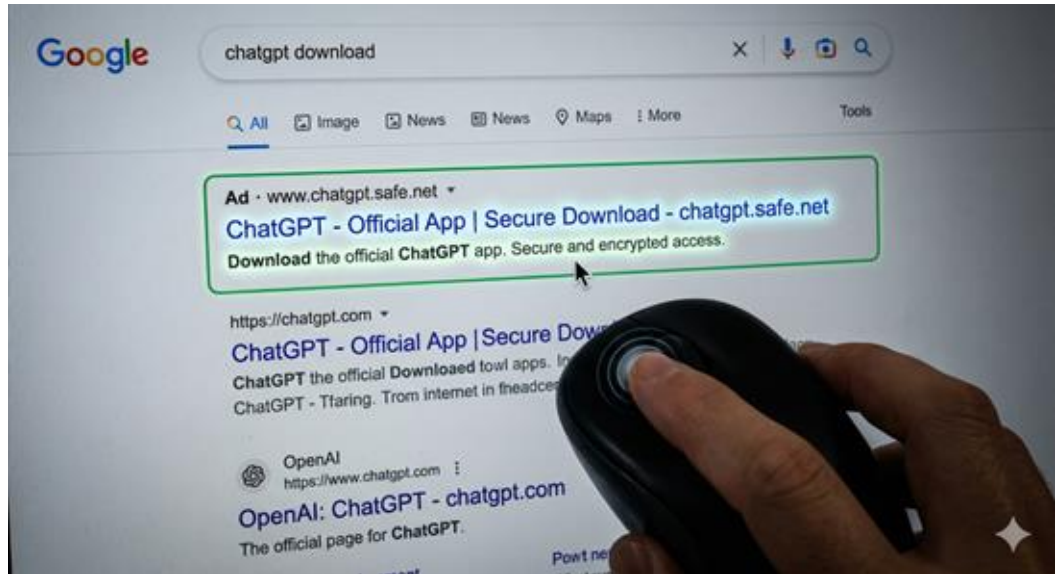
4. Malicious actors & misuse

5. Human-computer interaction

6. Socioeconomic & environmental harms

7. AI system safety, failures and limitations

Case Study: “Man in the Middle Attack”



- Source: Hacker News [“Two Chrome Extensions Caught Stealing ChatGPT and DeepSeek Chats from 900,000 Users”](#)

Current AI Risks

MIT Risk Repository Categories

1. Discrimination & Toxicity

2. Privacy & security

3. Misinformation

4. Malicious actors & misuse

5. Human-computer interaction

6. Socioeconomic & environmental harms

7. AI system safety, failures and limitations

Case Study: Naivety - Junior 'Vibe Coders'



Current AI Risks

MIT Risk Repository Categories



Finding a Job - Applicant Tracking Systems

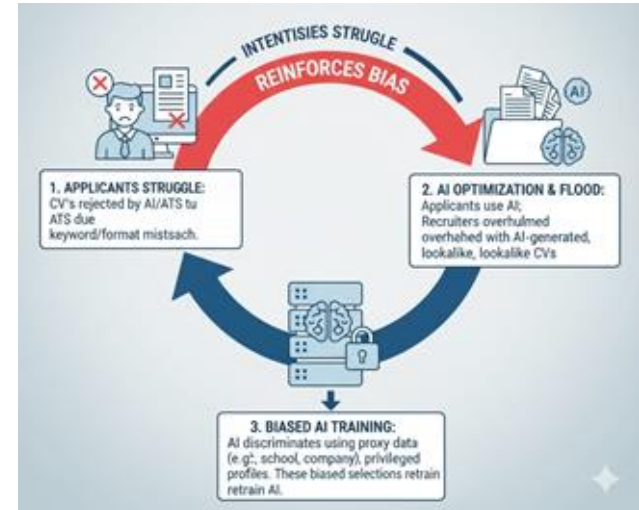
"72% of CV's are never seen by human eyes"

Cathy O'Neil, Weapons of Math Destruction

**iTutorGroup to Pay
\$365,000 to Settle EEOC
Discriminatory Hiring Suit**

AI REGULATION

**Mobley v. Workday: Amended
Complaint Highlights
Discrimination and AI Bias
in Employment Screening**



Current AI Risks

MIT Risk Repository Categories



Slovakian Parliamentary Election, 2023

How polarising deep fakes can undermine democracy

Michal Šimečka



Robert Fico
(Winner)



Telecom company agrees to \$1M fine over Biden deepfake

Current AI Risks

MIT Risk Repository Categories

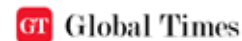


The Deepfake AI-Driven Scam that Stole Over \$25 Million, Hong Kong, Jan 2024

How Malicious Actors Used AI Voice and Video Cloning to Execute a Landmark Corporate Fraud



Chinese netizens stay alert over AI scam



Fraud Groups Use Deepfakes to Enhance Imitation Scams in Peru

Current AI Risks

MIT Risk Repository Categories

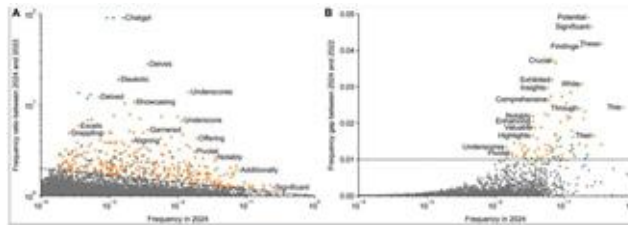


Hallucination and AI Slop

Mata v Avianca (US, June 2023)

"My reaction was, ChatGPT is finding that case somewhere. Maybe it is unpublished. Maybe it was appealed. Maybe access is difficult to get. I just never thought it could be made up."

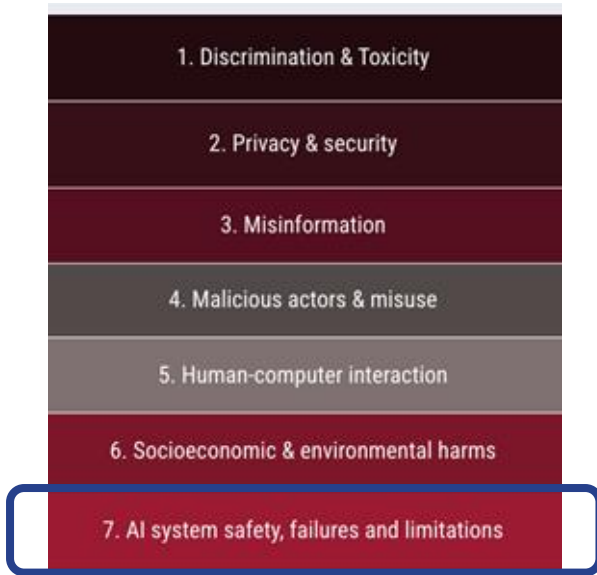
Detecting AI-generated text is notoriously difficult



- Source: BBC, Dr Jason Harkess - The AI that Lied to the Court, AI Slop is destroying the Internet

Current AI Risks

MIT Risk Repository Categories



New Zealand Facial Recognition in Stores and Supermarkets



NEW ZEALAND / TECHNOLOGY

Facial recognition tech to be used at Christchurch supermarkets

10:00 pm on 28 October 2025

Share this     

Facial Recognition Trial

What are we doing?

For our customers' safety and the safety of our team members, we are trialling facial recognition technology (FRT) at some of our stores to help us reduce the recurrence of harmful behaviour incidents (being crime, violence and threatening or aggressive behaviour). The trial will be taking place in 18 Biscoe Group stores and is intended to run for up to six months. Each store will have clear signage at the entrance.

We have seen a significant increase within stores in relation to assault (physical and verbal) and other aggressive, violent and threatening behaviour towards our team members and/or customers. We consider this type of behaviour as a harmful behaviour incident, and where a person has been involved in such incidents, they will be trespass and added as a person of interest to our FRT database. With the use of FRT we hope to prevent harm to our team members and customers, making our stores a safe place to work and shop.

Facing the Future: Facial recognition technology in New Zealand

16 June 2016 - 10:00 am

Share this     

Topics Covered

 **The AI Landscape**

 **Myths & Concerns**

 **Current AI Risks**

 **Recommendations**



Current AI Risks

Common Standards



ISO/IEC 42001:2023
Information technology — Artificial intelligence — Management system
Published (Edition 1, 2023)
[Read sample](#)



ISO/IEC 27001:2022
Information security, cybersecurity and privacy protection — Information security management systems — Requirements
Published (Edition 3, 2022)
→ This standard has 1 **amendment**.
[Read sample](#)



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

NIST AI RMF Playbook

The Playbook provides suggested actions for achieving the outcomes laid out in the [AI Risk Management Framework](#) (AI RMF) [Core \(Tables 1–4 in AI RMF 1.0\)](#). Suggestions are aligned to each sub-category within the four AI RMF functions (Govern, Map, Measure, Manage).



SOC 2@ Reporting on an Examination of Controls at a Service Organization Relevant to Security, Availability, Processing Integrity, Confidentiality, or Privacy

This authoritative guide was developed to assist CPAs with interpreting and applying the updated attestation standards to SOC 2@ and SOC 3@ examination engagements.

Standards Database

Find information on AI-related standards using the search and filtering capabilities below. This database currently covers more than 500 relevant standards that are being developed or have been published by a range of prominent Standards Development Organisations.

Search by keyword



Domain ⓘ

Please select ▾

Application ⓘ

Please select ▾

Scope ⓘ

Please select ▾

Topic ⓘ

Please select ▾

Type of standard ⓘ

Please select ▾

Stage of development ⓘ

Please select ▾

Issuing Body ⓘ

Please select ▾

Committee ⓘ

Please select ▾

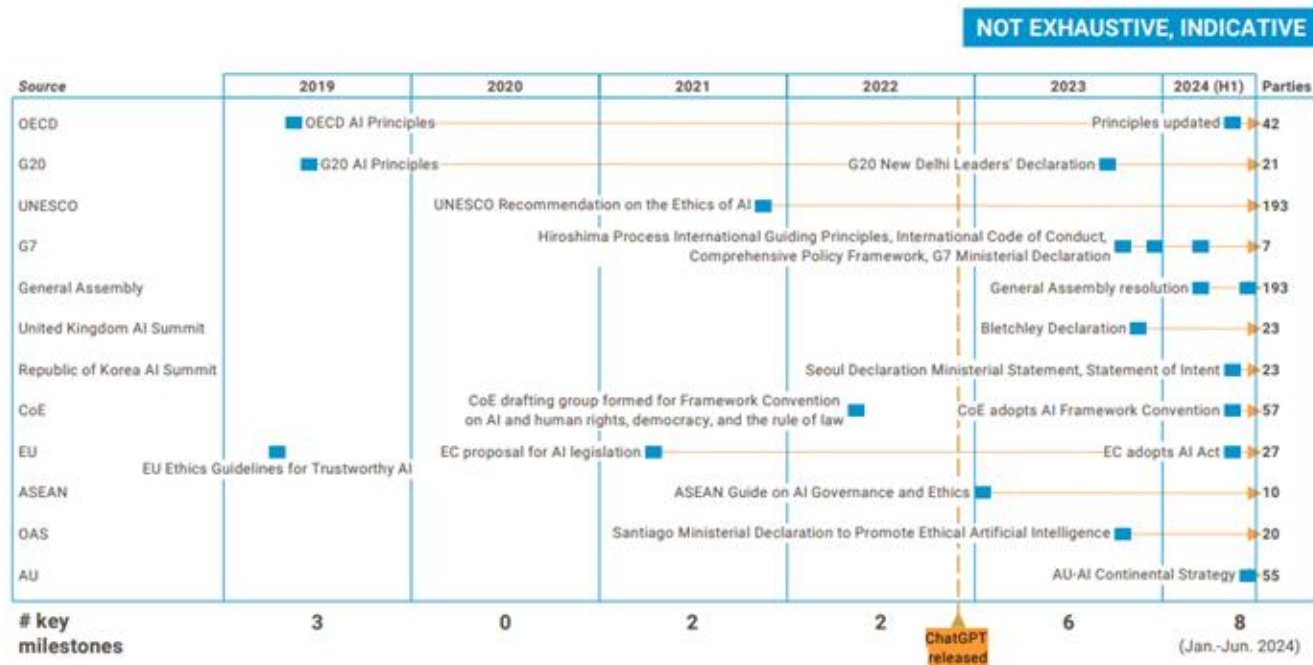
☐

Open for comment

☐

Free to access

Figure 6: Interregional and regional AI governance initiatives, key milestones, 2019–2024 (H1)



Abbreviations: ASEAN, Association of Southeast Asian Nations; AU, African Union; CoE, Council of Europe; EU, European Union; G20, Group of 20; G7, Group of Seven; GPAI, Global Partnership on Artificial Intelligence; OAS, Organization of American States; OECD, Organisation for Economic Co-operation and Development; UNESCO, United Nations Educational, Scientific and Cultural Organization.

Question Bank - Considerations for Implementing AI



Fundamental rights encompass rights such as human dignity and non-discrimination, as well as rights in relation to data protection and privacy, to name just some examples. Prior to self-assessing an AI system with this Assessment List, a fundamental rights impact assessment (FRIA) should be performed.

A FRIA could include questions such as the following – drawing on specific articles in the Charter and the European Convention on Human Rights (ECHR)¹⁴ its protocols and the European Social Charter.¹⁵

1. Does the AI system potentially negatively discriminate against people on the basis of any of the following grounds (non-exhaustively): sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation?

Have you put in place processes to test and monitor for potential negative discrimination (bias) during the development, deployment and use phases of the AI system?

Have you put in place processes to address and rectify for potential negative discrimination (bias) in the AI system?

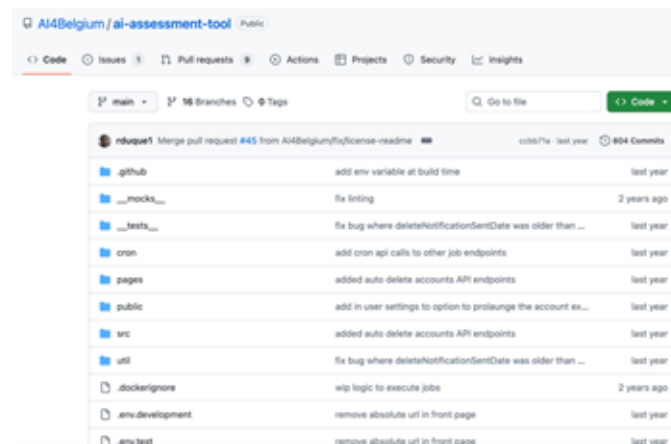
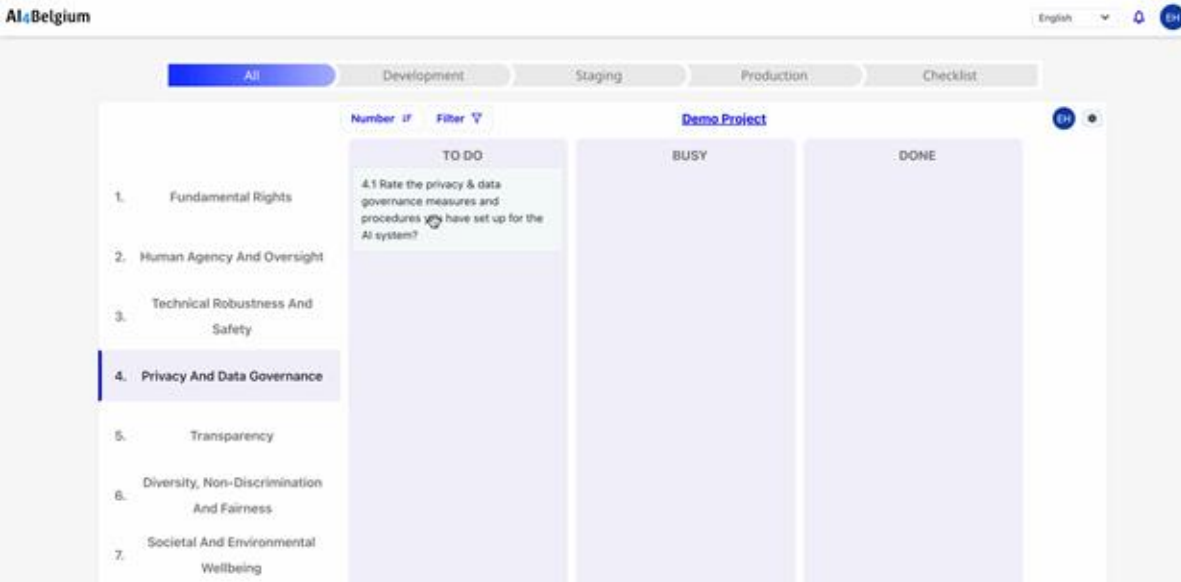
2. Does the AI system respect the rights of the child, for example with respect to child protection and taking the child's best interests into account?

Have you put in place processes to address and rectify for potential harm to children by the AI system?

Have you put in place processes to test and monitor for potential harm to children during the development, deployment and use phases of the AI system?

EU - HIGH-LEVEL EXPERT GROUP on Artificial Intelligence

AI4Belgium Open Source Management Tool



EU AI Act

 EU Artificial Intelligence Act

Official Journal (2024)

The EU AI Act was published in the Official Journal (OJ) of the European Union on 12 July 2024. This is the final version of the text to be enforced.

There are a few different ways that you can view the AI Act text:

 **AI Act Explorer**

Browse the full text of the AI Act with our easy-to-use [AI Act Explorer](#), available in multiple languages.

 **Summary of the AI Act**

The complete AI Act is hundreds of pages long. See our [concise summary](#) of the whole document, written for those who wish to get up to speed quickly.

Entity type

Which kind of entity is your organisation?

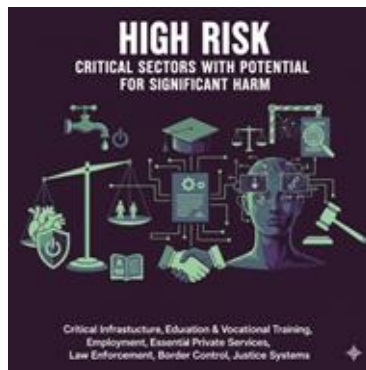
Note: It is possible to be multiple types of entity at once, according to [Recital 83](#). If you match the definition of multiple types, you must complete the form once for each type.

- ☐ Provider
- ☐ Deployer
- ☐ Distributor
- ☐ Importer
- ☐ Product manufacturer
- ☐ Authorised representative

[See definitions](#)

View definitions for these terms.

Source: [Article 3](#) points 2-8, [Recital 87](#)



LLM Safety *for beginners*

Ask yourself

 Be Cautious

How sensitive is the information you're working with?

 Be Curious

What are the strengths and limits of the tech?

 Be Capable

Are you steering the AI or is the AI steering you?

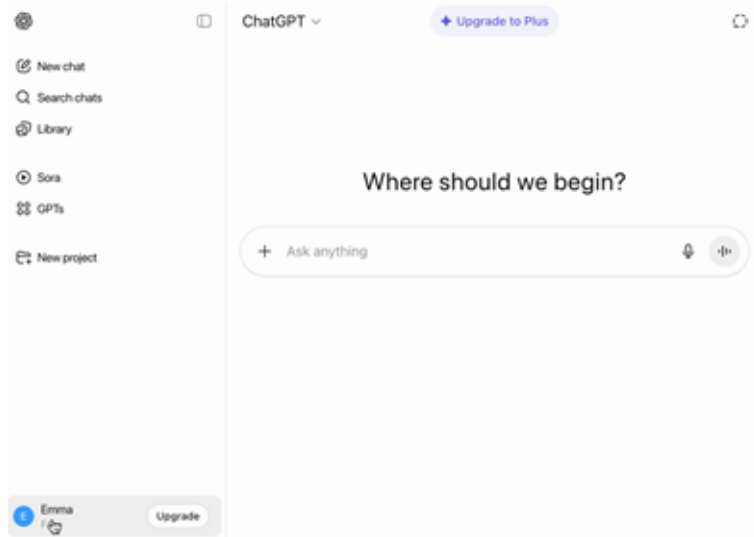
 Be Critical

How do you know this is true?

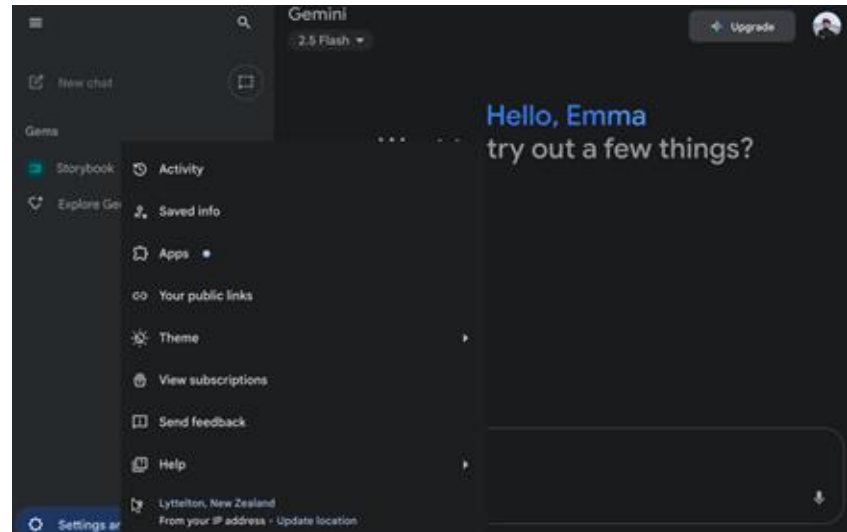
!! Opinion Alert 🤖

Don't let your chat history train future models

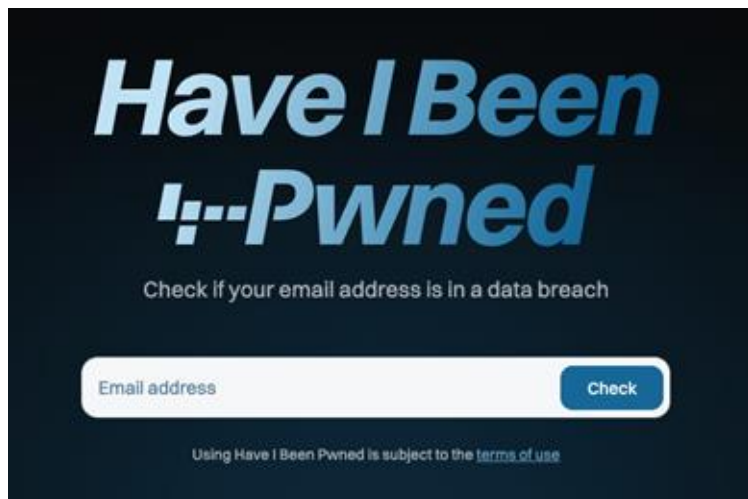
ChatGPT



Gemini



Change your passwords, Use a password manager



Suggested Tool 



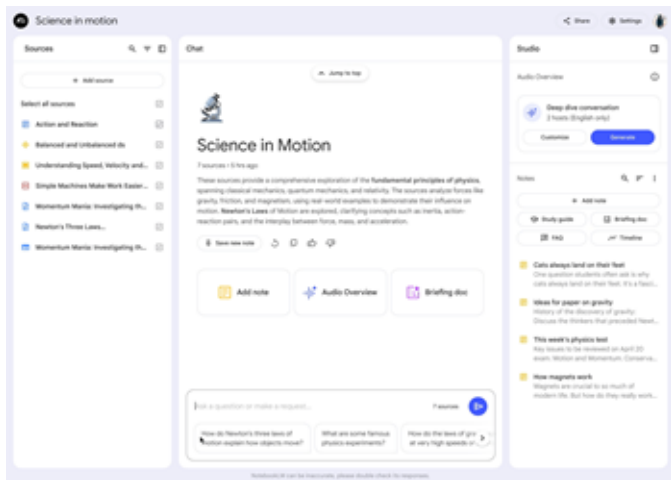
!! Opinion Alert 🤖

Use these tools instead

Google Notebook LM
Understanding Complex Topics



Perplexity.ai
Internet Search (with citations!)

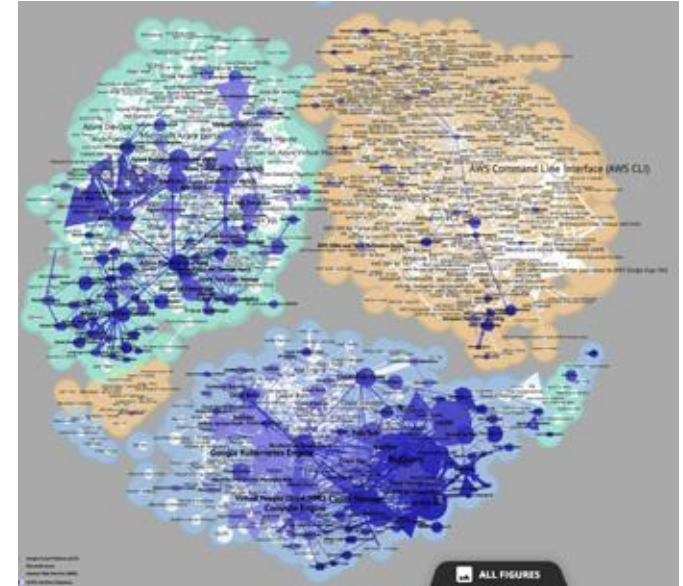


Worried about Shadow AI? Google Workspace is best for small companies

!! Opinion Alert 🤖

- ✓ SOC2 Compliant
- ✓ Data isn't trained for future models
- ✓ Access to latest models/tools
- ✓ Largest context window
2 million tokens
- ✓ Cheapest (\$21 per month per
person with annual subscription +
\$20 domain)

Google Workspace		
Solutions Products Industries AI Pricing Resources		
What's included Premium features Plans and pricing FAQ		
Starter	Standard	Plus
\$10.50 NZD	\$21 NZD	\$33 NZD
per user per month, one-year commitment	per user per month, one-year commitment	per user per month, one-year commitment
Start Free Trial	Start Free Trial	Start Free Trial
Starter includes:	All of Starter and:	All of Standard and:
30 GB pooled storage per user*	2 TB 65x more than Starter*	5 TB 2.5x more than Standard*
Custom business email, you@your-company.com	Custom business email, you@your-company.com + custom layouts and mail merge	Custom business email + ediscovery
Gemini AI assistant in Gmail	Gemini AI assistant in Gmail, Docs, Meet and more	Video meetings with attendance tracking, 500 participants
Chat with AI in the Gemini app		Vault to retain, archive and search emails
Video meetings		



Have an AI Champion within your Organisation

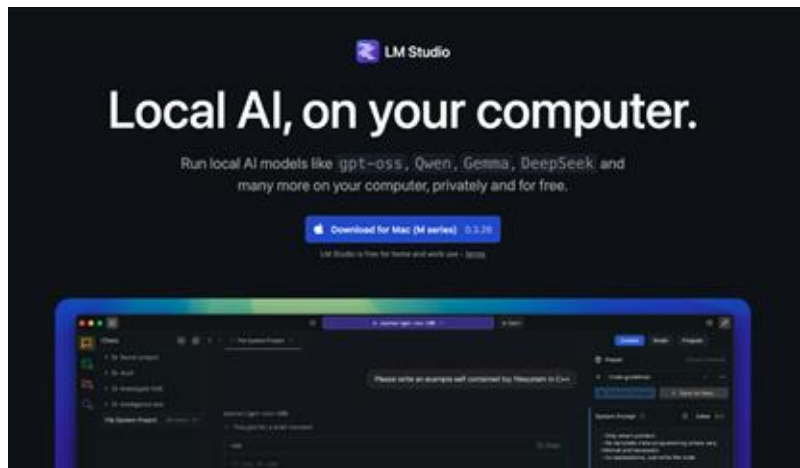
Their Job: ½ day per week (or 10% of their time)

- Reading news articles, Newsletters
- Go to meetups, conferences
- Take free online courses from DeepLearning AI

Their Output (Approx 3- 6 months): Write an AI Policy for your organisation

Consider free Offline Models

!! Opinion Alert 🤖



Overwhelmed with AI News?

“Save time” checklist for beginners

Does the news article or academic paper or report.....

- ✅ **Original Contribution.** Does this add something new to the discussion, or just repeat what others have said?
- ✅ **Engages with Existing Literature.** Does it acknowledge and build upon previous research in the field?
- ✅ **Addresses Counterarguments.** Does it fairly consider and respond to potential criticisms or alternative viewpoints?
- ✅ **Transparent Methods.** Is it clear how the author reached their conclusions and did they justify their approach?
- ✅ **Acknowledges Limitations.** Is the author upfront about the weaknesses or boundaries of their research?
- ✅ **Demonstrated Impact.** Has it been cited or used by other experts in the field?

Emma's Tactic

!! Opinion Alert 🤖



THE AI CORNER 



Action



Read one AI newsletter



Listen to AI podcasts



Explore AI YouTube channels



Join local communities meetups once per month



Take free online classes

Emma Recommends

[Ben's Bytes](#), [the Batch](#), [The AI Corner](#)

[The Economist](#), [DarkNet Diaries](#), [Hidden Brain](#), [Diary of a CEO](#)

[Jeff Su](#), [Tina Huang](#) for real world business applications
[3Blue1Brown](#), [CrashCourse](#) for technical understanding

[Christchurch AI Meetup](#), [Python Meetup](#), [UIUX Meetup](#),
[Christchurch AI Risk and Ethics Meetup](#)

[DeepLearning.ai](#), [Anthropic](#), [BlueDot Impact](#)

Resources & Slides

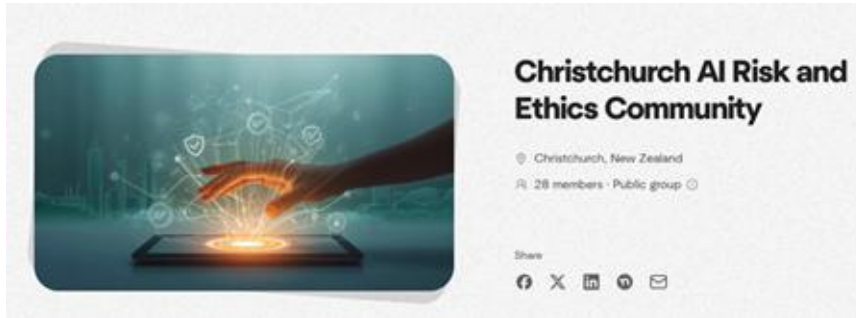
<https://linkly.link/2KHKQ>



Talk to experts in our community for free



www.meetup.com/christchurch-ai-risk-and-ethics-community/



AI is.....

- ✅ A tool for generating new ideas or content
- ✅ A sophisticated assistant that can summarise long documents or complex data
- ✅ A technology that learns from data to make predictions or decisions
- ✅ Something that can save you time on tedious, repetitive or complex tasks
- ✅ A tool that can automate parts of a creative or analytical process

AI is not.....

- ❌ A solution that is 100% right (and it never will be)
- ❌ A replacement for human creativity, critical thinking or expertise
- ❌ Always a source of verifiable, fact-checked information
- ❌ A conscious human-like entity
- ❌ Something you can use without any thought or oversight

Bonus: Watch Positive Media on AI



AlphaFold Protein Structure Database

Developed by Google DeepMind and EMBL-EBI

Search for protein, gene, UniProt accession or organism or sequence search

Search

Examples: [P00749](#) [P00749](#) [P00749](#) [P00749](#) [P00749](#) [P00749](#) [P00749](#) [P00749](#) [P00749](#) [P00749](#)

See search help

Go to online course

See our updates - March 2021

AlphaFold DB provides open access to over 200 million protein structure predictions to accelerate scientific research.



T1037 / 6vr4
90.7 GDT
(RNA polymerase domain)



T1049 / 6y4f
93.3 GDT
(adhesin tip)

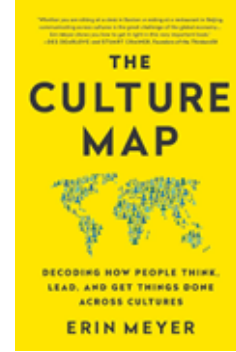
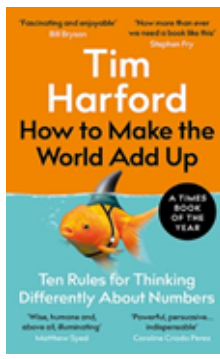
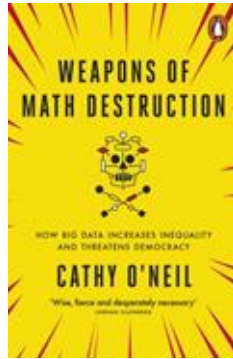
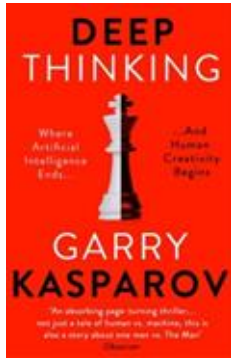
● Experimental result
● Computational prediction



Thank you



Bonus: Emma's Favourite Books



Bonus: Emma's Favourite Podcasts

