

Which AI, for whom, and under what conditions?

Brian Castellani, PhD. FAcSS

Professor of Sociology

Director, Durham Research Methods Centre

Co-Director, Wolfson Research Institute for Health & Wellbeing

Visting Professor, Nelson Mandela University

Adjunct Professor of Psychiatry Northeastern Ohio Medical University



My talk is grounded in two key concerns

1. What if *'AI'* is becoming one of the least useful categories in contemporary discourse about these newly emerging technologies?

- We have spent the last few years debating whether AI is intelligent, creative, transformative, biased, dangerous, or beneficial.
- Yet I increasingly wonder whether we are debating the wrong thing.
- What if 'AI' is becoming one of the least useful categories in contemporary discourse?

- We routinely group together LLMs, machine learning systems, coding assistants, generative image tools, recommendation algorithms, and decision-support systems, then speak about them as if they share a common logic and common consequences.
- My concern is that the category itself may now be obscuring more than it reveals.
- If that is true, then many of our debates about AI—whether optimistic or pessimistic—may be built upon a category that is too broad, too vague, and too analytically weak to support the claims we are asking it to carry.

2. Technological determinism

Manuel Castells and the Danger of Technological Determinism

Technologies do not determine social outcomes on their own. Their effects emerge through their interaction with institutions, cultures, political systems, economic incentives, social practices, and human agency.

Many contemporary AI debates assume a direct relationship between technological capability and social change. Castells reminds us that the same technology can produce very different outcomes depending on how it is adopted, governed, interpreted, resisted, and integrated into society.

Key Question

Which AI, for whom, and under what conditions?

***Which AI,
for whom
and under what conditions?***

Which AI?

Large Language Models (LLMs)

Examples include ChatGPT, Claude, Gemini, and Llama.

These systems are designed to work with language.

They can answer questions, generate text, summarise information, assist with coding, support brainstorming, and engage in conversation.

Much of the current public discussion about 'AI' is probably a discussion about LLMs, despite often being presented as a discussion about AI in general.

But even LLMs are all not the same . . .

Large Language Models (LLMs)

Frontier Models

Frontier models such as ChatGPT, Claude, and Gemini are large general-purpose LLMs trained on enormous datasets using vast computational resources.

They are designed to perform a wide range of tasks, including writing, coding, research, reasoning, and problem-solving.

Because of their scale, resources, and continuous development, they often outperform most open-source models across a broad range of benchmarks and everyday tasks.

Large Language Models (LLMs)

Specialist Models

Specialist models are trained or fine-tuned for particular domains such as medicine, law, finance, science, or software development.

Rather than trying to do everything, they are designed to perform specific tasks or operate within specific knowledge domains, often providing greater accuracy, consistency, or domain expertise than a general-purpose model.

Large Language Models (LLMs)

Open vs Closed Models

Open models, such as Llama, make their underlying model weights available for others to use, modify, and deploy.

Closed models, such as ChatGPT, Claude, and Gemini, remain proprietary and are developed, controlled, and distributed by private organisations.

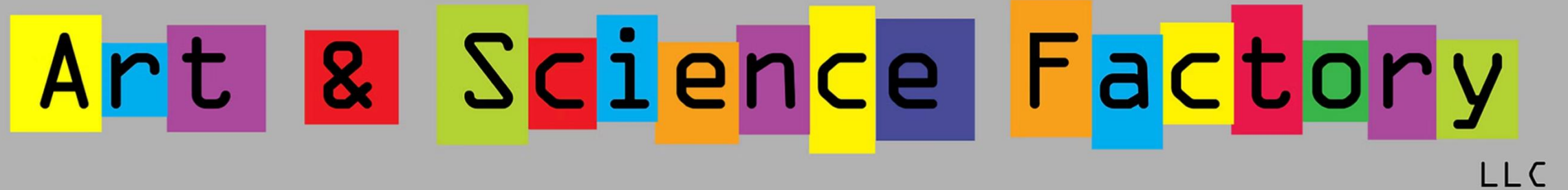
Generative Image Systems

Examples include Midjourney, DALL-E, Sora, Veo, and Runway.

These systems generate images from prompts and can also be used to edit, extend, and transform existing images and video.

They are increasingly being used in design, advertising, architecture, education, and artistic practice.

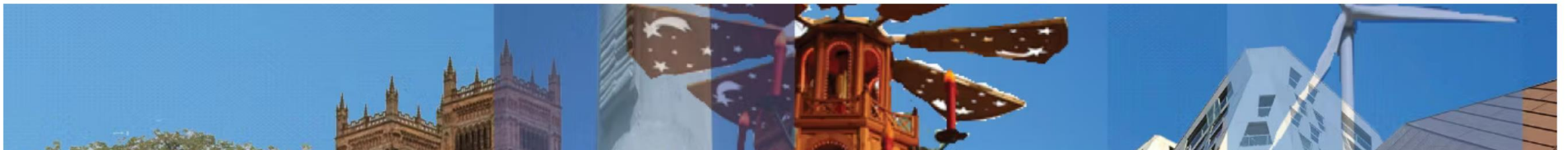
Questions about creativity, authorship, originality, and intellectual property often emerge most strongly in relation to these systems.

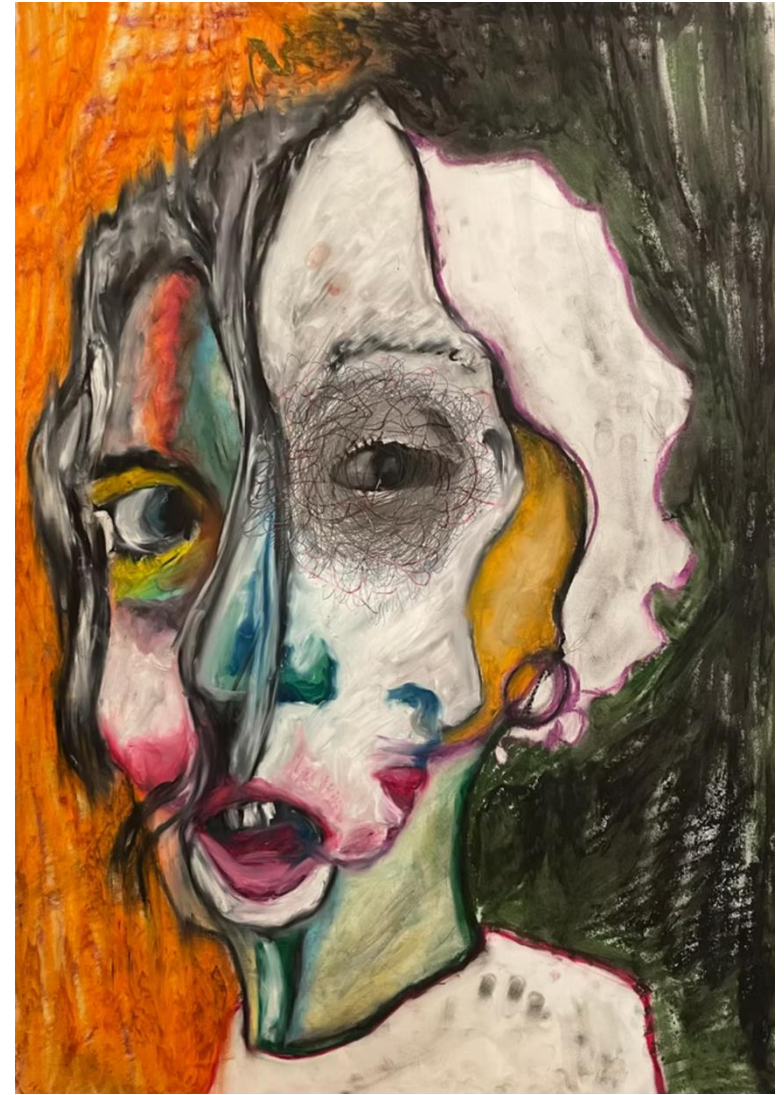


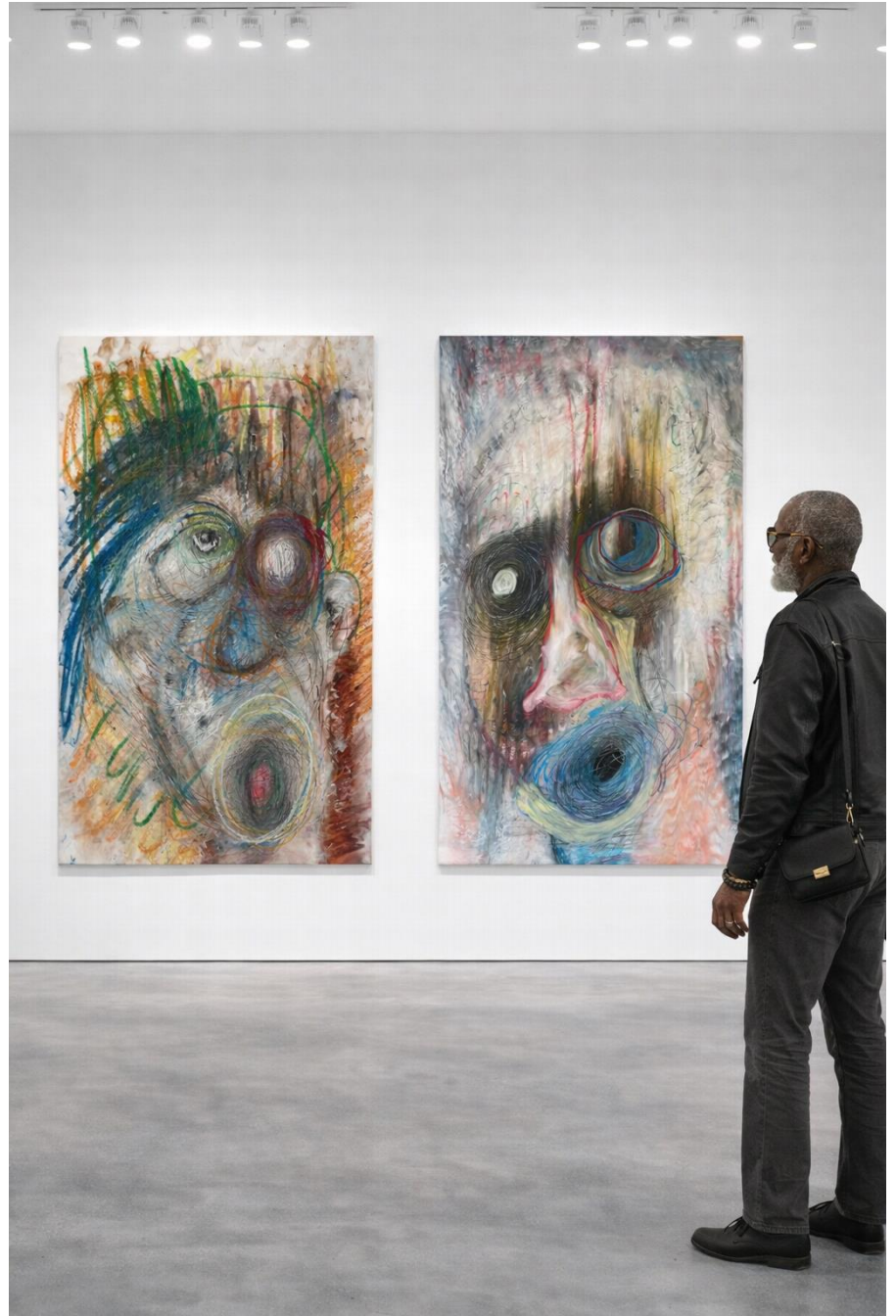
Art Website

Where complexity art and science meet

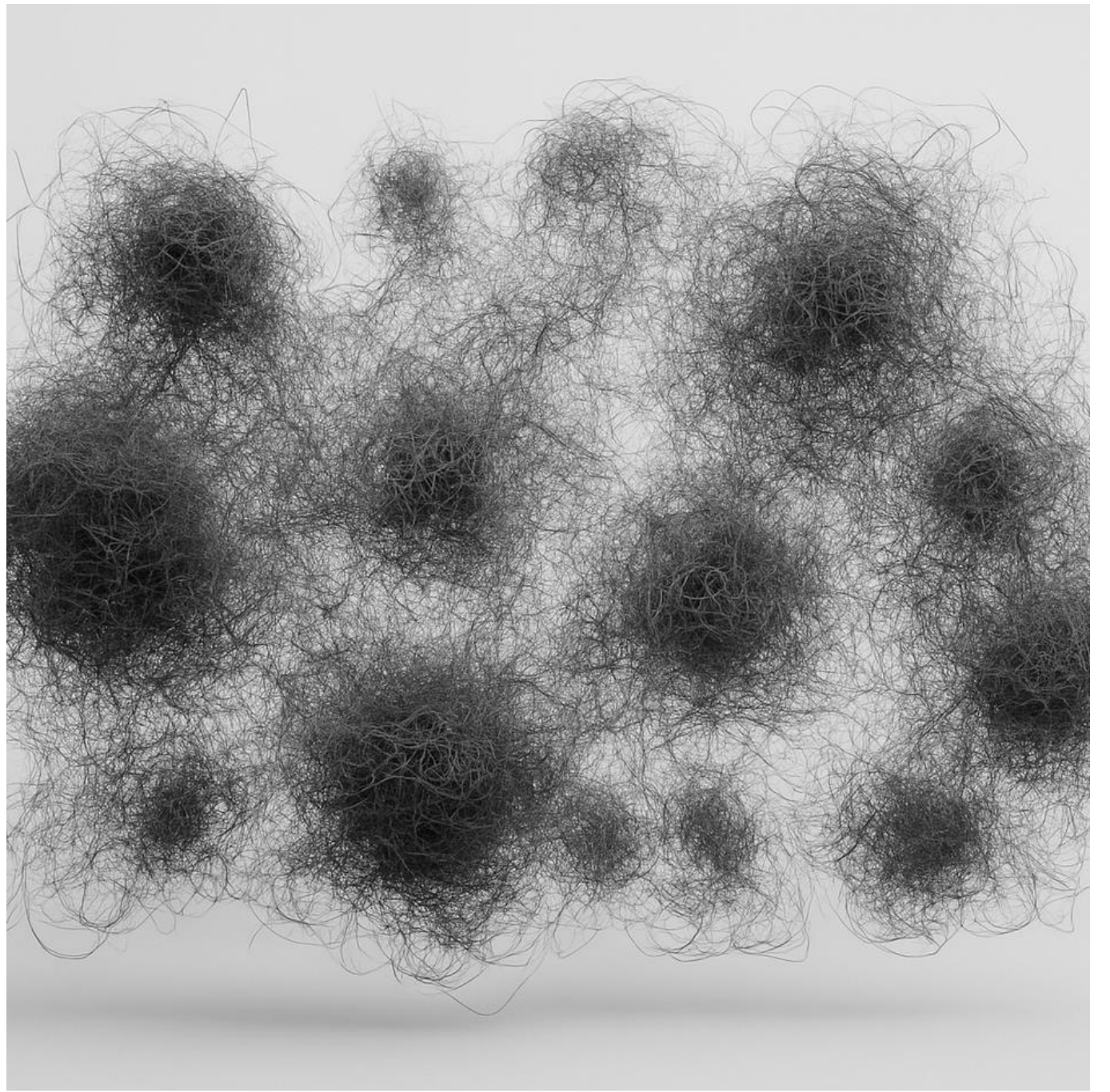
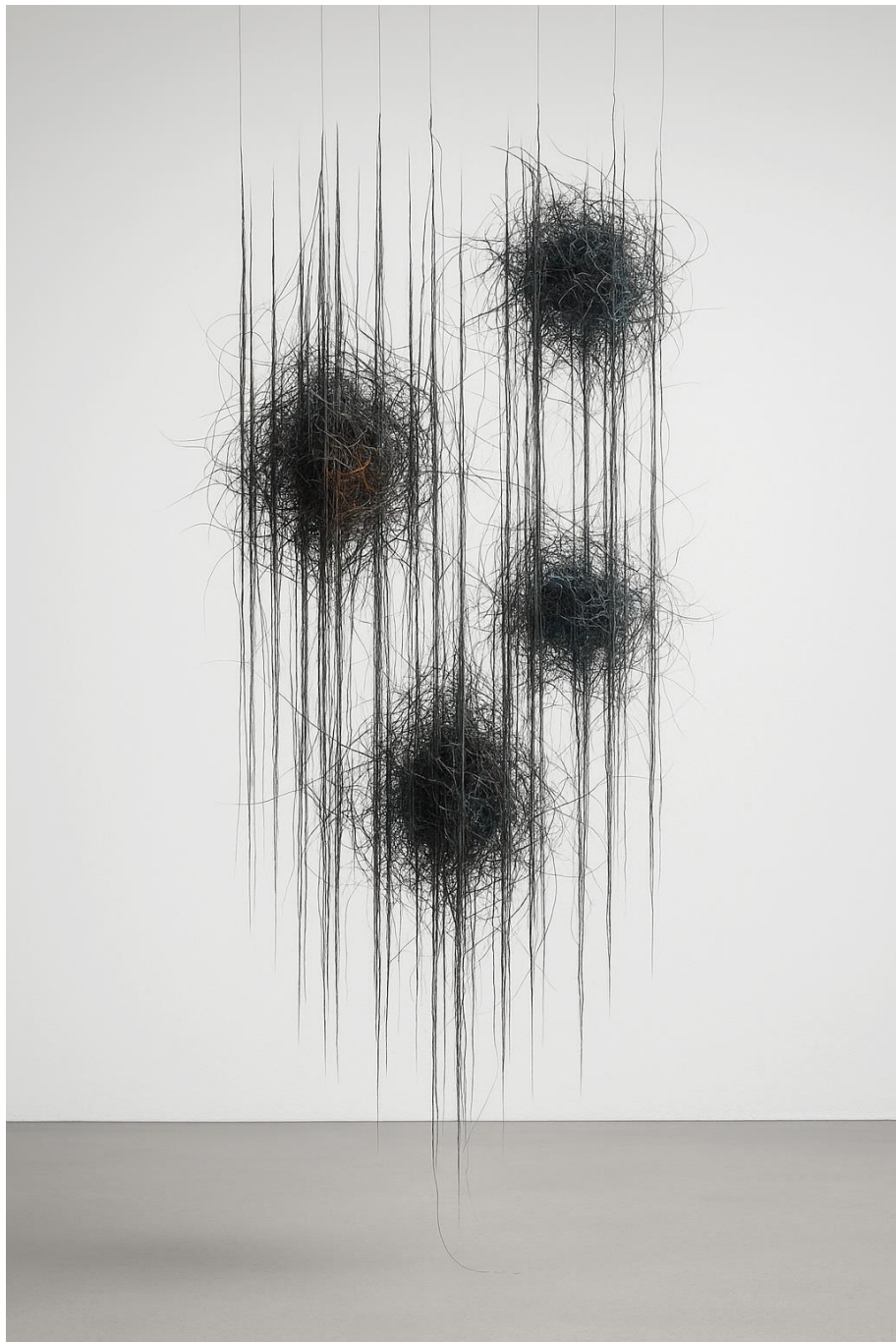
Founded in 2009, the Art & Science Factory brings together artistic and scientific methods to explore the complexity of contemporary global life













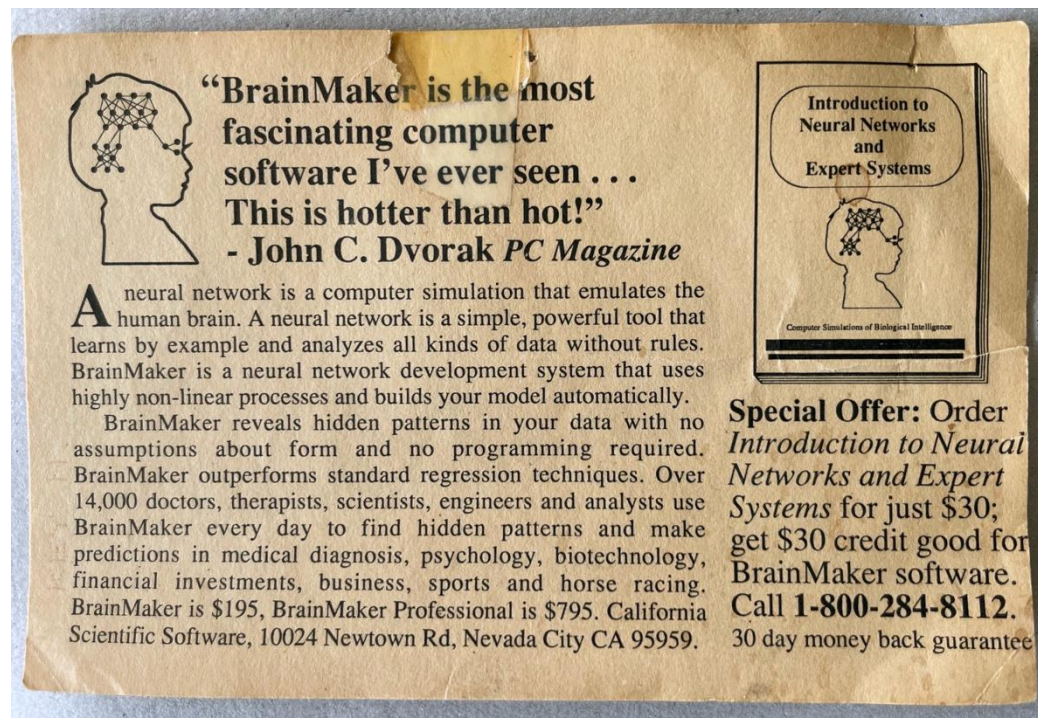
Machine Learning and Predictive Systems

- Examples include forecasting models, diagnostic systems, fraud detection systems, and risk prediction tools.
- These systems are designed to identify patterns within large datasets and make predictions, classifications, or recommendations.
 - Machine learning played an important role during the COVID-19 pandemic, helping researchers analyse viral genomes, predict protein structures, accelerate biomedical research, and support vaccine development and monitoring.
- Unlike LLMs, many machine learning systems are not conversational.
- They are often embedded within healthcare, finance, insurance, logistics, public services, and scientific research.

Machine Learning: A brief history

1950s–1980s: Early AI focused on rule-based systems and expert systems designed to mimic human reasoning through predefined rules.

1990s: A statistical learning revolution emerged, with neural networks returning alongside decision trees, Bayesian models, and support vector machines.



“BrainMaker is the most fascinating computer software I’ve ever seen . . . This is hotter than hot!”
- John C. Dvorak *PC Magazine*

A neural network is a computer simulation that emulates the human brain. A neural network is a simple, powerful tool that learns by example and analyzes all kinds of data without rules. BrainMaker is a neural network development system that uses highly non-linear processes and builds your model automatically. BrainMaker reveals hidden patterns in your data with no assumptions about form and no programming required. BrainMaker outperforms standard regression techniques. Over 14,000 doctors, therapists, scientists, engineers and analysts use BrainMaker every day to find hidden patterns and make predictions in medical diagnosis, psychology, biotechnology, financial investments, business, sports and horse racing. BrainMaker is \$195, BrainMaker Professional is \$795. California Scientific Software, 10024 Newtown Rd, Nevada City CA 95959.

Introduction to Neural Networks and Expert Systems

Computer Simulation of Biological Intelligence

Special Offer: Order *Introduction to Neural Networks and Expert Systems* for just \$30; get \$30 credit good for BrainMaker software. Call **1-800-284-8112**. 30 day money back guarantee

Postcard sent to me in 1992 when I was working as a statistician.

Machine Learning: A brief history

2000s: Machine learning quietly expanded into everyday life through fraud detection, credit scoring, search engines, customer analytics, and recommendation systems.

2010s: The rise of Big Data and deep learning transformed image recognition, speech recognition, translation systems, and large-scale predictive analytics.

2020s: Generative AI brought AI into public view through systems such as ChatGPT, Claude, Gemini, Midjourney, and DALL-E.

Machine Learning and Predictive Systems

When considering what AI, for whom, and under what conditions, a key concern is the predictive capacity of machine learning, particularly unsupervised approaches.

- who gets a loan,
- who gets insurance,
- who is flagged for fraud,
- which patient is high risk,
- which customer is targeted,
- which shipment is delayed,
- which neighbourhood receives resources.

That is a very different AI story from the current LLM story.



COMPLEX-IT

Run Online or Download for R-Studio



<https://www.complex-it-data.org/>

Select the number of clusters

3

Do you want to set a seed for reproducible results?

Build Your Model

1. Build database and import your cases
2. Cluster your cases
3. Confirm & Explore Your Model
4. Use AI to confirm your cluster solution
5. Compare and visualize your results

Run Scenario Simulations

6. Simulate your scenarios, policies, and interventions

Run Data-forecasting/classification

7. Use AI to predict the cluster membership of new cases

Systems Mapping Tab

8. Using Sytems Mapping To Explore Cluster Variables

Export Your Results

9. Generate your report

[Help](#)

Fuel.poverty	Social.Isolation	Get.NHS.Health.check
10.091	45.233	32.759
11.614	48.155	30.53
10.646	45.562	37.842
10.567	45.854	34.402

(x = km@uclusters, dist = dissM)

3 clusters C_j

$j : n_j | ave_{i \in C_j} s_i$

1 : 57 | 0.41

2 : 22 | 0.41

COMPLEX-IT

Team

As a team we are committed to advancing a **case-based complexity approach** to research, policy and practice in an effort to advance the study of social complexity and to support decision making. We each bring to the team a wide range of methodological and programming expertise and are proud of the truly transdisciplinary and international makeup of our work.



Brian Castellani
Durham Research Methods Centre
Durham University



Corey Schimpf
Department of Engineering Education
University at Buffalo, SUNY



Michael Ball
Computer Services
Kent State University, Ashtabula



Christopher Caden
Durham Research Methods Centre
Durham University



COMPLEX-IT 1.0.1 Beta



Import Your Cases and Map Your Theory <

Build, Confirm and Explore Your Model ▾

» Cluster Your Cases

» Use AI to Confirm Clusters

» Compare and Visualise Your Results

Run Scenario Simulations <

Forecast New Data <

Explore Systems Map <

Export Your Results <

Help <

COMPLEX-IT is part of the AM-Smart movement in methods

- Advances in the integration of smart technology with interdisciplinary methods has created a new genre, **approachable modelling and smart methods** – AM-Smart for short.
- AM-Smart platforms address a major challenge for applied and public sector analysts, educators and those trained in traditional methods: accessing the latest advances in interdisciplinary (particularly computational) methods.
- AM-Smart platforms do so through nine design features. They are
 - (1) bespoke tools that
 - (2) involve a single or small network of interrelated (mostly computational) methods
 - (3) they also embed distributed expertise
 - (4) scaffold methods use
 - (5) provide rapid and formative feedback
 - (6) leverage visual reasoning
 - (7) enable productive failure
 - (8) promote user-driven inquiry
 - (9) while counting as rigorous and reliable tools

Recommendation Systems

Examples include Netflix, Spotify, Amazon, TikTok, YouTube, and social media feeds.

These systems recommend content, products, services, or behaviours based on patterns of user activity.

They have become some of the most influential forms of AI in everyday life, shaping what people watch, read, purchase, and engage with online.

Decision-Support Systems

Examples include clinical decision support tools, business intelligence systems, policy modelling tools, and operational dashboards.

These systems are designed to assist human decision-making rather than replace it.

They often combine data analysis, prediction, visualisation, and recommendations.

Questions of trust, accountability, expertise, and professional judgment are central to their use.

Autonomous Agents

Examples include AI assistants capable of performing tasks across multiple applications and workflows.

Unlike systems that simply generate outputs, agents can increasingly take actions on behalf of users.

They may schedule meetings, conduct research, manage workflows, interact with software systems, and complete complex sequences of tasks.

This raises new questions about delegation, oversight, responsibility, and control.

Robotics and Embodied AI

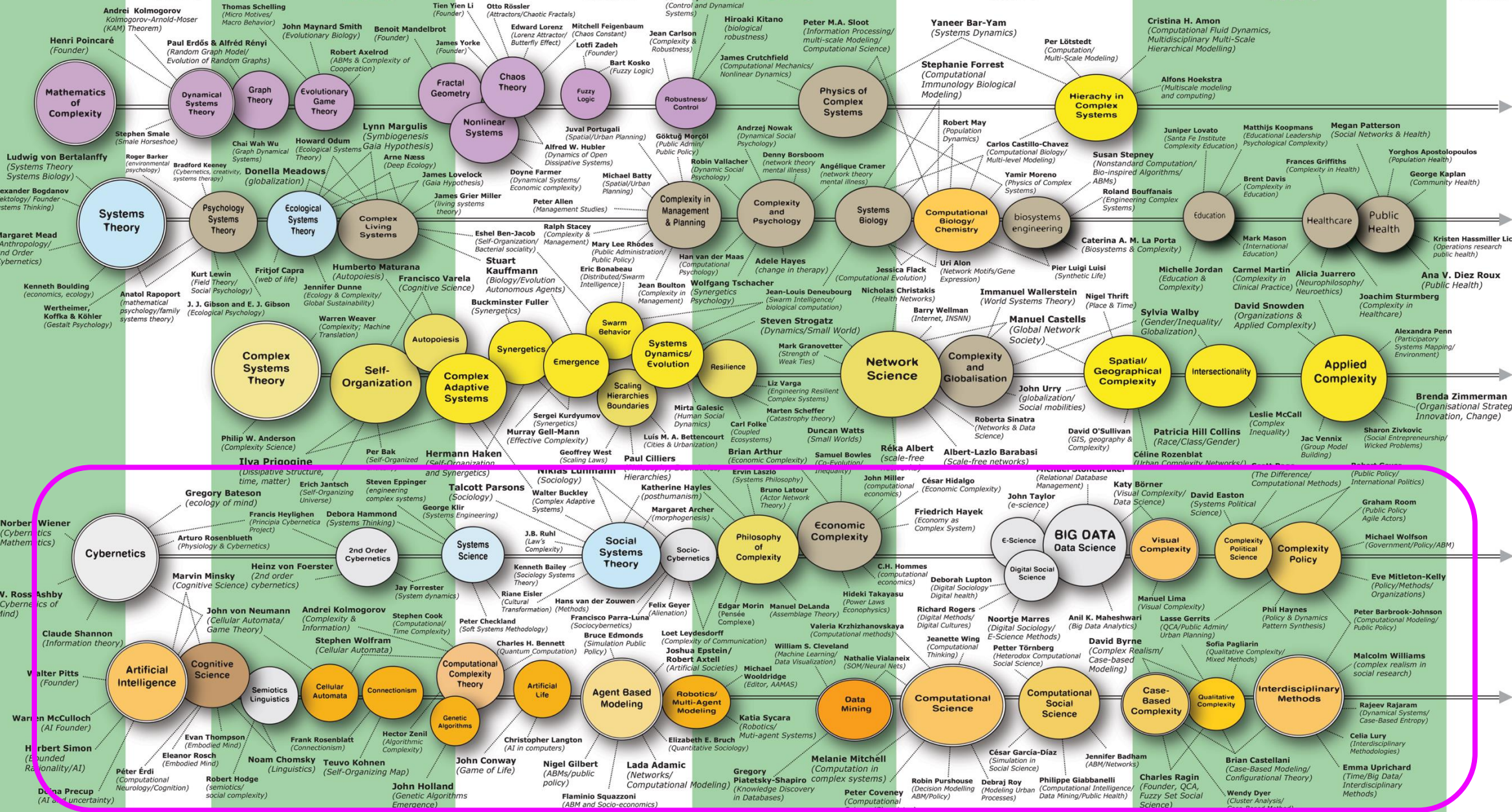
Examples include autonomous vehicles, warehouse robots, manufacturing systems, drones, and service robots.

These systems combine AI with physical action in the real world.

Their decisions have direct material consequences, making issues of safety, reliability, accountability, and human-machine interaction particularly important.

Unlike most other forms of AI, embodied systems must operate within physical environments rather than purely informational ones.

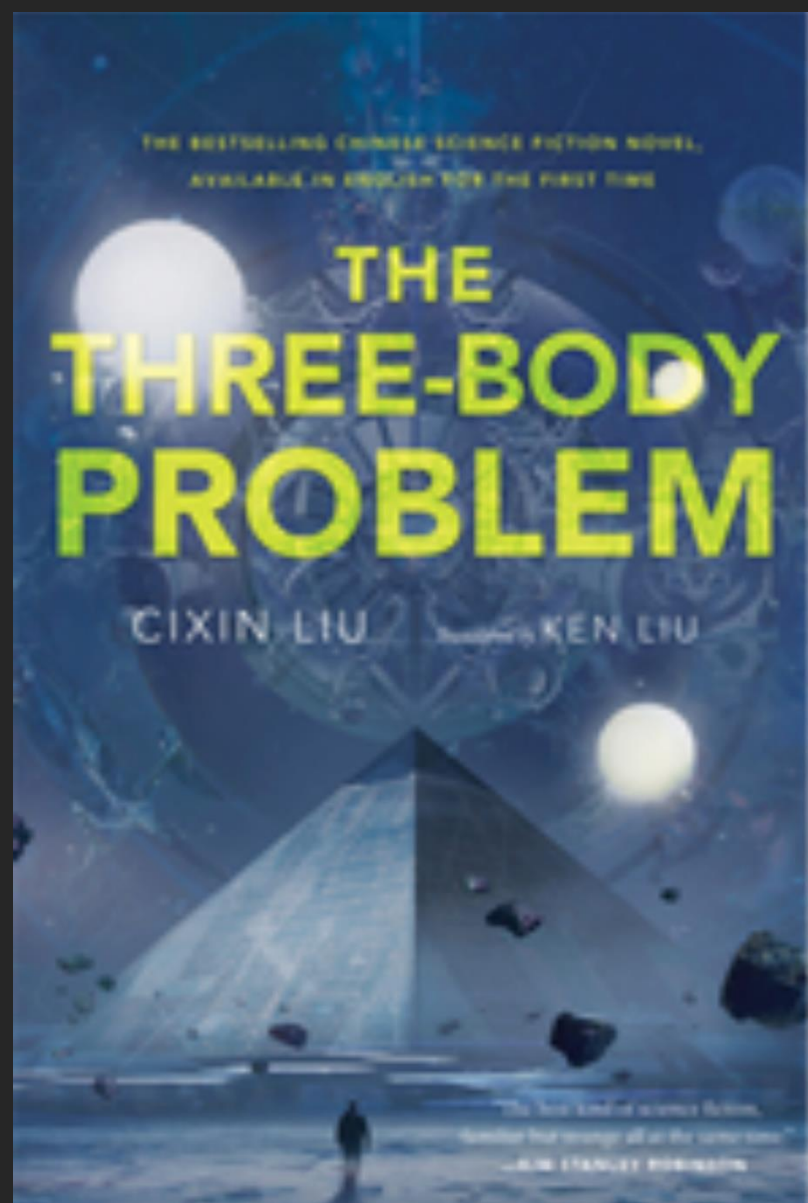
1940-1950s 1960's 1970's 1980's 1990's 2000's 2010's 2020's



AI TECHNOLOGIES ALSO EVOLVE – Again, what are we talking about?

Not only are there multiple forms of AI, but the technologies continue to evolve, overlap and newly emerge.

Which takes us to the dangers of *futurecasting*



Liu Cixin and the Challenge of Futurecasting

Remembrance of Earth's Past reminds us that the future is rarely shaped by a single technology. It emerges from the interaction of technologies, institutions, cultures, incentives, resources, politics, accidents, and human choices.

Key Insight

The further we move into the future, the more possible futures become available and the less confidence we should have in any single forecast.

Rather than asking what AI *will* do, we might ask:

Under what conditions might different AI futures emerge?



FUTURECASTING

The challenge is that **futurecasting** often shifts the conversation away from what AI is doing now and towards speculative claims about what it might do in the future.

This is slightly different from preparing for what one can see on the horizon or what one is presently dealing with – such as technology usage amongst adolescents.

As a result, debates can become driven as much by assumptions, expectations, and narratives as by evidence.

GOOGLE SEARCH

Originally (late 1990s–2000s)

- Primarily a search and ranking system.
- Heavy use of machine learning and statistical methods.
- Not really a recommendation system in the modern sense.

2010s

- Increasing use of machine learning and predictive systems.
- Predicting search intent.
- Ranking results and personalisation.
- Language understanding.

Today

- A hybrid system.
- Search engine and recommendation system.
- Machine learning and predictive system.
- Increasingly an LLM-powered assistant through AI Overviews and Gemini integration.

The Paperless Office

Throughout the 1980s and 1990s, many experts predicted that computers and digital communication would eliminate paper from organisations. Instead, paper use often increased as people printed emails, reports, presentations, and online documents.

Lesson: The technology changed work, but not in the simple way people predicted.

Radiologists and AI

In the late 2010s, many commentators predicted that advances in AI image recognition would largely replace radiologists. Instead, AI became integrated into radiology workflows, helping with image analysis and detection while radiologists continued to interpret findings and make professional judgments.

Lesson: The debate focused on whether AI would replace radiologists. A better question would have been: Which AI, performing which tasks, for whom, and under what conditions?

For whom?

Robotics and Embodied AI

Robotics and embodied AI are used by manufacturers, logistics companies, healthcare providers, defence organisations, and households to perform physical tasks, support human activity, and operate within real-world environments.

Autonomous Agents

Autonomous agents are increasingly used by organisations, businesses, knowledge workers, and consumers to automate tasks, manage workflows, conduct research, coordinate activities, and perform actions on behalf of users.

Decision-Support Systems

Decision-support systems are used by clinicians, policymakers, managers, public agencies, and other professionals to assist with diagnosis, planning, resource allocation, operational management, and complex decision-making.

Recommendation Systems

Recommendation systems are used by technology platforms, retailers, advertisers, and media companies to shape what people watch, read, buy, listen to, and engage with online.

Machine Learning and Predictive Systems

Machine learning and predictive systems are primarily used by organisations such as banks, insurers, hospitals, governments, retailers, and research institutions to forecast outcomes, identify risks, detect fraud, allocate resources, and support scientific discovery.

LLMs

Increasingly used by everyone.

But, again, there are different types of LLMs for different people, from specialist to frontier models.

What, then, is our point here?

1. There is no generic AI user.

When people discuss the risks, opportunities, benefits, or harms of AI, they often talk as if everyone is encountering the same technology in the same way.

But the experiences of a civil servant using decision-support tools, a teenager using TikTok recommendations, a scientist using machine learning, a warehouse worker interacting with robotics, and a student using ChatGPT are fundamentally different.

2. When we discuss the benefits, risks, opportunities of AI, we must ask: for whom?

Different people encounter different forms of AI, for different purposes, with different levels of expertise, power, access, responsibility, and vulnerability.

As a result, the opportunities, risks, and consequences of AI are unevenly distributed rather than uniformly experienced.

***Under what
conditions?***

'Context' tells us where something happens.

'Conditions' -- in contrast -- force us to specify what makes a claim true, false, useful, risky, or misleading.

When someone says 'AI works', 'AI is dangerous', 'AI improves productivity', or 'AI threatens expertise', the right response is not agreement or disagreement.

The right response is: under what conditions?

That means specifying the task, the user, the setting, the stakes, the level of expertise, the available oversight, the consequences of failure, and the wider social, cultural, linguistic, and institutional pressures.

Under What Conditions?

- What task is the AI (in all its forms) being used for?
- What level of expertise, training, and experience do they have?
- **What generation is using the AI?**
 - Children
 - Adolescents
 - University students
 - Early-career professionals
 - Experienced professionals
 - Older adults
- What access do they have to technology, connectivity, and digital infrastructure?
- In what part of the world do they live and what are the laws around usage? And what social, cultural, linguistic, and political contexts shape its use?

Under What Conditions?

- Are they using a free or paid version of the system?
- What information, data, and resources are available to them?

- What organisational, professional, or institutional pressures shape its use?
- What forms of oversight, verification, and accountability are available?
- What are the stakes if the AI gets something wrong?
- Who benefits from its use?
- Who may be disadvantaged, excluded, or harmed?
- What are the short- and long-term consequences of relying upon it?

- Are we in the present tense, the near future, the distant future?

***Which AI,
for whom
and under what conditions?***

**Thank you for
listening.**