

AI, Ethics, and Integrity

Exploring Trust, Authorship, and Governance

Dr Fotini Toso | Research and Innovation Portfolio | fotini.toso@rmit.edu.au

Dr David Blades | Research and Innovation Portfolio | david.blades@rmit.edu.au

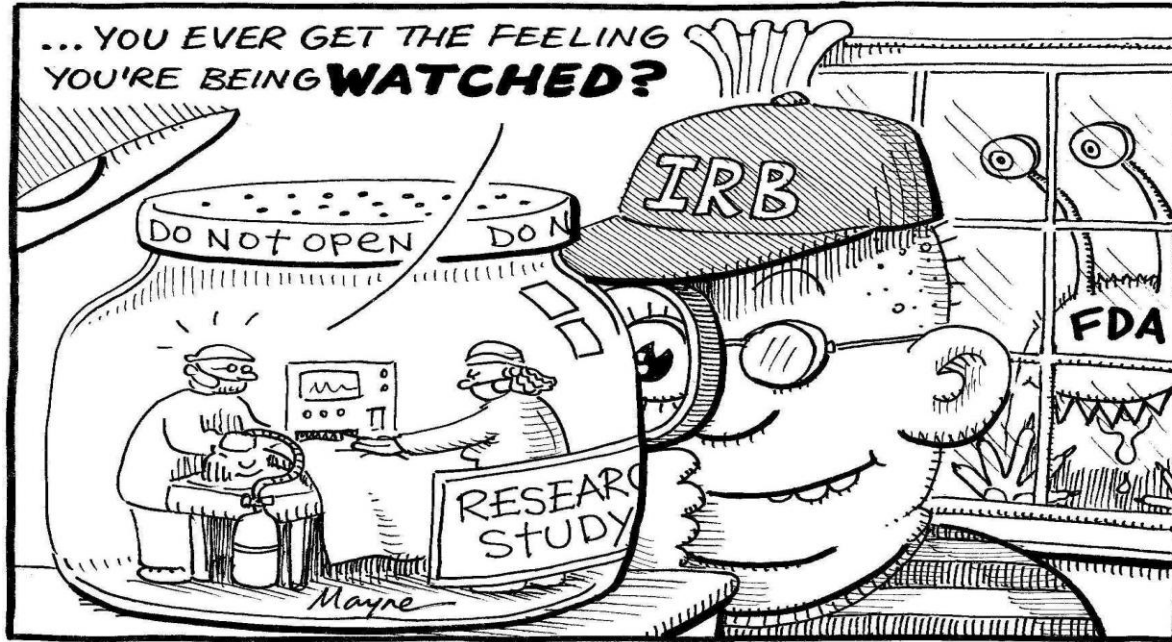


**We respectfully acknowledge our elders and the community
of this land and all elders and communities
of this continent and neighbouring islands.
We thank them for their courage, strength,
integrity and values.**

AI and Research Ethics

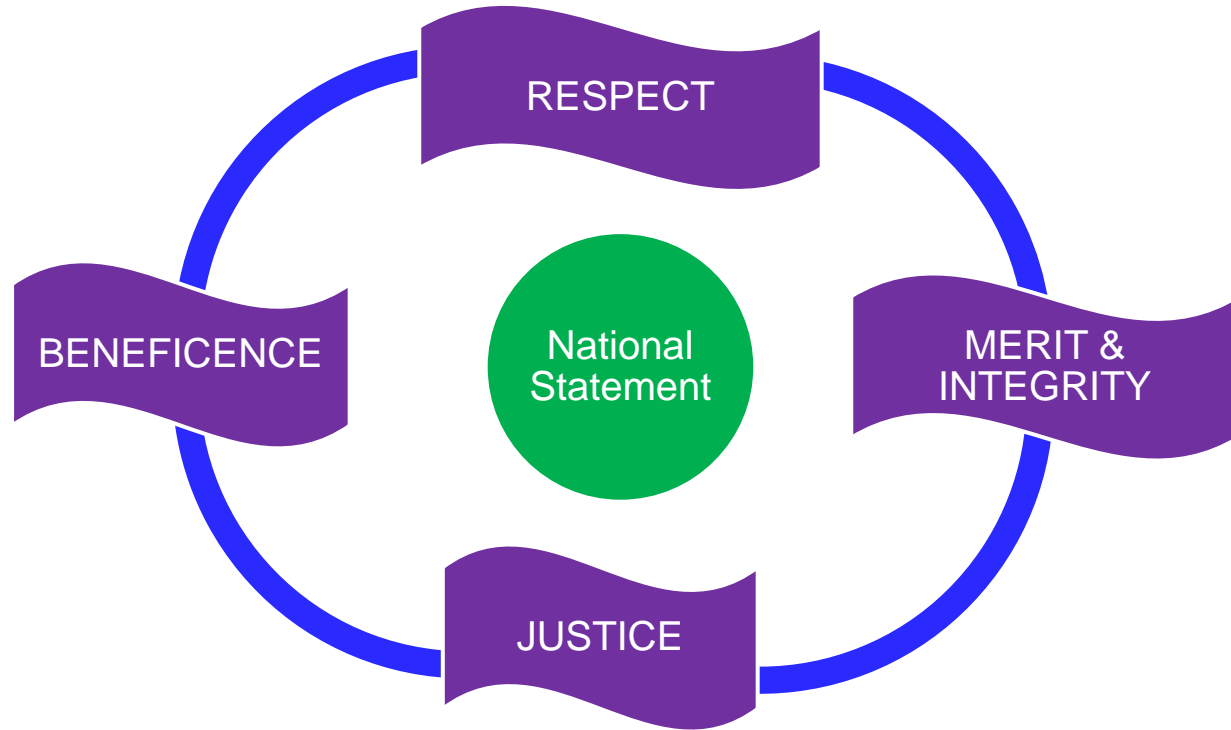
Dr Fotini Toso

What is human research?

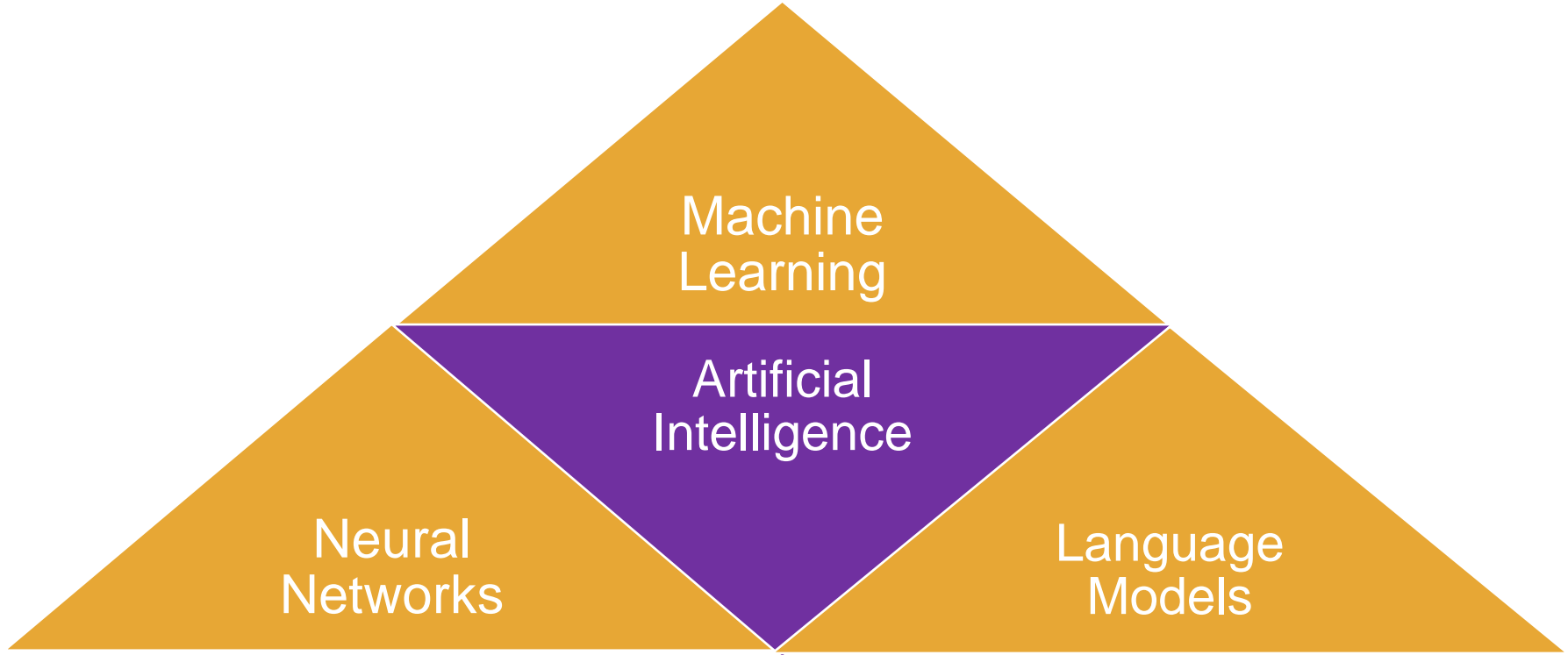


cartoon by Don Mayne, published on January 6, 2023, Friday Afternoon's Funny, AHRECS

Governance of Human Research



What is AI?



What is AI research?

- Research with existing AI
- Research to develop AI



IMPACT OF AI

SCIENCE / TECH / ARTIFICIAL INTELLIGENCE

AI suggested 40,000 new possible chemical weapons in just six hours



An instructor at the Fort Leonard Wood Chemical School, who is designated as an agent handler, carries the VX nerve agent to contaminate a jeep in one of the eight chambers used for training chemical defense on April 18, 2003 at Fort Leonard Wood, Missouri. Photo by Brendan Smialowski/Getty Images

/ 'For me, the concern was just how easy it was to do'

By [Justine Calma](#), a science reporter covering the environment, climate, and energy with a decade of experience. She is also the host of the Hell or High Water podcast.

Mar 18, 2022, 6:06 AM GMT+1 | [0 Comments](#) / [0 New](#)



<https://www.theverge.com/2022/3/17/22983197/ai-new-possible-chemical-weapons-generative-models-vx>



Regulation of AI

Current state

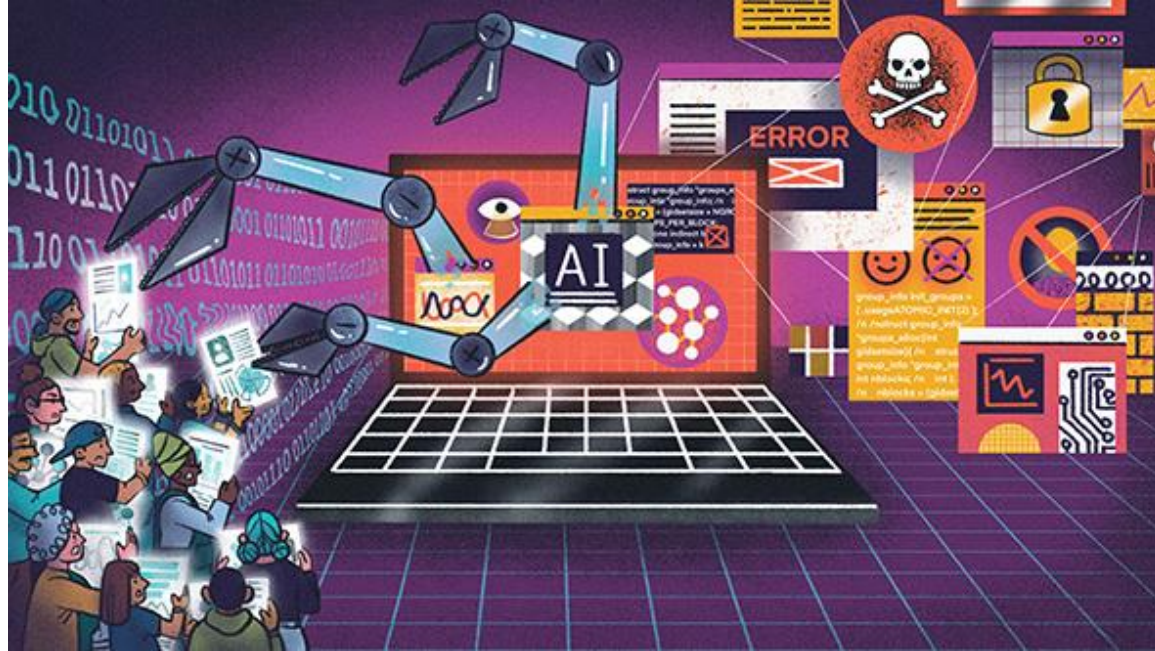


Illustration by Vreni Stollberger - Stanford Social Innovation Review



Why is regulation necessary?

Physical harms

- including injury, illness, pain;

Psychological harms

- including feelings of worthlessness, distress, guilt, anger or fear related, for example, to disclosure of sensitive or embarrassing information;

Devaluation of personal worth

- including being humiliated, manipulated or in other ways treated disrespectfully or unjustly;

Social harms

- including damage to social networks or relationships with others; discrimination in access to benefits, services, employment or insurance; social stigmatisation; and findings of previously unknown paternity status;

Economic harms

- including the imposition of direct or indirect costs on participants;

Legal harms

- including discovery and prosecution of criminal conduct.



AI and Research Integrity

Dr David Blades



No libraries, no roads,
no power lines.



A creative boundary

Is creativity required?

Are AI tools a source of data, like an archive?

Are AI tools a co-creator of knowledge, like an interviewee?

Are AI tools a co-contributor, like a co-author?

'one male professor one female professor speaking with robot in style of byzantine mosaic'

Midjourney, response to prompt from author, 9 March 2023





Criteria of authorship

Significant intellectual or scholarly contribution
(Authorship supporting guide 2019)

Contributor role taxonomy (Allen et. al. 2019)

Continuity and creditworthiness (Jenkins and Lin
2023)

'robot writing a book in style of byzantine mosaic'

Midjourney, response to prompt from author, 9 March 2023





Accountability

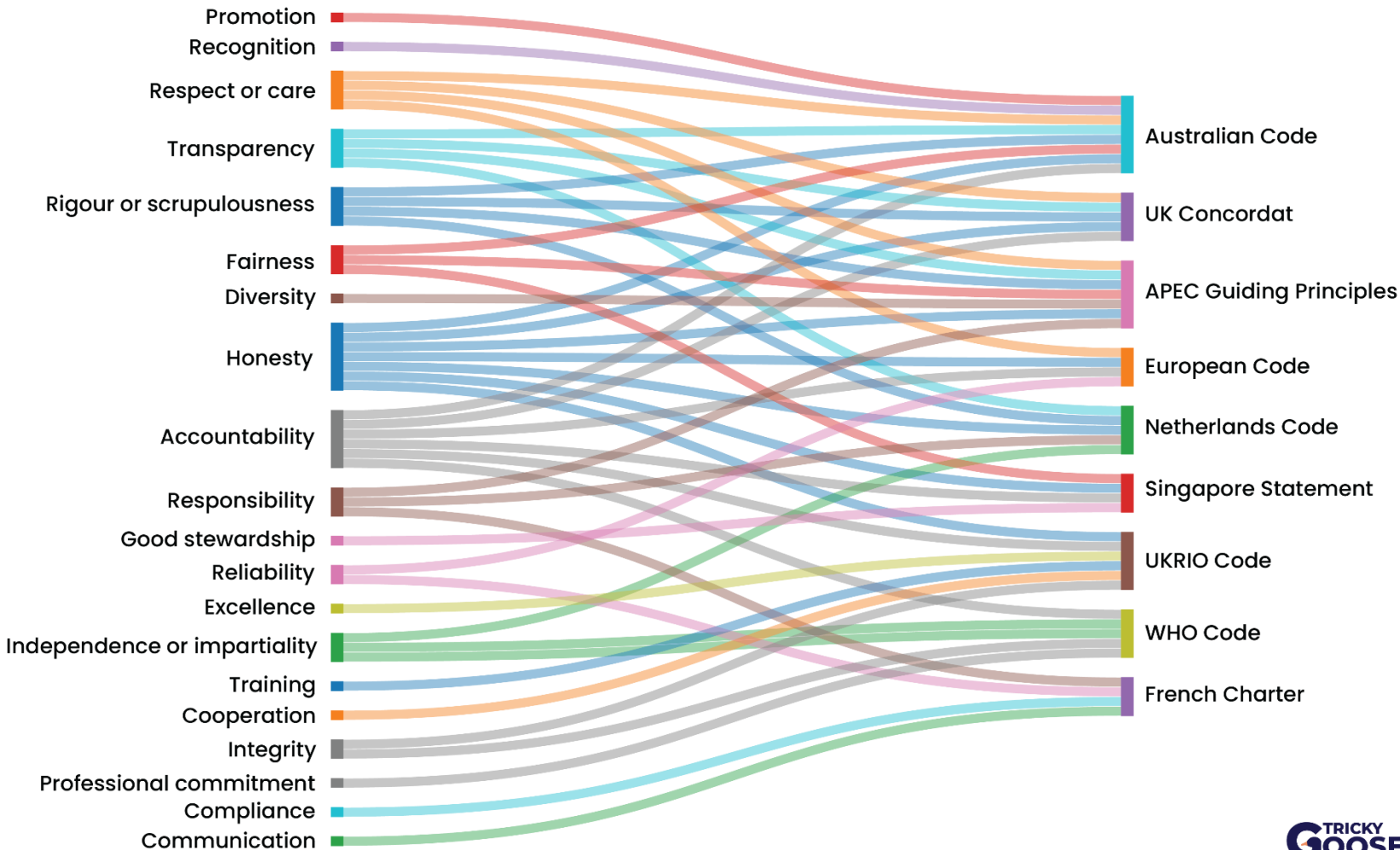
Can ChatGPT be accountable for what it writes?

As an AI language model, ChatGPT is not capable of being held accountable for what it writes because it is an algorithm that generates responses based on its training data and the input it receives from users. ChatGPT's responses are not based on personal beliefs or intentions, and it cannot be held responsible for any harm or misinformation that may arise from its responses.

'<byzantine mosaic> imposing AI'

Midjourney, response to prompt from author, 9 March 2023





Thank you!

Dr Fotini Toso | fotini.toso@rmit.edu.au

Dr David Blades | david.blades@rmit.edu.au

Parts of this presentation have been presented elsewhere.

