

Title: Ethics PROBES: Artificial Intelligence, Machine Learning, & Quantum Computing

Speakers: Joshua August Skorburg


Collection: Machine Learning for Quantum Many-Body Systems

Date: June 14, 2023 - 3:30 PM

URL: <https://pirsa.org/23060042>



Ethics PROBES: Artificial Intelligence, Machine Learning, & Quantum Computing



Joshua August (Gus) Skorburg

Assistant Professor of Philosophy

Academic Co-Director, CARE-AI

University of Guelph

skorburg@uoguelph.ca



[← All Open Letters](#)

Pause Giant AI Experiments: An Open Letter

We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.

Signatures

31810

[Add your signature](#)

Contents

[Statement](#)[Signatories](#)[Sign the statement](#)

Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war.

Signatories:

AI Scientists Other Notable Figures

Geoffrey Hinton

Emeritus Professor of Computer Science, University of Toronto

Yoshua Bengio

Professor of Computer Science, U. Montreal / Mila

Demis Hassabis

CEO, Google DeepMind

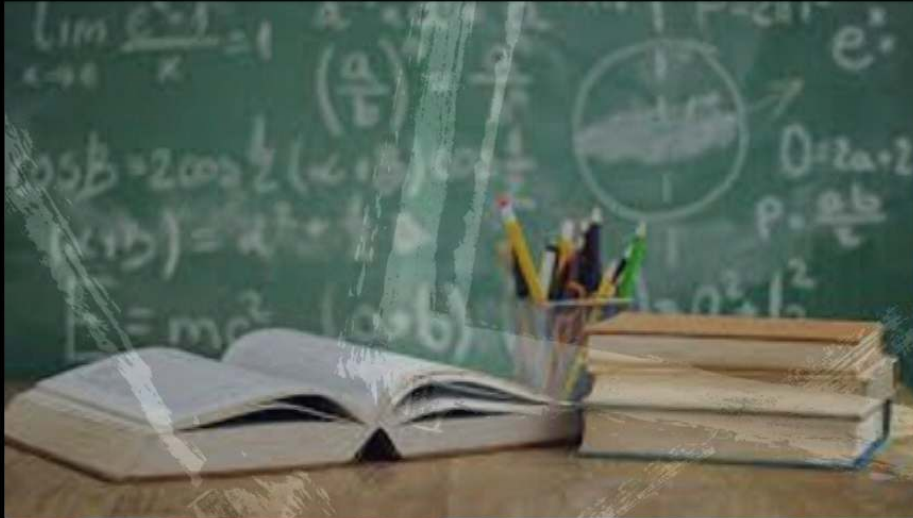
Sam Altman

CEO, OpenAI

Dario Amodei

CEO, Anthropic

Taking a step back from X-Risk



Probing the Ethics of AI/ML

- Privacy

Probing the Ethics of AI/ML

- Privacy
- Responsibility
- Opacity
- Bias
- Equity
- Sustainability

Training data, privacy and the “nothing to hide” argument



Training data, privacy and the “nothing to hide” argument



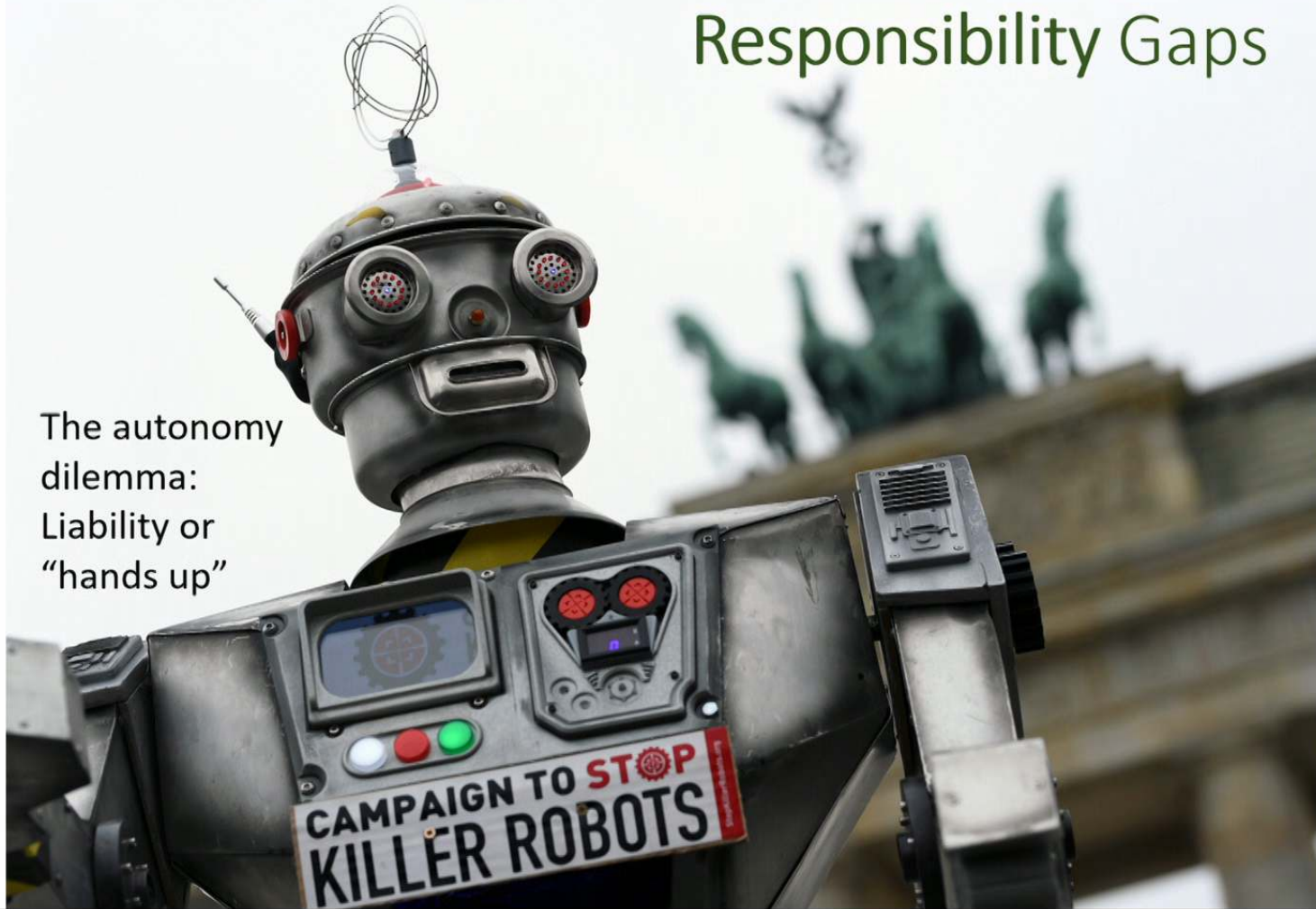
Value, mother-in-laws, social roles

Responsibility Gaps



Responsibility Gaps

The autonomy
dilemma:
Liability or
“hands up”



Opacity



Opacity

“The game of giving and asking for reasons”

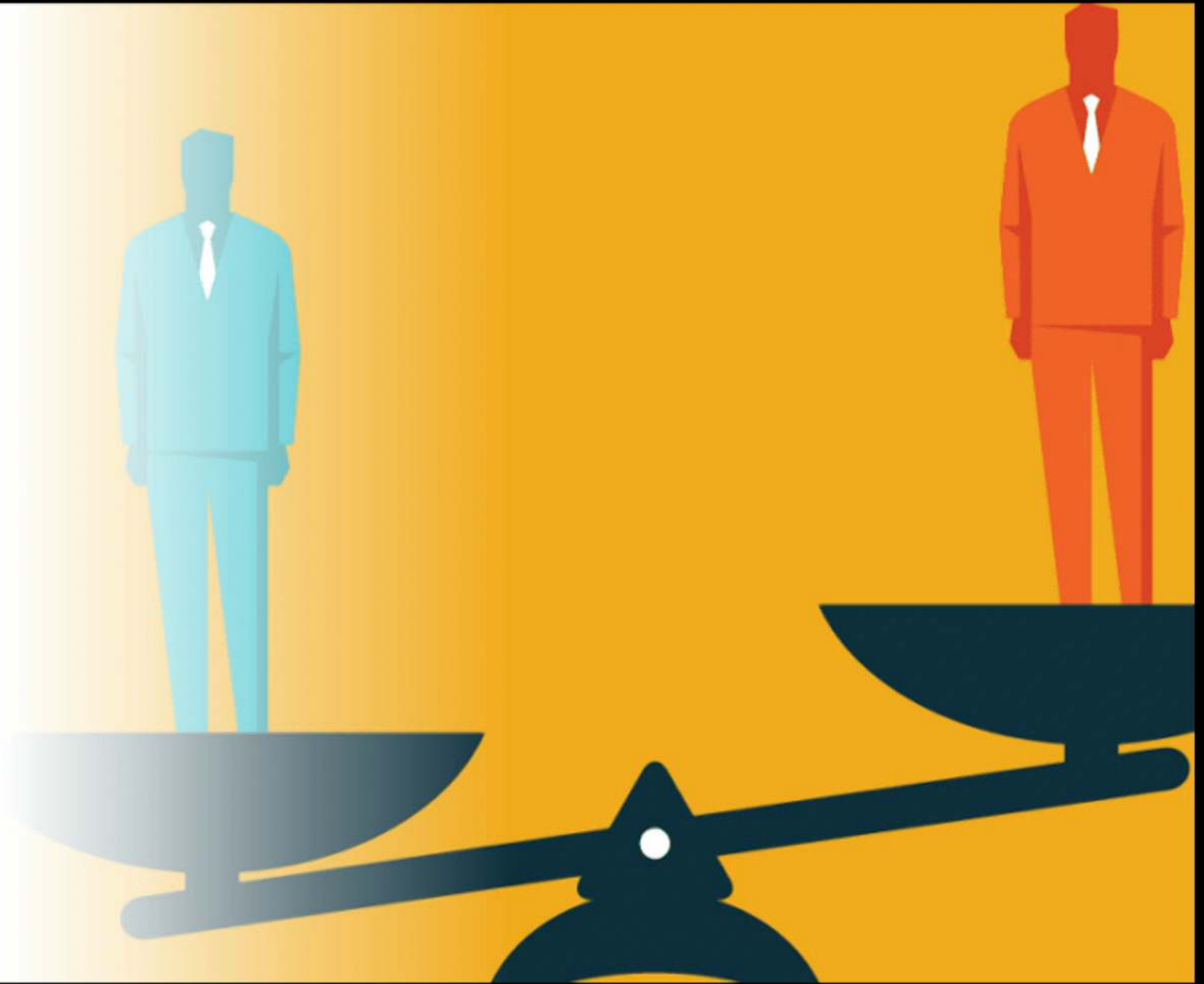


Opacity

“The game of giving and asking for reasons”

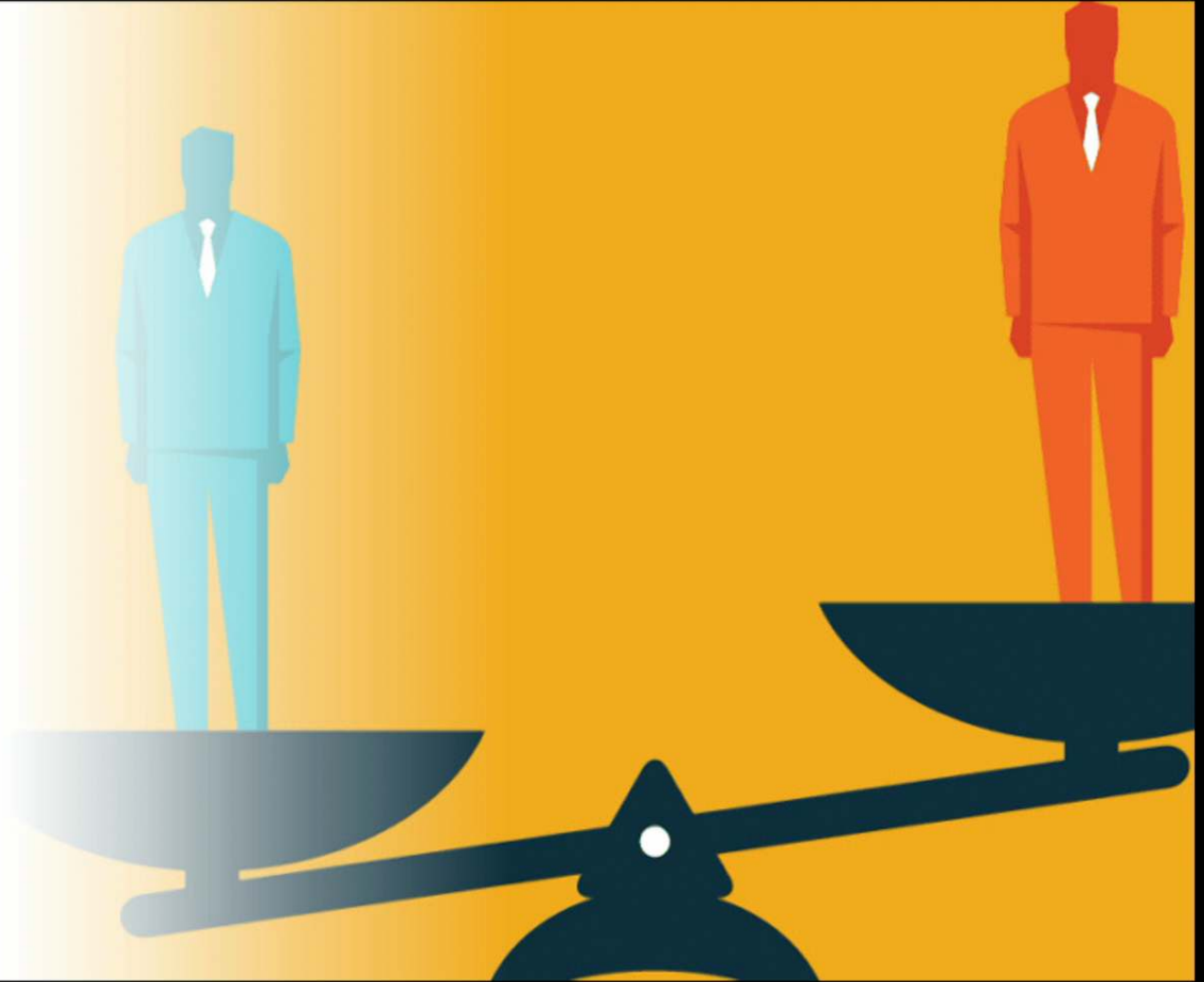
MNIST, the double-edged sword,
and double-standards

Bias



Bias

- Fazelpour & Danks (2021): Admissions



Bias

- Fazelpour & Danks (2021): Admissions
- **In problem specification**
 - Target variable: what is student success?
- **In data**
 - Systemic barriers, holding causes constant



Bias

- Fazelpour & Danks (2021): Admissions
- **In problem specification**
 - Target variable: what is student success?
- **In data**
 - Systemic barriers, holding causes constant



Bias

- Fazelpour & Danks (2021): Admissions
- **In problem specification**
 - Target variable: what is student success?
- **In data**
 - Systemic barriers, holding causes constant
- **In modelling and validation**
 - Error distribution



Bias

- Fazelpour & Danks (2021): Admissions
- **In problem specification**
 - Target variable: what is student success?
- **In data**
 - Systemic barriers, holding causes constant
- **In modelling and validation**
 - Error distribution
- **In deployment**
 - Feedback loops and data evolution



Equity and the digital (health) divide



Sustainability and the hidden environmental and human costs

Sustainability and the hidden environmental and human costs

Common carbon footprint benchmarks

in lbs of CO2 equivalent

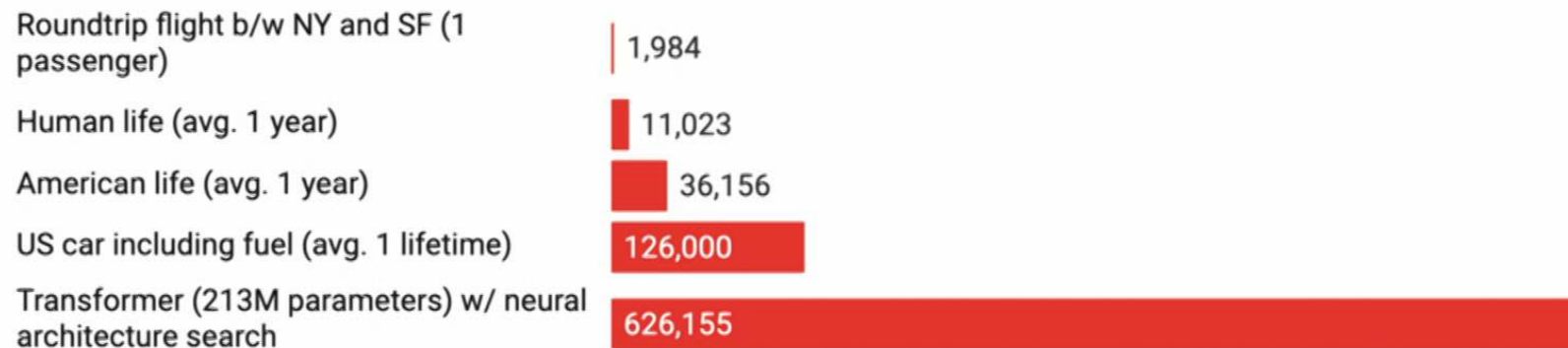


Chart: MIT Technology Review • Source: Strubell et al. • Created with Datawrapper

Continuity?

Continuity?

- Consider the rule “don’t harm innocents”
 - Stones vs firearms

Continuity?

- Consider the rule “don’t harm innocents”
 - Stones vs firearms
- Now consider broadcast vs targeted ads

Continuity?

- Consider the rule “don’t harm innocents”
 - Stones vs firearms
- Now consider broadcast vs targeted ads
- Are credible applications of AI/ML/QC more like the former or the latter?

Continuity?

- Consider the rule “don’t harm innocents”
 - Stones vs firearms
- Now consider broadcast vs targeted ads
- Are credible applications of AI/ML/QC more like the former or the latter?
- If the latter, who regulates?





If not you, then
who?

Quantum Applications?

- Privacy
- Responsibility
- Opacity
- Bias
- Equity
- Sustainability



Thanks for your attention