

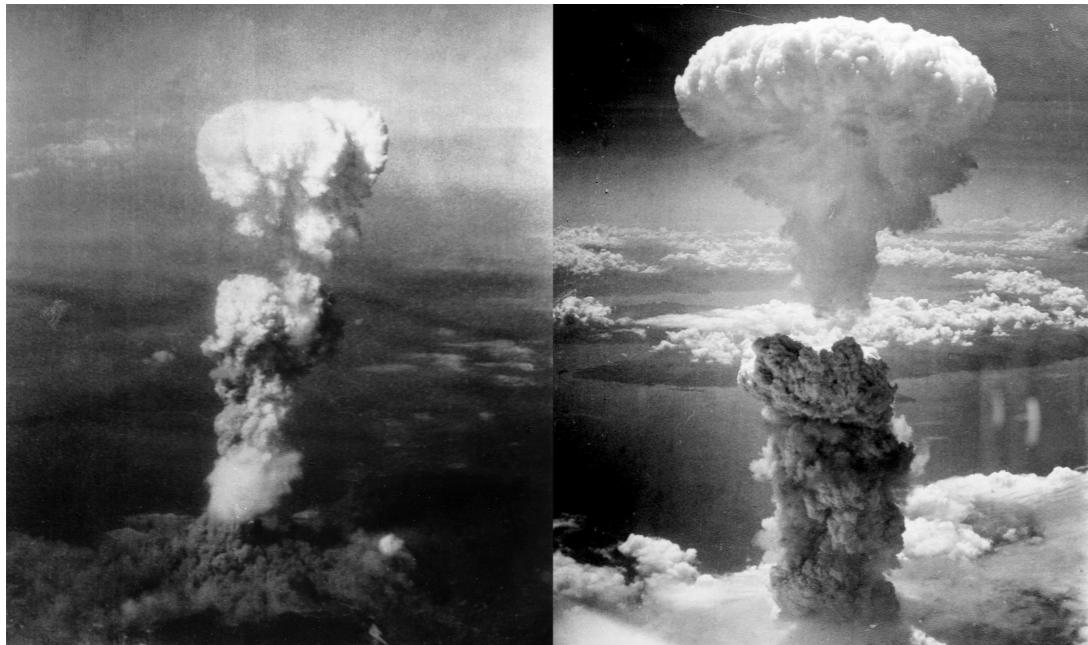
Introduction to Ethical and Social Implications of Artificial Intelligence (background)

22-Sep-2020
Dr. Atoosa Kasirzadeh

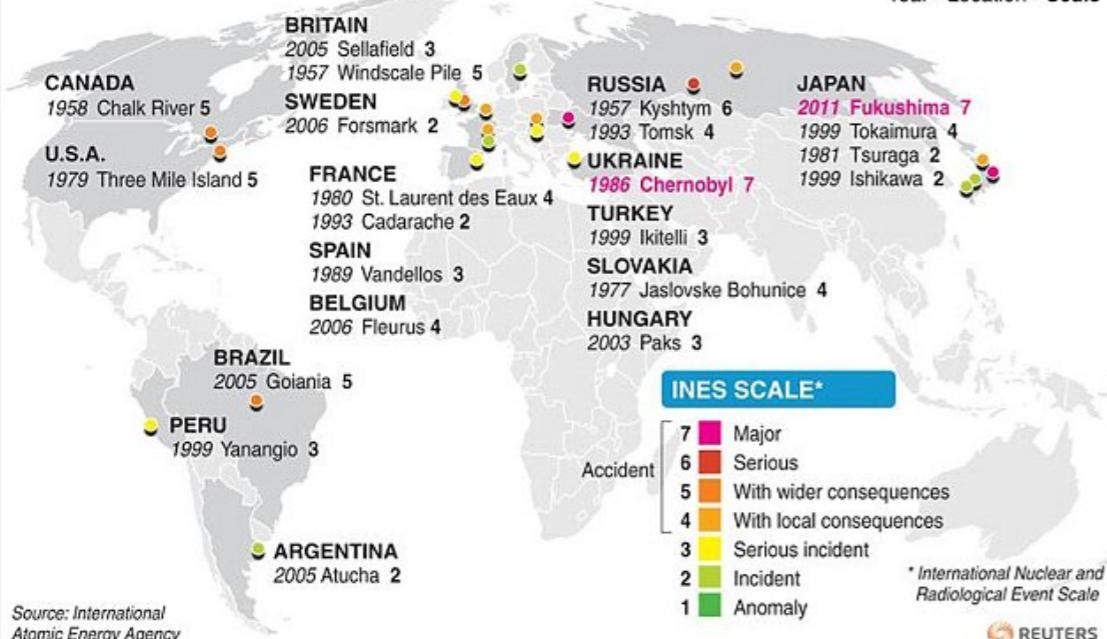
- The content of the lectures from week 7 to week 12 requires student engagement and discussion.
- I strongly encourage the students to come (= zoom in) to the lectures on Tuesdays and Wednesdays, and get engaged in the discussions. For this reason, I have decided not to live-record the lectures in class (except for our two guest lecturers), such that everyone feels more comfortable to talk knowing that nothing is recorded.
- After each lecture, I post the video lectures. They exclude class discussions and questions asked by students, but include the exact same material and slides that we go through in the class.
- I will hold office hours every week (starting from next week) for additional questions/discussions.
- More about the logistics, the course plan, office hours, evaluation methods, etc at the end of today's lecture.

Concerns of various sorts (ethical, social, political) always happen
in response to
the emergence of **new technologies**

Nuclear technology: energy, atomic bombs, accidents



GLOBAL NUCLEAR INCIDENTS SINCE 1956



Op-Ed: U.S. leaders knew we didn't have to drop atomic bombs on Japan to win the war. We did it anyway



Paper lanterns are floated on the Motoyasu River in Hiroshima every year to mark the anniversary of the dropping of an atomic bomb on the city by the U.S. (Toru Yamanaka / AFP/Getty Images)

By GAR ALPEROVITZ AND MARTIN J. SHERWIN
AUG. 5, 2020 | 3:05 AM

- SUBSCRIBERS ARE READING
- CALIFORNIA
'Like Armageddon': Rotting food, and chaos at postal facilities amid
- CALIFORNIA
Firefighters 'running on fumes,' b equipment, manpower amid fire
- SPORTS
Plaschke: I had COVID-19, and the things nobody tells you
- COMPANY TOWN
Kamala Harris' husband, Doug Emhoff, also make history
- LAKERS
Plaschke: Lakers' LeBron James g mode, becoming Playoff LeBron ii



Biotechnology

NEWS · 30 NOVEMBER 2018

First CRISPR babies: six questions that remain

Startling human-genome editing claim leaves many open questions, from He Jiankui's next move to the future of the field.

David Cyranoski



Science that inspires

Trends in Biotechnology

FEATURE | VOLUME 18, ISSUE 1, P8-10, JANUARY 01, 2000

Ethical issues in biotechnology

John C Polkinghorne

DOI: [https://doi.org/10.1016/S0167-7799\(99\)01392-X](https://doi.org/10.1016/S0167-7799(99)01392-X)

Abstract

Keywords

Keywords

References

Article Info

Related Articles

Abstract

New ethical questions have arisen from our ability to intervene in the structure of the genome. Responsible use of this technique requires ethical evaluation in which experts, potential beneficiaries and the general public should all participate. The examples of genetically modified food and of human genetics help to illustrate the issues involved.

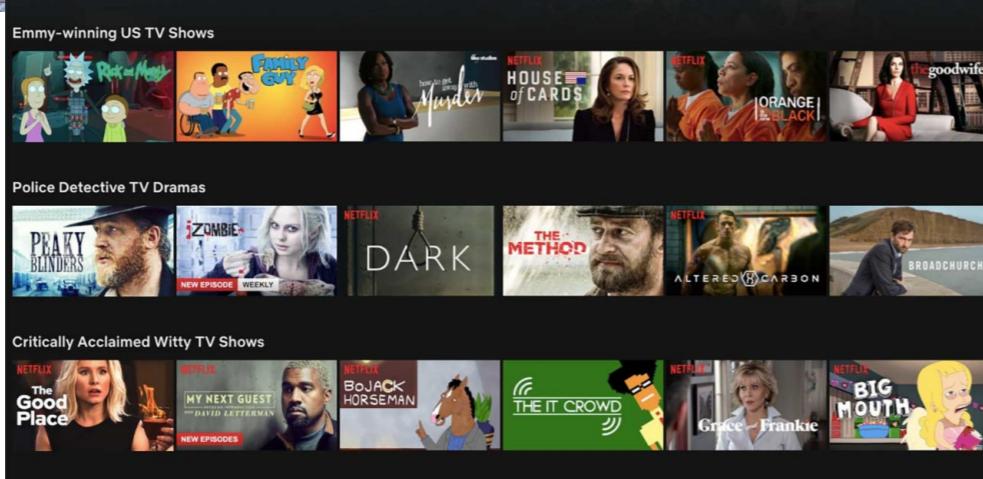
Some of these concerns are

Bizarre (trains are too fast for our souls)

Wrong (telephones will destroy personal communication)

 Broadly correct and moderately/deeply relevant (should we have nuclear technology? How should we effectively regulate AI?)

Digital technologies empowered by AI are increasingly used for



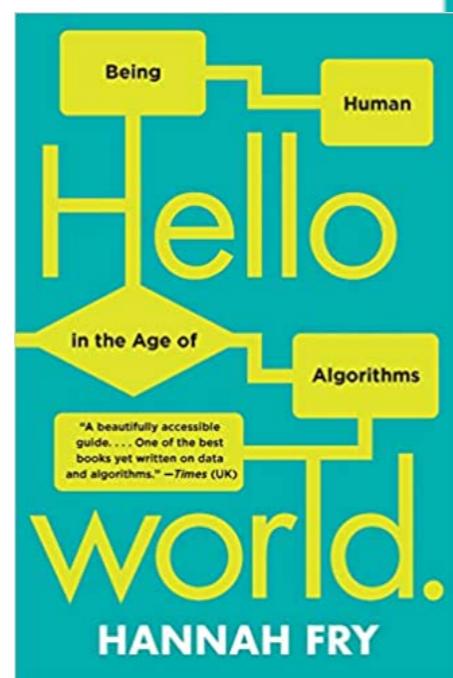
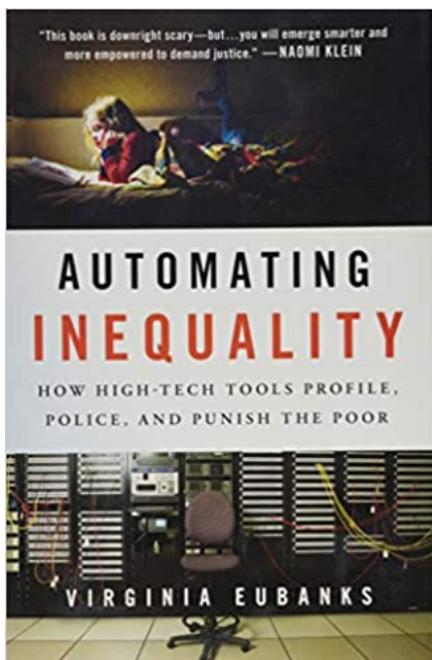
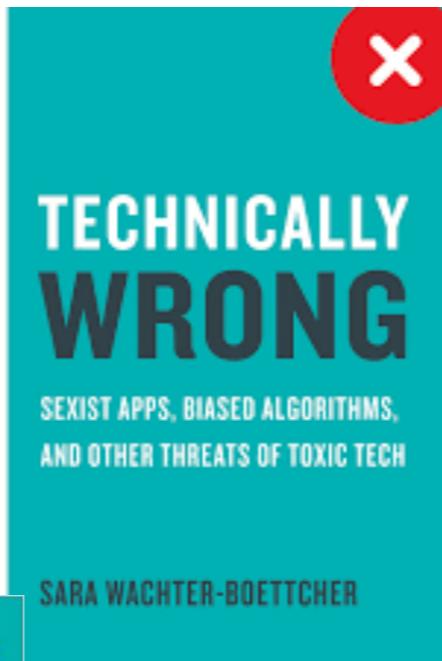
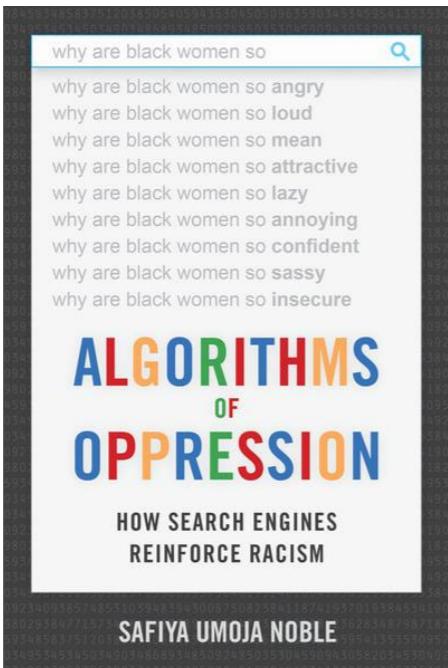
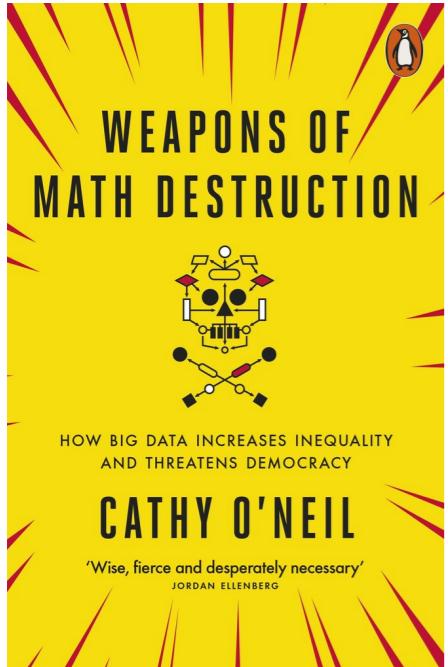
Following

@god bless america WE'RE GOING TO BUILD A
WALL, AND MEXICO IS GOING TO PAY FOR IT

RETWEETS LIKES
3 5

1:47 AM - 24 Mar 2016

Increasing public media attention to the ethical problems of AI



Here Come the Fake Videos, Too

Artificial intelligence video tools make it relatively easy to put one person's face on another person's body with few traces of manipulation. I tried it on myself. What could go wrong?



By Kevin Roose

March 4, 2018

Amazon

Amazon ditched AI recruiting tool that favored men for technical jobs

Specialists had been building computer programs since 2014 to review résumés in an effort to automate the search process



▲ Amazon's automated hiring tool was found to be inadequate after penalizing the résumés of female candidates.
Photograph: Brian Snyder/Reuters

How robots could help care for Japan's ageing population

"More people live alone these days, and a robot can be a conversation partner for them. It will make life more fun."

Photography, Kim Kyung-Hoon | Monday 09 April 2018 14:27



Artificial intelligence Jun 26

A new US bill would ban the police use of facial recognition



The news: US Democratic lawmakers have introduced a bill that would ban the use of facial recognition technology by federal law enforcement agencies. Specifically, it would make it illegal for any federal agency or official to "acquire, possess, access, or use" biometric surveillance technology in the US. It would also require state and local

HOME » NEWS » WORLD NEWS » EUROPE » GERMANY

Robot kills man at Volkswagen plant in Germany

A 22-year-old worker was grabbed by the robot and crushed against a metal plate



The accident happened at one of Volkswagen's production plants in Germany Photo: DPA/Alamy

By Justin Huggler, Berlin

1:56PM BST 02 Jul 2015

A robot has killed a man at a Volkswagen car factory in Germany.

The 22-year-old worker died from injuries he sustained when he was

A difficult problem: utopian benefits and dystopian problems

- Difficulty: AI can have positive and negative impacts
- Image recognition can help diagnosing diseases such as cancer and Alzheimer
- **Should** autonomous lethal weapons be allowed at all?
- Who is **responsible** when something goes wrong?



- Long term concerns: existential threat, super intelligence, singularity, global catastrophic risk
- Short—medium term concerns: is it good to use facial recognition system? Most of our discussions in the next 6 weeks focus on such concerns.
- (And, of course, there is always an interplay between short-medium and long-term goals)

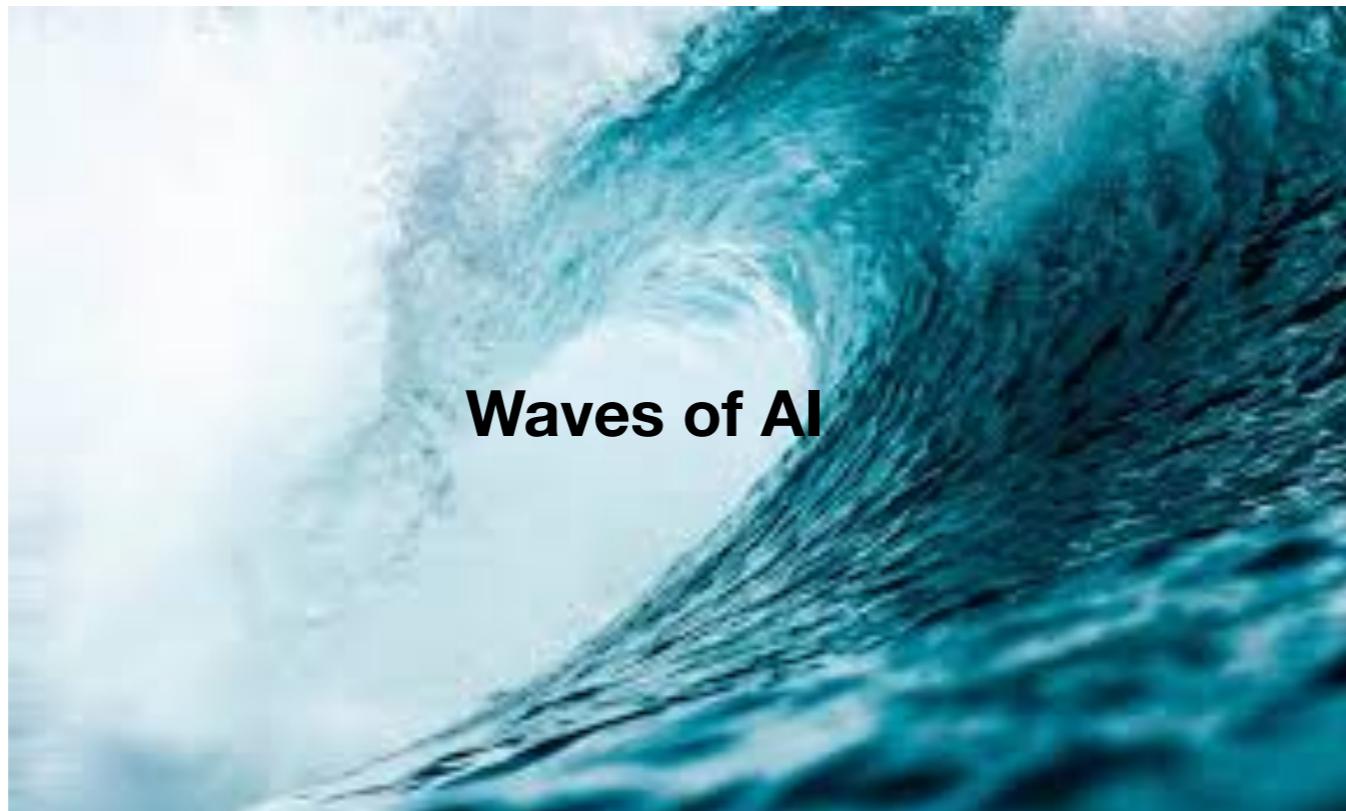


What do we mean by AI?

What do we mean by Ethics of AI?

What do we mean by Ethics?

What do we mean by AI?



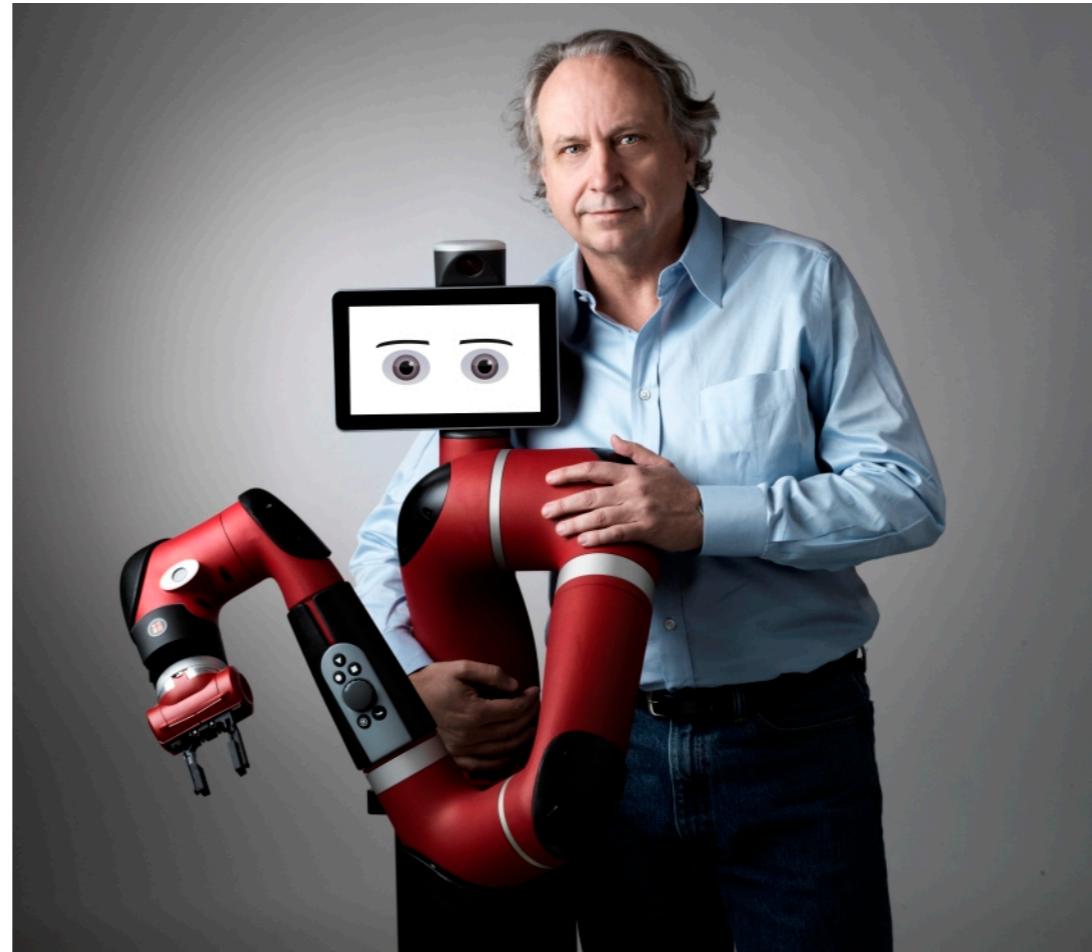
(Not mutually exclusive groups)

Origin: 1956 Dartmouth College

1. Logic-based symbolic manipulation (Intelligence with representation)

- The activity of the system as “reasoning” about an external world
- The reasoning process is implemented by two components
 - A. an inference system operating over
 - B. a model of the task domain

2. Embodied robotics



Use the world as its own model; Intelligence without representation

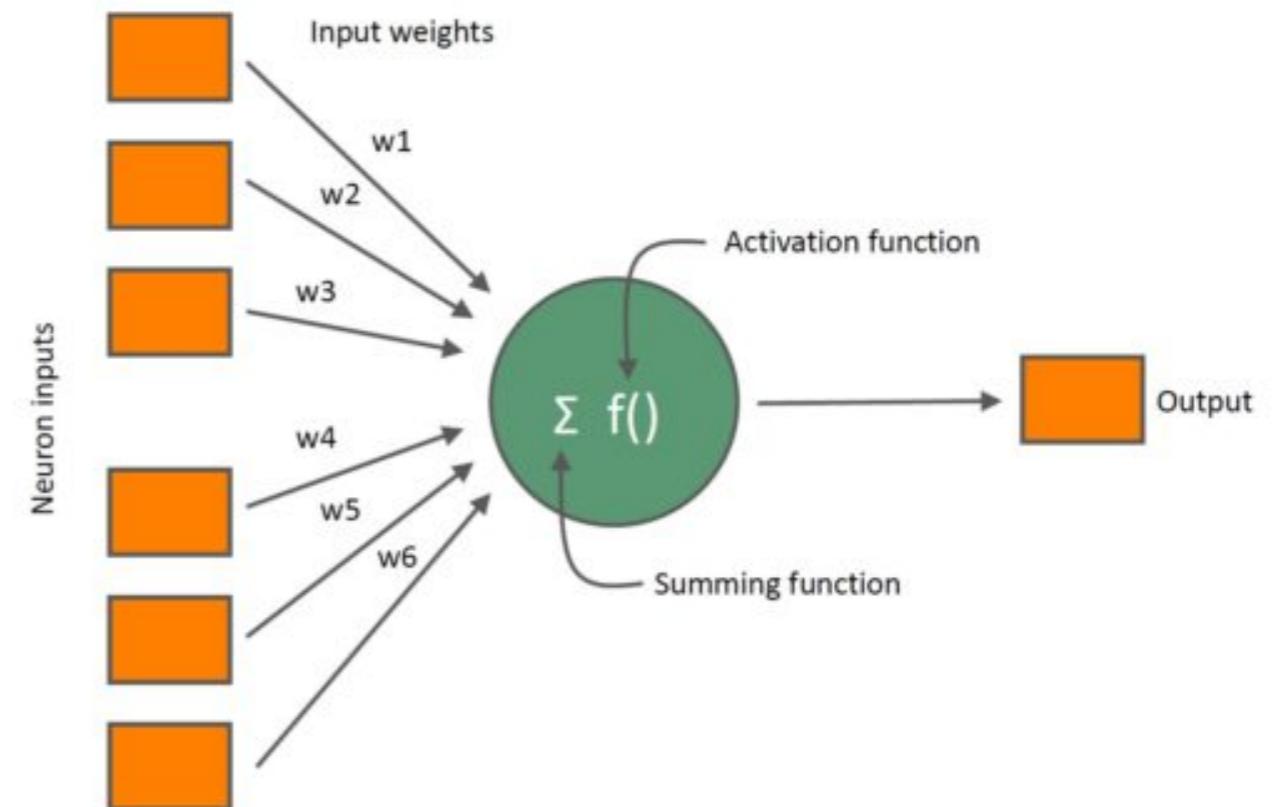
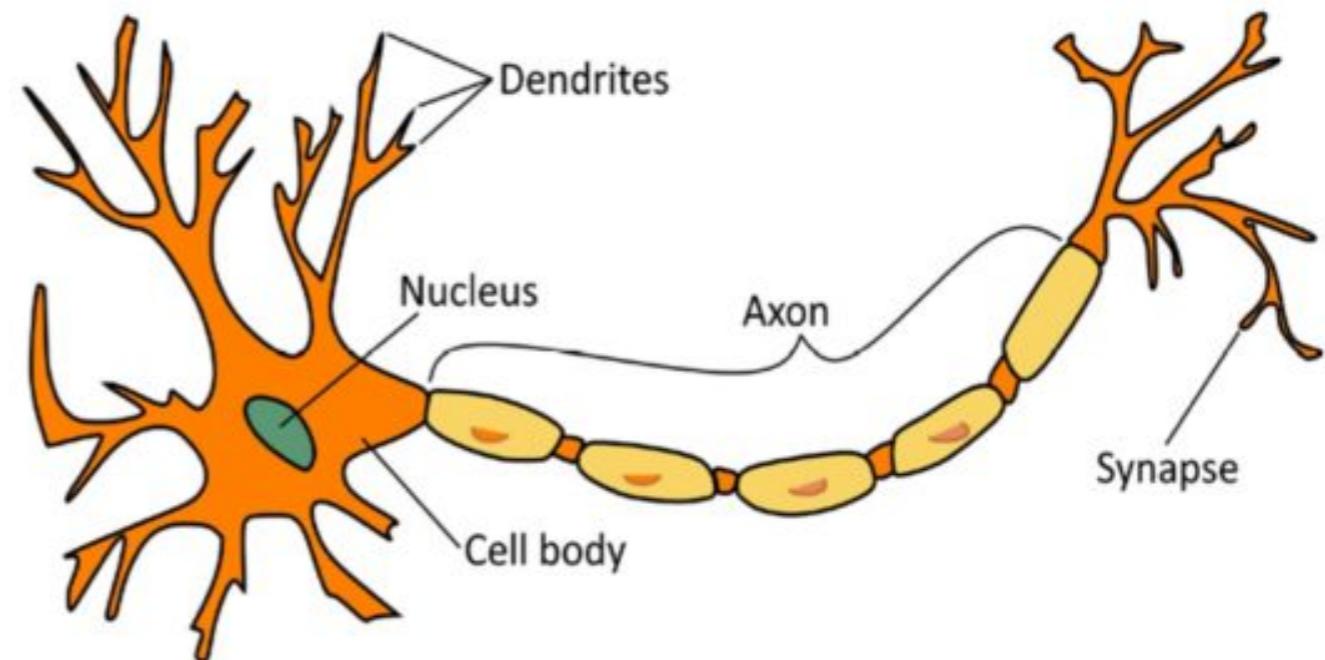
Don't try to build what we humans think an “intelligent robot” would be like

Instead, incrementally build up the capabilities of intelligent systems from the bottom, constructing complete systems at each step of the way and thus automatically ensure that the pieces and their influences are valid

3. Statistical learning

Exemplar:

Neural networks + large datasets



Let AI systems learn about ethical behaviour

Force AI researchers follow ethical codes

Align AI systems with human values

Code ethical theories in AI systems

What do we mean by Ethics of AI?

Define some ethical principles and operationalise them into AI systems

Resolve the problem of the centralised power (surveillance, privacy, power held by a few mega corporations)

Regulate AI (example: General Data Protection Regulation)



Align AI systems with
human values

What could we mean by Ethics of AI?

(Next week)

What could we mean by Ethics of AI?



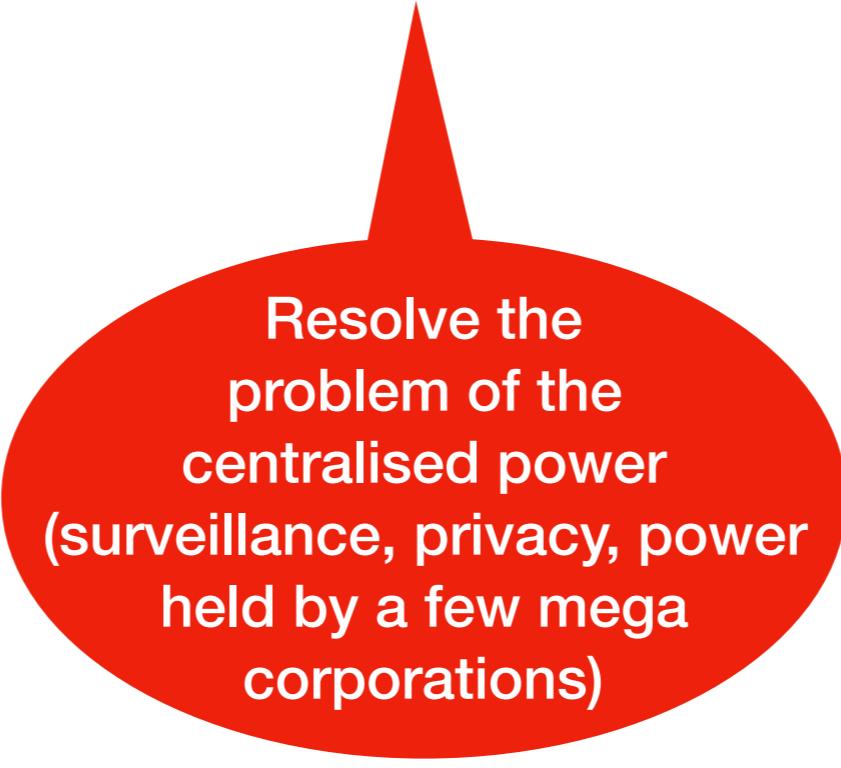
Define some ethical principles and operationalise them into AI systems

(Week 9 & Week 10)

What could we mean by Ethics of AI?



Define some ethical principles and operationalise them into AI systems



Resolve the problem of the centralised power (surveillance, privacy, power held by a few mega corporations)

(Week 11 & a bit of Week 12)

Now to office hours, evaluation, and assignments

- Zoom Office hours (AEST, every Wednesday from the 30th of September, 4pm—5pm). I will post the link to my virtual zoom meeting room on the announcements. If you cannot make it to these hours, write to me at atoosa.kasirzadeh@anu.edu.au requesting an appointment. Please start the title of your email by “COMP4620— ...”
- I have posted some optional materials, with links to some youtube videos. For those who aren’t in Australia, can you let me know what website you use for watching English videos and English news? Please share some information with me on this issue.
- Those who require special services, please write to me about your needs.

Evaluation

- 10%: Participation + 4 weekly questions
- 15%: Assignment 1
- 25%: Assignment 2: Final term paper

Participation + 4 weekly reflections (10%)

- Give a comment in class or in my office hours or post to the Wattle discussion forum
 - You should post any interesting news item related to the “ethical and social implications of AI” that you happen to read to the wattle discussion forum, and take a moment to reflect on their significance and relevance to our in-class topics and conversations. Or you could respond to any question in class, engage in a conversation, or come to my office hours with a question/discussion. At least one time engagement (either in class, or in my office hours, or on wattle discussion forum is required. 2%
- Every Friday (starting from the 25th of September for four weeks), I post a question in the quiz format on Wattle. You have time until Monday 5pm to submit your answers. The purpose of the question is to nudge you thinking about the lecture’s topic before I give the presentation (the question prepares you for the presentation). You can spend up to 15 minutes to write your reflections in response to the answer. 2% for each question. 8% for the 4 questions.

Assignment 1 (15%)

Critically assess a proposed AI project (~500 words)

- Drawing on the readings for the class and the topics we cover in lectures, you will critically assess a proposed AI project. I will provide you with a brief description of the project, and you need to identify 3 potential problems with the proposed project that could raise ethical and social concerns. Your answer should take the form of 3 paragraphs, each paragraph including (1) the identification of the source of the problem and (2) explaining how the social and ethical issues are at stake. The problems that you identify should be as distinct as possible.

Details about the submission procedure will be posted by next week.

Assignment 2: term paper (25%)

Write a final paper that revisits a recent controversy (~2000 words)

(I will give you a lecture on how to write such a paper [Week 9 probably] and will point you to some further readings on writing such a paper.)

- In a double-spaced paper, you will revisit a recent controversy involving different course themes. You can choose a case from a set of 5 predetermined cases. Your paper should draw extensively from the course materials, lectures, and in-class discussions, and present a comprehensive plan for undertaking the project in a way that addresses ethical principles such as privacy, transparency, and fairness. The paper should also point to some difficult tensions and unavoidable trade-offs. You should explain why it is not possible to satisfy and have all the ethical requirements if the project is going to run. After that, you should provide a thoughtful justification for the specific trade-offs that you suggest.

Details about the submission procedure will be posted by next week.

- Any questions?
- Tomorrow, we talk about the main debates in “Ethics of AI” or the “Ethical and social implications of AI”