Ethics before data ethics

INTRODUCTION TO DATA ETHICS



Shalini Kurapati, PhD
Co-founder and CEO, Clearbox AI

What is ethics?



- Greek word *ethos* - way of life
- Moral philosophy branch, standards for right and wrong behavior
- Set of standards, rules, values
- Sense of right and wrong beyond laws- e.g. respecting others, offering a seat
- Guide decisions and actions- food, work, politicians

[1] Icon made by iconixer from www.flaticon.com

Why study ethics



- Decision making is not black and white
- Ethical dilemmas
  - Friendship vs. professional integrity

Innovation vs. affordable medicine



- Ethical dilemmas
  - Pharma innovation vs. unaffordable medicine
- Impact on others
  - Individual vs. societal benefits
  - Financial market dynamics, climate change

[1] Photo by Louis Reed on Unsplash

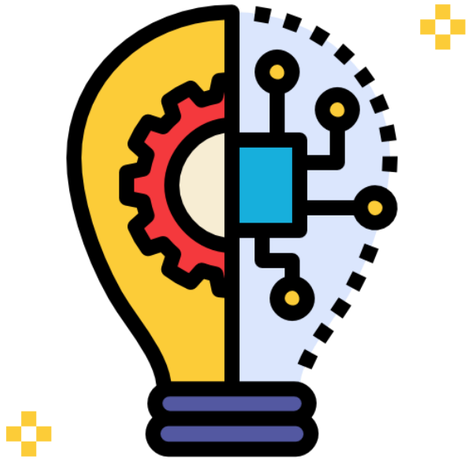
Ethics everywhere



- Related to laws, rights, dignity, civilization
- Unknown impact of progress- powerful technologies

[1] Photo by George Pagan III on Unsplash

Ethics of technology and innovation



- Risks, benefits, impact of new technologies
- Responsible use of technologies
- E.g. Stem cell research, drone warfare
- Data is the game changer

[1] Icon made by geotatah from www.flaticon.com

The data economy- is data the new oil?



- Oil and energy companies most valuable in the 20th century, replaced by data-driven tech companies in the 21st century
- Massive data 90% created in the past 5 years, endless applications

[1] https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data [2] https://www.visceralbusiness.com/latest-news/ten-years-digital-transformation

Data drives (almost) every innovation and application



- Data-driven innovation in the 21st century
- Social media, smart devices, cloud, AI
- Economic growth, societal progress
- Great power, great responsibility- data ethics

[1] https://digital-strategy.ec.europa.eu/en/policies/strategy-data

Data ethics



- Branch of ethics that studies moral problems related to data
- Ethics is wide-ranging
- Decisions, value, impact
- Data ethics focus on processing, sharing, use, application of data

Data ethics foundations

- Privacy and data protection
- Informed consent, individuals rights and agency
- Transparency and trust
- Fairness, non-discrimination
- Accountability and oversight
- Benefit to others and society

[1] Mittelstadt, B. D. (2019). Principles alone cannot guarantee ethical AI. Nature Machine Intelligence, 1(11), 501-507.
[2] Icon made by Freepik from www.flaticon.com

Let's practice!

INTRODUCTION TO DATA ETHICS

Beyond do no harm

INTRODUCTION TO DATA ETHICS



Shalini Kurapati, PhD
Co-founder and CEO, Clearbox AI

datacamp

Legal or ethical?

- Ethics and laws are closely related
- Ethics, go beyond laws:
  - Evolving sense of right and wrong
  - Society, religion, progress, civilization
- Ethics are guidelines, laws are enforced
- Legal does not equate to ethical, ethics go beyond laws



[1] Icon made by Freepik from www.flaticon.com

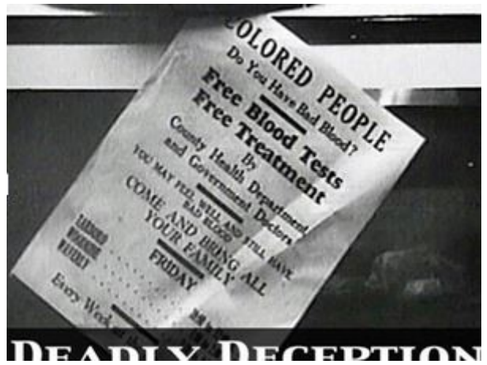
Lessons from the past

HeLa Cells



- Immortal human cells, crucial in advancing cancer research
- From Henrietta Lacks, without her consent from a biopsy
- Privacy, dignity, respect- ethical issues

Tuskegee syphilis study



- Effects of untreated syphilis, 1932-1972
- African American participants, not informed well, no access to treatment
- 100+ died as a result, despite treatment options

Guiding ethical principles



- **Do not harm**- minimize risks for individuals
- **Beneficence**- Maximize benefits for others and society
- **Respect, autonomy, agency**- consent, privacy, information, control, agency
- **Justice**- Legality, fairness and non-discrimination
- IEEE, ACM address data ethics - Responsible use of data, preventing misuse

[1] Icon made by Freepik from www.flaticon.com [2] https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html [3] http://ethics.acm.org/code-of-ethics/ [4]

Data ethics: Old problems new cover

Facebook's emotional contagion study


Latest Newsletters

The Atlantic

TECHNOLOGY

Everything We Know About Facebook's Secret Mood-Manipulation Experiment

It was probably legal. But was it ethical?

- Data scientists skewed newsfeeds of 700,00 users to influence behavior
- Users not informed, nature of research and purpose of data collection

Google's project nightingale


THE WALL STREET JOURNAL.

Subscribe Sign In

English Edition | Print Edition | Video | Audio | Latest Headlines | More

Home World U.S. Politics Economy Business Tech Markets Opinion Books & Arts Real Estate Life & Work Style Sports Search

WSJ NEWS EXCLUSIVE | TECH

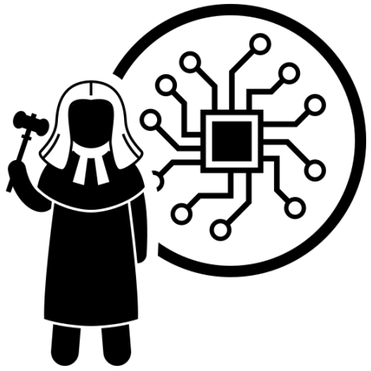
Google's 'Project Nightingale' Gathers Personal Health Data on Millions of Americans

Search giant is amassing health records from Ascension facilities in 21 states; patients not yet informed

- Google handling of sensitive data
- Large scale medical data without patient consent or doctor's knowledge

[1] https://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/ https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7059004/

Encoded data ethics in laws

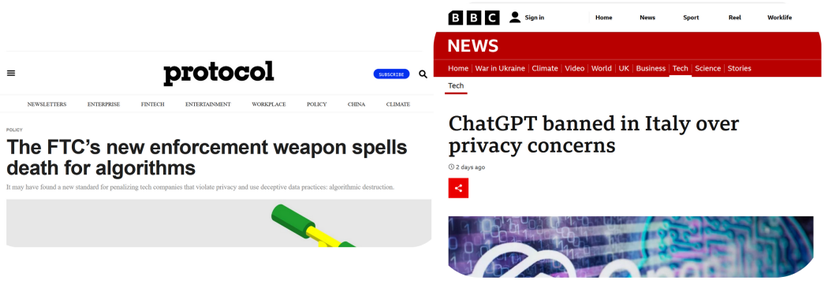


- Data protection laws like GDPR, CCPA embed several data ethics principles
- Privacy, data security, transparency, informed consent, non-discrimination, accountability
- Not mere guidance, but enforced
  - GDPR fines
  - EUR 20 million or 4% of the worldwide turnover

[1] Icon made by Leremy from www.flaticon.com

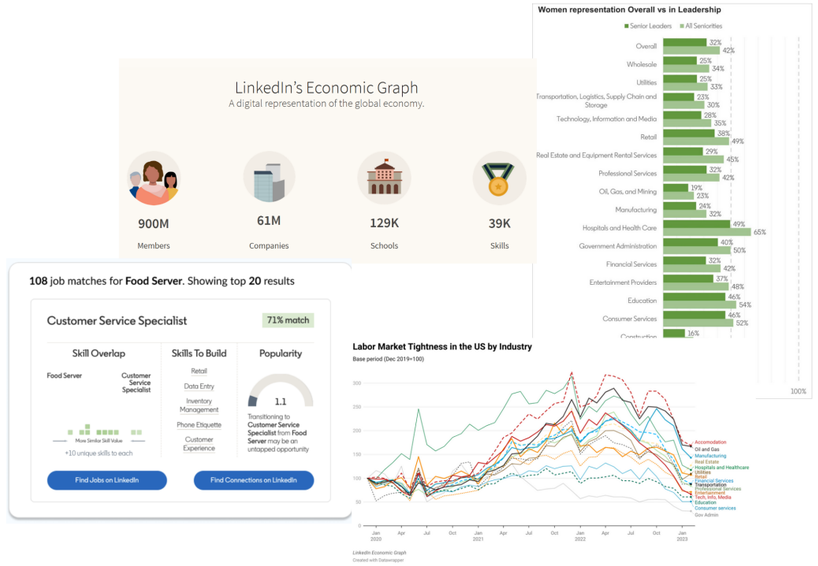
The business case for data ethics

Compliance and reputation



- Regulators move in for lax data practices- FTC, Italian DPA
- Following high data ethics standards boosts compliance
- Attractive to clients and employees

Responsible innovation



- LinkedIn Economic Graph in collaboration with governments, policy makers
- Benefits clients, society and the company

[1] https://economicgraph.linkedin.com/

Let's practice!

INTRODUCTION TO DATA ETHICS

Data ethics and AI ethics

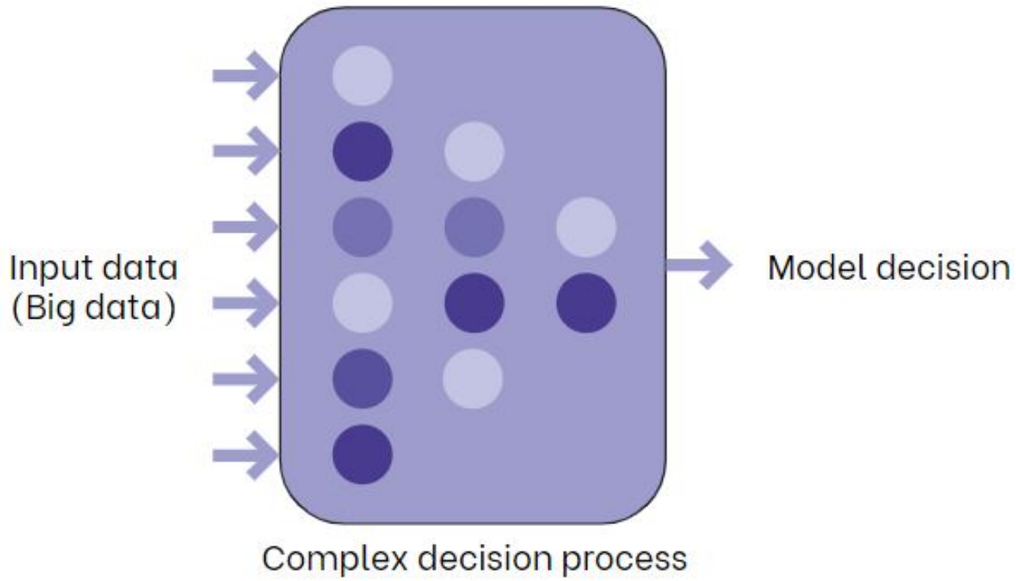
INTRODUCTION TO DATA ETHICS



Shalini Kurapati, PhD
Co-founder and CEO, Clearbox AI

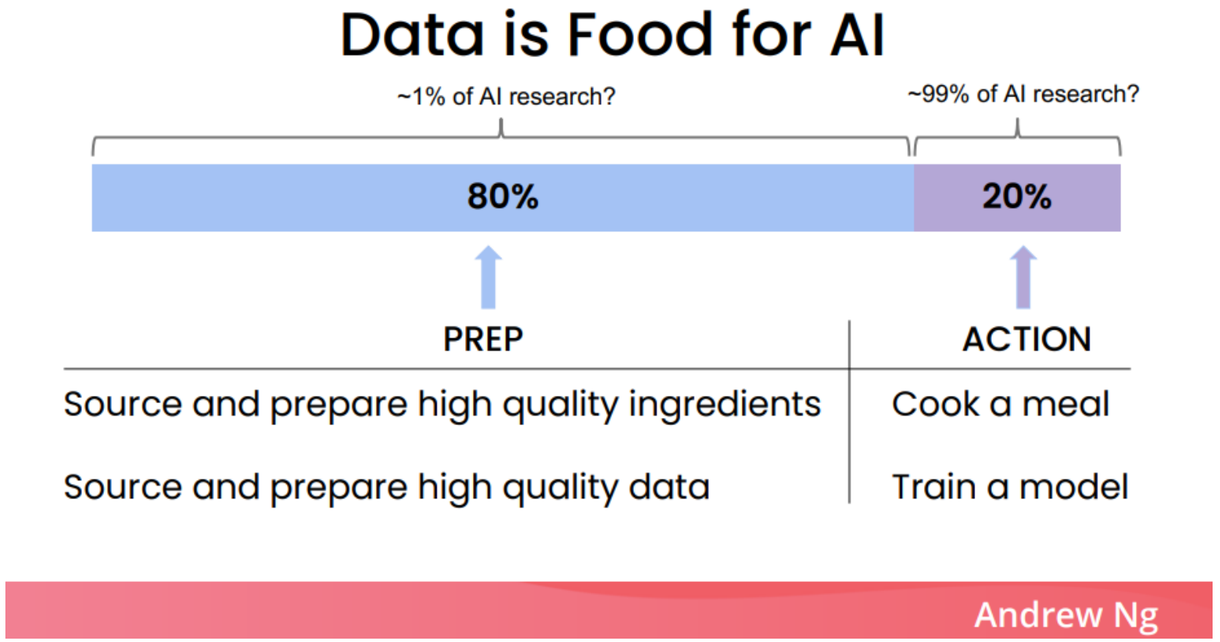
datacamp

What is AI?



- AI combines algorithms and robust dataset to solve problems
- Powerful AI uses large datasets
- Massive potential, double-edged sword, ethical considerations

Data is fuel for AI



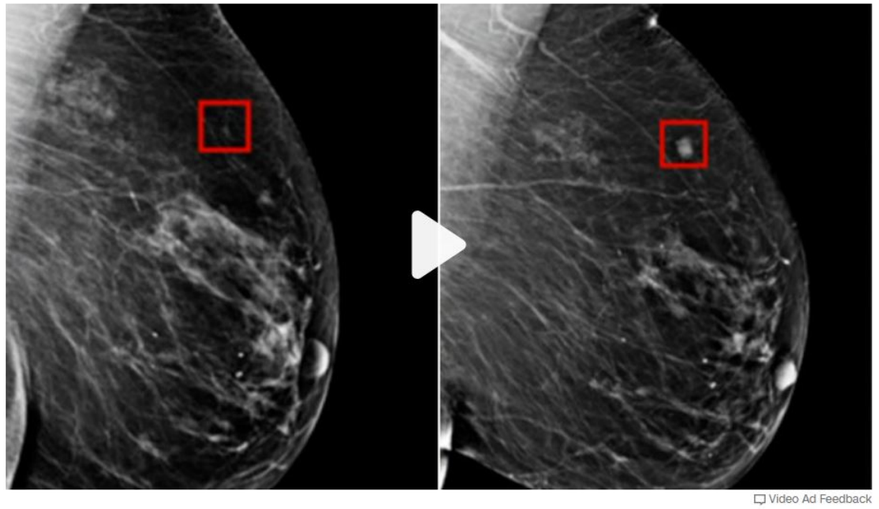
- AI models need massive amounts of data, most of AI efforts spent on data
- Poor data leads to poor AI outcomes
- AI ethics wide-ranging than data ethics, important to know the similarities and differences

[1] Photo source: Andrew Ng's webinar, MLOps: From Model-centric to Data-centric AI

The good and the amazing



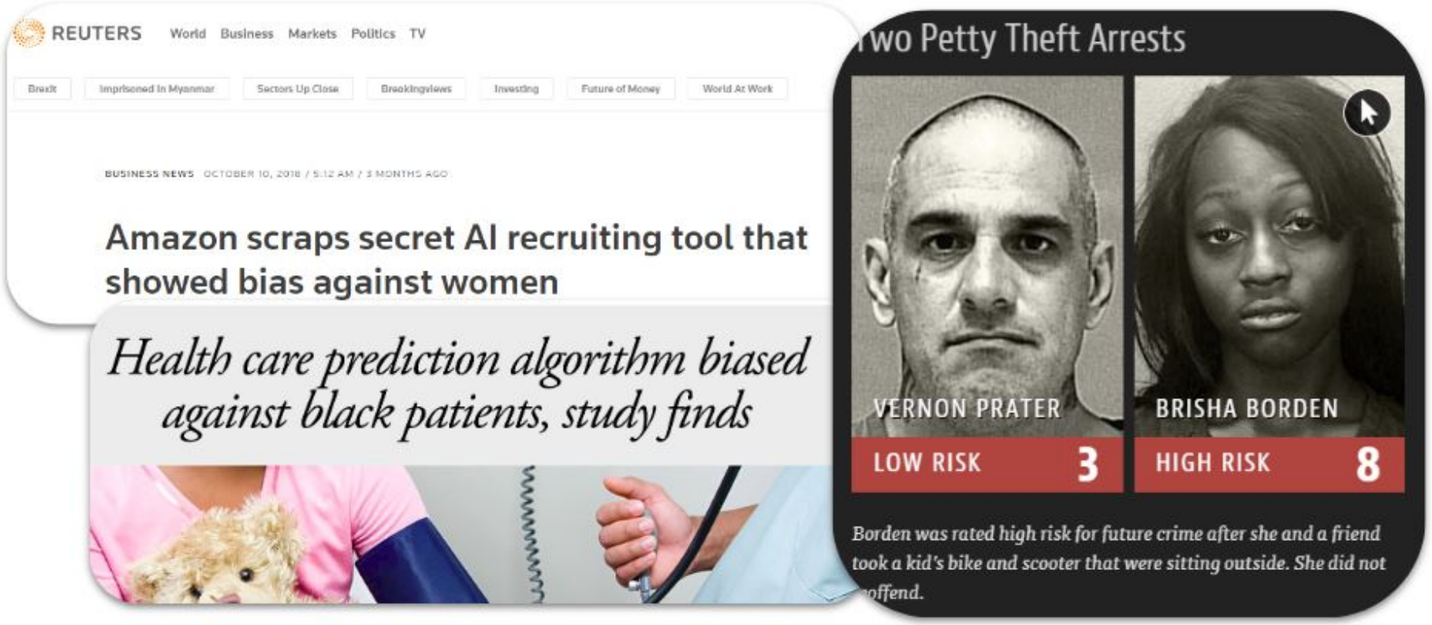
- Voice assistants, conversational AI
- Movie recommendations



Images show AI detecting breast cancer 4 years before it developed

- AI for good
- Early cancer diagnosis
- Precision agriculture- reduced water & pesticide use

The bad and the ugly

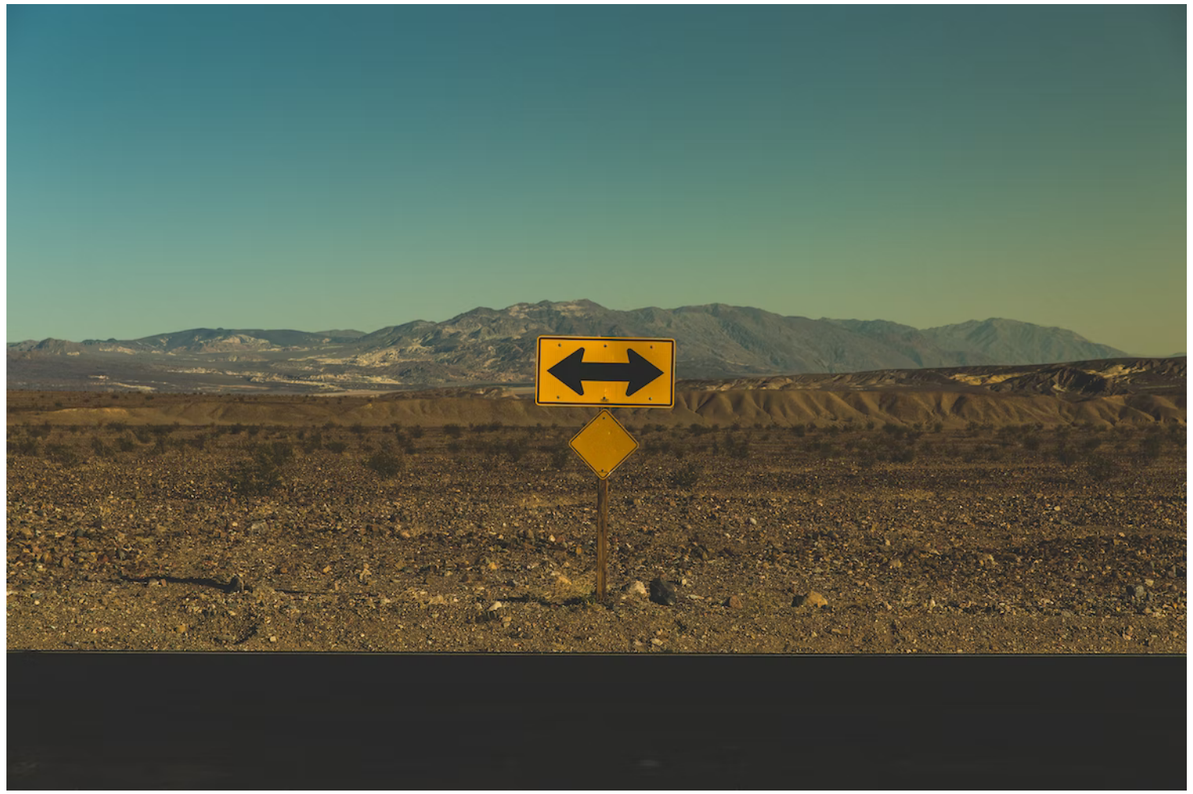


- Discrimination in hiring, access to healthcare, and fair trials in court
- Many reasons- biased datasets, complex and blackbox AI models, lack of explanations
- AI ethics go beyond data ethics

AI ethics dilemmas

AI ethics go beyond data ethics

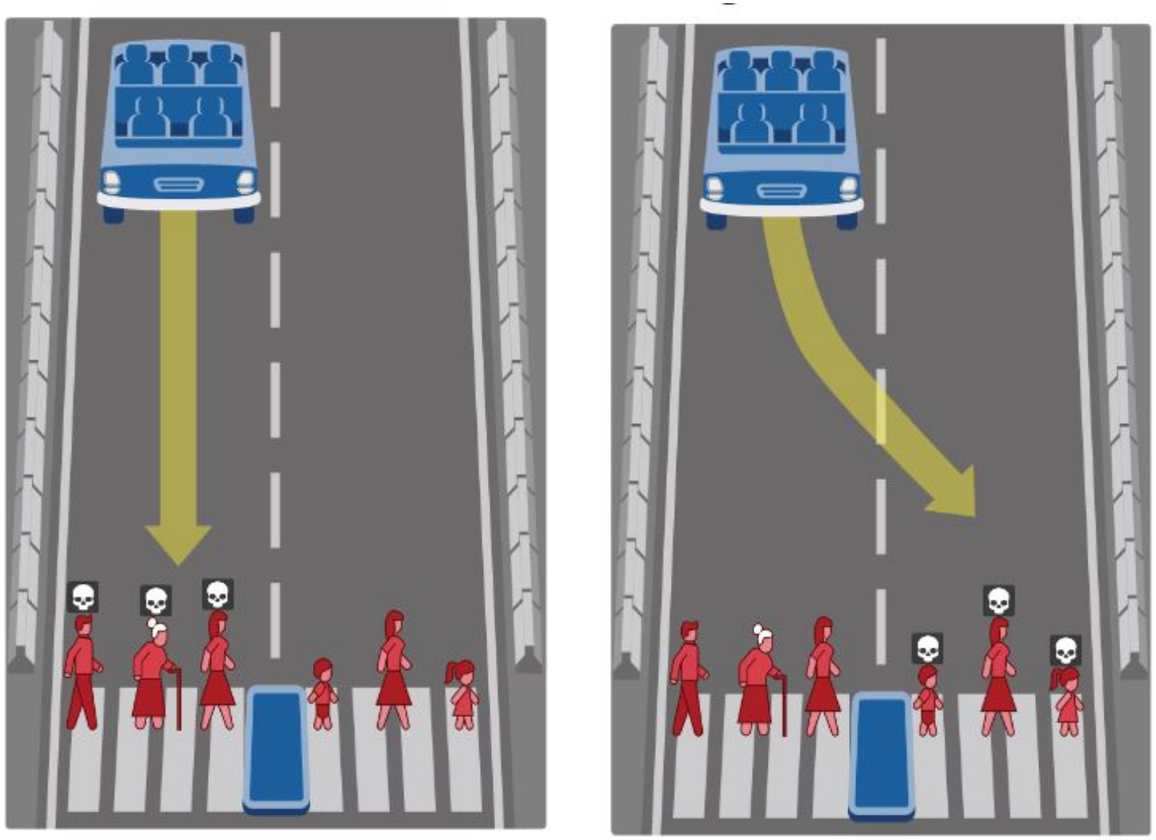
- Consequences if AI decision making:
  - Credit access/loans
  - Hiring
- AI discrimination
- AI/Robots taking over jobs
- Will robots/drones claim fewer human lives in a war?
- What does it mean to be human?
- Can we encode morality into AI?



[1] Photo source: unsplash.com

What should an AI do?

- Selfing driving car with sudden brake failure
- Continue ahead or swerve
- Continue will kill a woman, man and an older woman
- Swerve will kill a young woman and two children
- Who is responsible?



[1] https://moralmachine.mit.edu/

AI ethics guidelines in a nutshell

Many organizations and bodies working on AI ethics guidelines

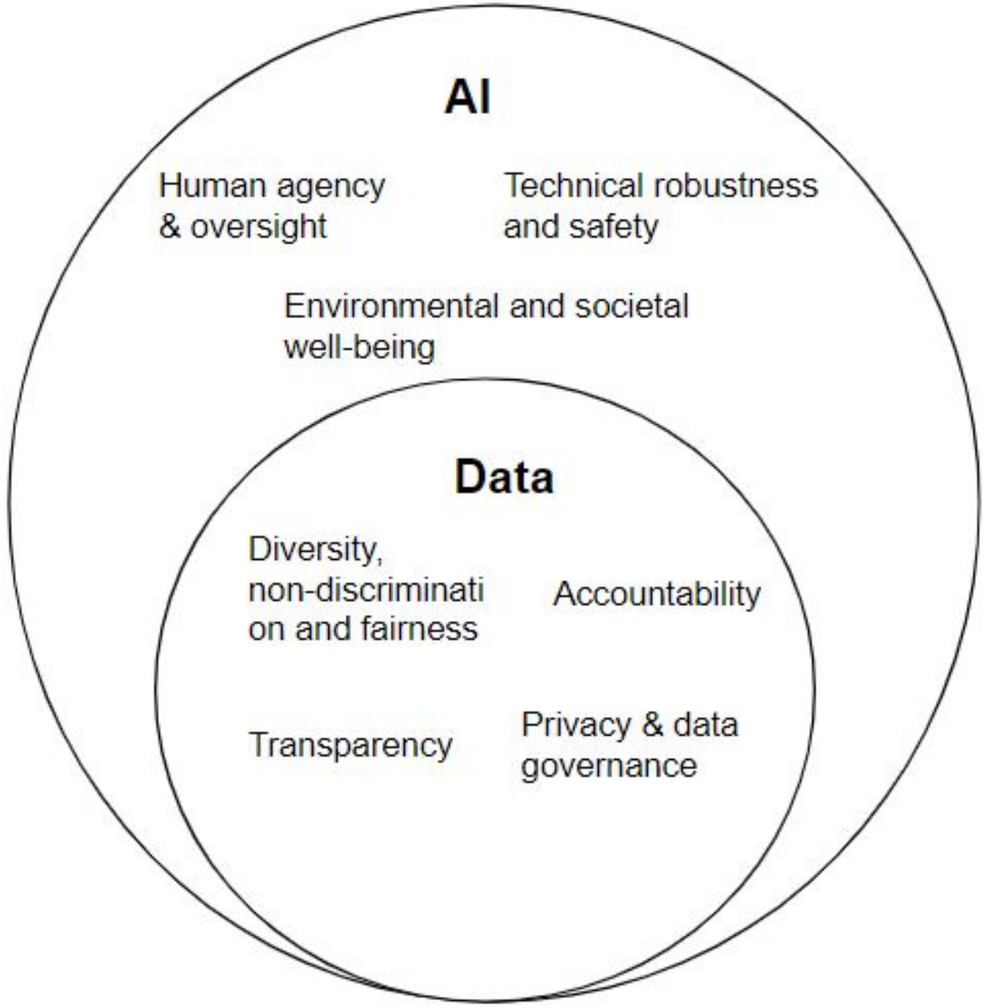
- Human agency and oversight
- Technical robustness and safety
- Privacy and data governance
- Transparency
- Diversity, non-discrimination and fairness
- Environmental and societal well-being
- Accountability



[1] https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment [2] https://inventory.algorithmwatch.org/

The data overlap

- Privacy, security, data protection
- Bias and fairness
- Transparency
- Accountability



Let's practice!

INTRODUCTION TO DATA ETHICS