(Mis)Handling Data!

AI Ethics KTH
Kostis S-Z

Data in A.I.

- Using ready packaged algorithms on datasets becomes all the more easier.
- Handling personal data comes with a responsibility.
- Collaborating with field experts (medical experts, economists etc)
- A crucial and difficult task:
    - Interpreting results
    - Understanding the limitations of data

-> Being aware of limitations and possible dangers is vital!

Dangers when (mis)handling Data

- Knowingly mishandling data
  - Collecting sensitive information without users' consent
  - Personally Identifiable Information (PII) include: face images, home address, ID / driver's / Passport number, fingerprint, health information
  - Too many scandals… (Facebook, Google …[1, 2])
- Unknowingly mishandling data
  - Building technologies on incomplete, biased, corrupted data

A) Collecting data

Developing a new technology / model:

- Enhance integration, user feedback, personalized optimization, ...

Some examples:

- 5G networks, Self-driving cars, smart homes etc
- Medical diagnosis
- Personalised Content Marketing

Approach? Try not to collect PII? De-Identify PII (k-anonymity [3])?

Even when securely storing sensitive information, things can go wrong [4]

B) Biased Data

"Any time you have a dataset of human decisions, it includes bias..."

- *Roman Yampolskiy, Cybersecurity Lab at the University of Louisville*

A general problem when developing technologies *(e.g: seatbelts [5])*

- Gender
- Race
- Societal / Financial status
- Cultural etc

B) Biased Data

Dangers of wrong models

- Miscorrelation (Confusing human characteristics - Crime Prevention [6])
- Promoting stereotypes (Advertising better jobs to men- Linkedin [7])
- Overfitting (Training mostly with white people photos - Google [8])
- Wrong diagnosis from Medical Data (Training on subsamples for disease diagnosis that might have race specific preconditions [9])
- Many, many more...

Sources:

[1] Facebook Cambridge Analytica Scandal

[2] Google's Secret Microphone (and many more)

[3] De-Identifying PII (K-anonymity)

[4] Health Net Data Breach (but also so many more data breaches, leaks, hacks etc)

[5] Building Cars for Men (thanks Bas!)

[6] China using AI for Crime Prevention, Risks of AI Crime Detection

[7] Linkedin promoting high-paying jobs more often to men than women

[8] Google facial recognition AI classified black men as gorillas

[9] Disease risks vary amongst races

Inspiration for Biased Data Discussion