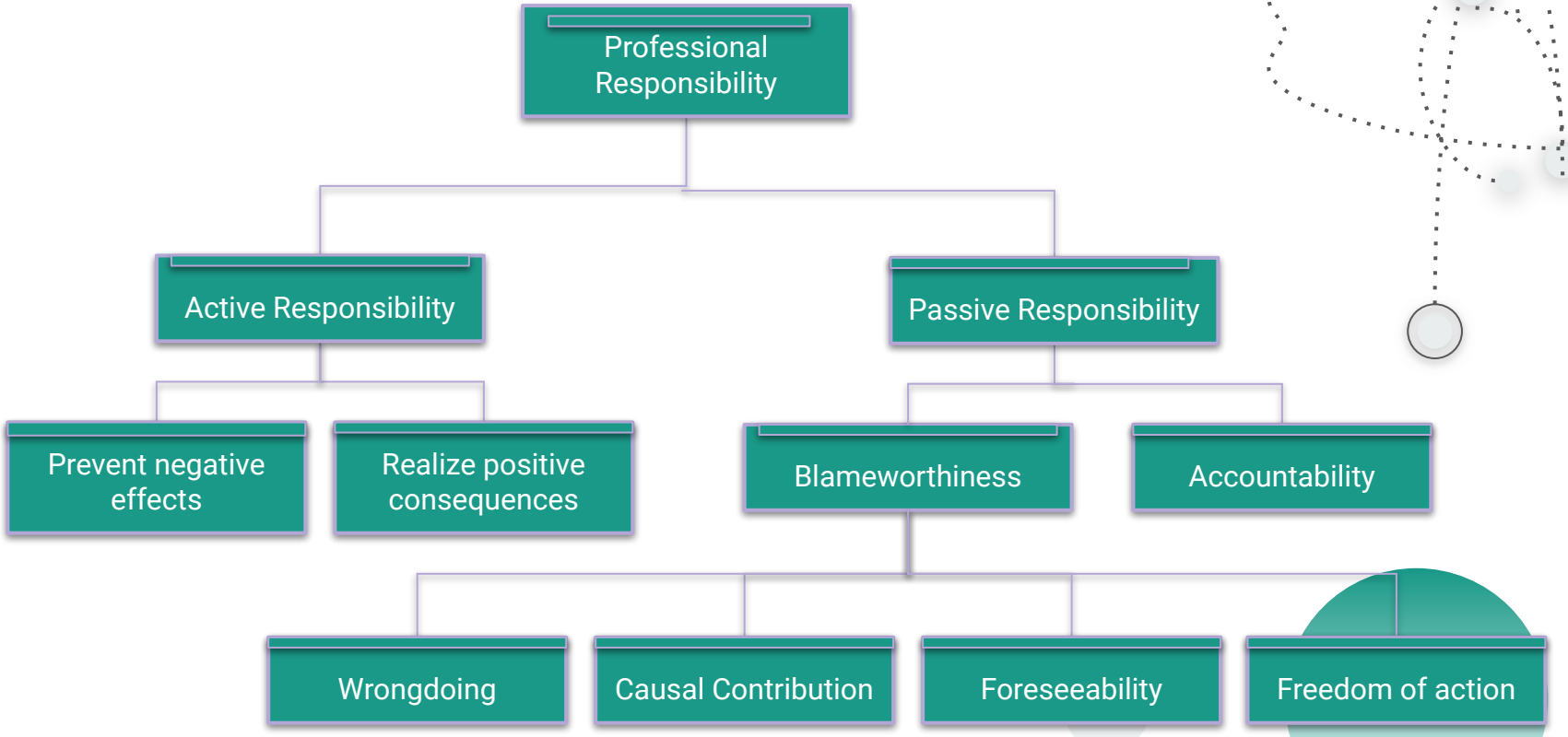
Ethics and AI

Lecture 13

Responsibility Tree



Blameworthiness

Wrongdoing

To blame a person or an institution, it is necessary that they did something wrong.

Example: Developers did not test enough

Foreseeability

To be responsible we must be able to know the consequences of actions.

Example: Developers must know consequences beforehand

Causal Contribution

Causal contribution refers not only to action but also to a failure to act.

Example: Engineers did not stop company to commercialize a faulty machine

Freedom of Action

They must have freedom of action, so not acting under compulsion.

Example: Developers must have freedom to voice opinions

01

Autonomous Vehicles

Autonomous Vehicles

- Operative Autonomy: Can perform tasks without human supervision.

- 5 levels of autonomy:
  - L1: Control basic functions i.e., acceleration
  - L5: Control everything

- For unavoidable collisions, they need to make moral decisions. For this, there is ethical frameworks.
  - For example, if a car has to crash with any of 2 bicycles where one is wearing a helmet. (utilitarianism)

- These analysis needs to be concrete policies. (German Ministry of Transport and Infrastructures has a code of conduct with 20 guidelines)

02

Digital Medicine

Digital Health and Medicine Applications

01

Health Monitoring

02

Diagnosis Support

03

Predictive Algorithms for treatment

04

Public Health

Digital Medicine

- AI Diagnosis has ethical issues related to **privacy, data protection, transparency, and representativeness** of training data.
- Biomedical data are **mostly representative of male Caucasian adults** which is a violation of ethical standards.
- The **transparency issue** raises questions on medical responsibility and informed consent since the **algorithm cannot be fully explained**.
- The adoption of AI tools in medicine should be seen as **supporting** traditional clinical practices rather than **substituting** them.

03

Cybersecurity

Cybersecurity

- **Security and ethics are interconnected**, as unauthorized access to systems can lead to **breaches of privacy and proprietary rights**.
- **AI predictive security measures** shape computing environments, influencing order.
- **Ethical dilemmas can arise in ethical hacking**, where hackers try to break into computer systems with a pro-social goal, such as testing security.
- Ethics in Cybersecurity tries to balance between prevention of harm and respect for human autonomy

Hackers' arguments defending hacking

01

Information should be free

02

Hackers demonstrate the faults in the system

03

They do not change anything

04

Keeps Big Brother at bay

04

Moralizing Technologies

If we want people to pay a ticket to use the subway, we can educate them to buy a ticket before entering it, but we can also design metro barriers that tell people: “Buy a ticket before you enter the subway”.

Moralization of Technologies:
Deliberate development of technologies in order to shape moral action and moral decision-making.

(Active Responsibility)

Factors affecting moralization

Uncertainty

AI tech is called "experimental technology" as it lacks operational experience to assess its social benefits and risks, caused by the uncertainty around its introduction and interaction with humans.

Invisibility

Most of the time and under most conditions, computer operations are invisible.
- Invisibility of abuse
- Invisibility of programming values
- Invisibility of complex calculations

THANKS!