Risks, Harms, Moral Responsibility and Technology

- **Agent:** The entity that performs the action and causes something to happen
- **Patient:** The entity that is affected by the action
- **Moral responsibility** deals with the link between the agent and the patient
- This link is less clear when both humans and technology interact and affect each other

- When assigning moral responsibility to some person or group, what are important characteristics we should look for in their actions and the outcomes?
  - **Causality:** a casual connection between the person/group and the outcome of the actions
  - **Knowledge:** i.e. of the possible consequences of their actions
  - **Free will:** i.e. the ability to freely choose how to act
  - **Economics(?)**: is the outcome constrained / influenced by money
  - **Constraints:** e.g. of time

- What methods are available to us to hold people and groups accountable for their actions? How effective are they?
  - **Law/lawsuits:** effectiveness is variable
  - **Regulatory bodies:** could be effective if rules are enforced
  - **Public relations / public pressure**

- How do we deal with the many hands problem?
o Divide responsibility up by modules (hierarchy of responsibilities). Drawback: encourage pointing fingers to other parties
o Holding the entire group responsible: pros (quality control encouraged for the whole group), cons (fairness?)
 o Whistleblowers: Mechanism to allow reporting of violations, protect whistleblowers

• How can we deal with technologies that can make it hard to understand or consider the outcomes?
  o Software testing
  o Proactive considerations of outcomes
  o Being proactive about updating the technology
  o Corporate culture
  o Standards, best practices

• What about cases where we make decisions based on the outputs of technologies that we don’t fully understand?
  o Authority-in-the-loop requirement
  o Explain it (GDPR)

Therac-25
• Radiation machine killed several people in the 80s by delivering a severe radiation overdose, mainly due to software problems
• Who was morally responsible?
  o Software developer? Assembly code was buggy but some the code was legacy code from a previous version
  o Testers? Were they testing under the right conditions
  o Management? Should do a recall
Medical staff / doctors? Do they understand the risks, operation
Government? Did they have a law in place to prevent this?
Designers? User interface could prevent issues

- What actions should have been taken to hold them accountable for that harm?
  - Government oversight
  - Quality assurance
  - Firings? Criminal charges due to negligence?
  - Civil lawsuits

- What lessons can we take away from Therac-25?
  - Overconfidence in software
  - Reliability is not the same as safety
  - Lack of defensive design
  - Unrealistic risk assessments
  - Inadequate investigations/followups
  - Inadequate software and system engineering practices
  - Software reuse
  - Safe vs “Friendly” user interfaces
  - User and govt oversight and standards