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. Circumstances for ascribing moral responsibility not always clear.
- This link is also 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 causal connection between the person/group and the outcome of their actions
  - **Knowledge**: i.e. of the possible consequences of their actions
  - **Free will**: i.e. the ability to freely choose how to act
  - Others?

- What methods are available to us to hold people and groups accountable for their actions? How effective are these methods, and what are their drawbacks?

- How do we deal with the **many hands problem**, i.e. the fact that it is difficult to determine who was responsible for what when many individual people contributed to the outcomes?
• How can we deal with the fact that technologies can make it hard to understand or consider the outcomes of our actions?
  ○ E.g. an individual software developer may not be able to foresee errors that might occur when their software is integrated into a larger system.

• What about cases where we make decisions based on the outputs of technologies we don't fully understand?

 Therac-25

• The Therac-25 was a radiation machine that killed several people in the 1980’s by delivering a severe radiation overdose, mainly due to problems with the machine’s software.
• Who is morally responsible for the harm done in the Therac-25 case? AECL, the company that produces Therac-25, did not consider software design during its assessment of risk and reliability.

• What actions should have been taken to hold them accountable for that harm?

• What lessons can we take from the case of the Therac-25?
  ○ Overconfidence in software
  ○ Reliability is not the same as safety
  ○ Lack of defensive design
  ○ Unrealistic risk assessments
  ○ Inadequate investigations/follow-ups?
  ○ Inadequate software and system engineering practices
  ○ Software reuse
  ○ Safe versus “friendly” user interfaces
  ○ User and government oversight and standards