Ethics + AI

4modIA 2022-2023

English class

Becky Coles coles@insa-toulouse.fr
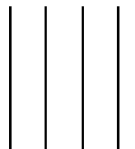
Definitions (Source: Cambridge Dictionary)

Ethics: *The study of what is morally right and wrong, or a set of beliefs about what is morally right and wrong.*

Ethical issue: *A situation where there are competing alternatives and the right thing to do is not obvious or clear. Sometimes terms such as good/bad and better/worse are used to consider the effect of particular actions on our lives, society, nature and the environment.*

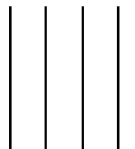Ethical behaviour: *Acting according to accepted standards of conduct*

Ethics of AI

= the branch of the ethics of technology specific to artificially intelligent systems (Muller, 2020)

Concerns the moral behavior of *humans* as they design, make, use and treat artificially intelligent systems, and the behavior of *machines,* in machine ethics.
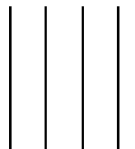
Ethics vs Morals (Source: Merriam Webster Dictionary)

Morals: one's particular values concerning what is right and what is wrong, includes an element of subjective preference

It would go against my morals to help you cheat on the test.

Ethics can refer broadly to moral principles, but often applied to question whether or not an action is responsible within a relatively narrow area of activity
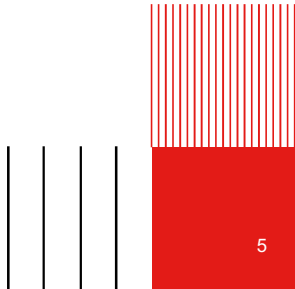
Our class had a debate over the ethics of genetic testing.

Question?

- Name a current application of AI in society?

- Who benefits from its implementation?

- What are the associated risks and/or injustices associated with its use?

- ….

Use AI ethical principles (or value-based criteria) when discussing the + / - consequences of using an AI solution.

List of guiding principles (adapted from Jobin et al. (2020) and the WEF (2019):

Transparency

Justice and fairness

Non-maleficence

Accountability

Privacy

Beneficence

Human autonomy

Compliance
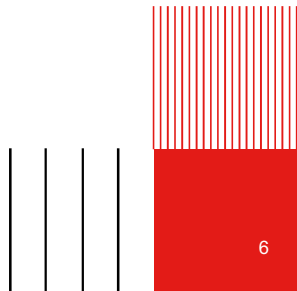
Sustainability

Safety

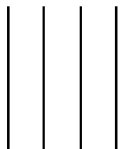Task 1: Worksheet: AI Ethical Principles

INSA Toulouse

Survey

A survey* of 602 technology innovators, developers, business and policy leaders, researchers and activists responded to this specific question:

*By 2030, will most of the AI systems being used by organizations of all sorts employ ethical** principles focused primarily on the public good?*

68% said they will not

* Pew Research Center and Elon University 2021, non-random sample

**"ethical" = adopting AI in a manner that is transparent, responsible and accountable.
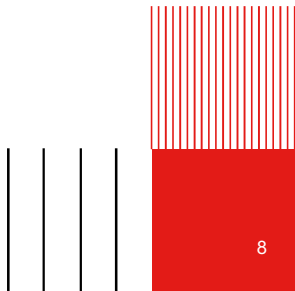
AI case studies

- Facial recognition software at airports

- AI and medicine (e.g. to detect cancer tumours)

- Lethal autonomous weapon systems

- AI Recruitment Tools

- Chatbots

- Self driving cars

Working in groups of 3, watch the video, read the articles and complete the analysis sheet ready for a mini-presentation

- Start your mini-presentation with a hook

- State the AI and its aim

- What are the main advantages/disadvantages of this AI, in the context of the AI guiding principles

- Conclude - how to address these ethical concerns?