



RESPONSIBLE ARTIFICIAL INTELLIGENCE

Virginia Dignum

Chair of Social and Ethical AI - Department of Computer Science

Email: virginia@cs.umu.se - Twitter: [@vdignum](https://twitter.com/vdignum)



UMEÅ UNIVERSITY

RESPONSIBLE AI: WHY CARE?

- AI systems act autonomously in our world
- Eventually, AI systems will make *better* decisions than humans

AI is designed, is an artefact

- We need to sure that the **purpose** put into the machine is the purpose which **we really want**

Norbert Wiener, 1960 (Stuart Russell)

King Midas, c540 BCE



TAKING RESPONSIBILITY

- **Ethics in Design**
 - Ensuring that development processes take into account ethical and societal implications of AI as it integrates and replaces traditional systems and social structures
- **Ethics by Design**
 - Integration of ethical reasoning abilities as part of the behaviour of artificial autonomous systems
- **Ethics for Design(ers)**
 - Research integrity of researchers and manufacturers, and certification mechanisms



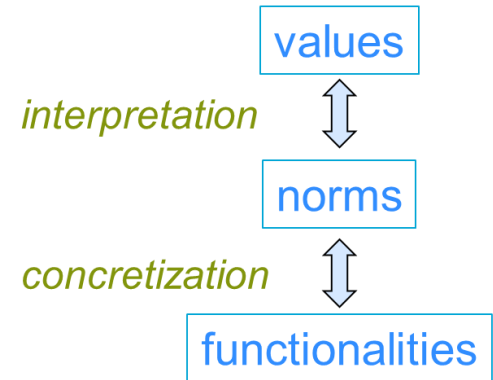
ETHICS IN DESIGN: AI – DOING IT RIGHT

- Principles for Responsible AI = ART
 - Accountability
 - Responsibility
 - Transparency



ETHICS IN DESIGN: AI – DOING IT RIGHT

- Principles for Responsible AI = ART
 - Accountability
 - Explanation and justification
 - Design for values
 - Responsibility
 - Transparency



ETHICS IN DESIGN: AI – DOING IT RIGHT

- Principles for Responsible AI = ART
 - Accountability
 - Explanation and justification
 - Design for values
 - **R**esponsibility
 - Autonomy
 - Chain of responsible actors
 - Human-like AI
 - Transparency



ETHICS IN DESIGN: AI – DOING IT RIGHT

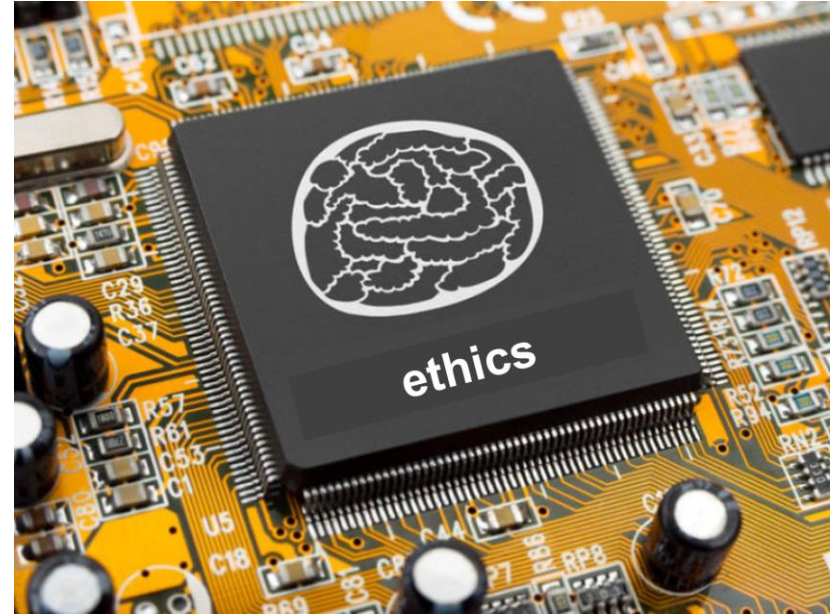
- Principles for Responsible AI = ART
 - Accountability
 - Explanation and justification
 - Design for values
 - Responsibility
 - Autonomy
 - Chain of responsible actors
 - Human-like AI
 - Transparency
 - Data and processes
 - Algorithms
 - Choices and decisions



ETHICS BY DESIGN

- Can AI artefacts be build to be ethical?
- What does that mean?
- What is needed?

- Understanding ethics
- Using ethics
- Being ethical



ETHICAL REASONING?

- Value: “human life”
- Reasoning:
- Utilitarian car
 - The best for most; results matter
 - **maximize lives**
- Kantian car
 - Do no harm
 - **do not take explicit action if that action causes harm**
- Aristotelian car
 - Pure motives; motives matter
 - **Harm the least; spare the least advantaged (pedestrians?)**

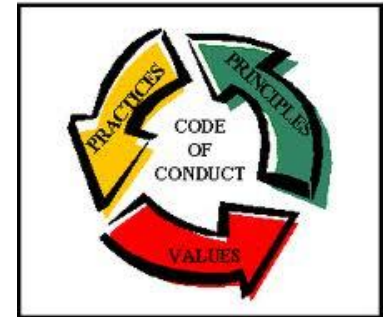
Ethical theories

- Many different theories, each emphasizing different points
 - Utilitarian, Kantian, Virtues....
- Highly abstract
- None provide ways to resolve conflicts
- Deontology and Virtue Ethics focus on the individual decision makers while Teleology considers on all affected parties.



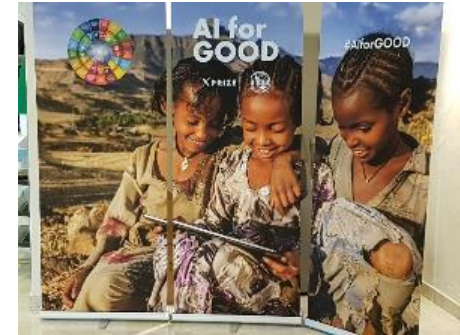
ETHICS FOR DESIGN(ERS) – REGULATION, CONDUCT

- A code of conduct clarifies mission, values and principles, linking them with standards and regulations
 - Compliance
 - Risk mitigation
 - Marketing
- Many professional groups have regulations
 - Architects
 - Medicine / Pharmacy
 - Accountants
 - Military
- Is what happens when society relies on you!



AI FOR GOOD, AI FOR PEOPLE,...

- Harness the positive potential outcomes of AI in society, the economy
- Ensure inclusion, diversity, universal benefits
- Prioritize UN2020 Sustainable Development Goals
- The objective of the AI system is to maximize the realization of human values



European Commission > Strategy > Digital Single Market > Policies >

Digital Single Market

POLICY

High-Level Expert Group on Artificial Intelligence

<https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>



<https://ethicsinaction.ieee.org/>



UMEÅ UNIVERSITY



<http://www.ai4people.eu>

TAKE AWAY MESSAGE

- AI systems are artefacts built by us for our own purposes
 - Our decision, our responsibility
- AI influences and is influenced by our social systems
 - Design is never value-neutral
 - Society shapes and is shaped by design
- Knowing ethics is not being ethical
 - Not for us and not for machines
 - Different ethics – different decisions
- Artificial Intelligence needs ART
 - Accountability, Responsibility, Transparency
 - Be explicit!



RESPONSIBLE ARTIFICIAL INTELLIGENCE

WE ALL ARE RESPONSIBLE

