RESPONSIBLE ARTIFICIAL INTELLIGENCE

Prof. Dr. Virginia Dignum
Chair of Social and Ethical AI - Department of Computer Science
Email: virginia@cs.umu.se - Twitter: @vdignum
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

• **in** Design
  o Ensuring that development **processes** take into account ethical and societal implications of AI as it integrates and replaces traditional systems and social structures

• **by** Design
  o Integration of ethical reasoning abilities as part of the **behaviour** of artificial autonomous systems

• **for** Design(ers)
  o Research integrity of **researchers** and manufacturers, and certification mechanisms
ETHICS IN DESIGN

• Doing the right thing
• Doing it right
• Design for values
• Design for all

“Do things right, and do the right things.”

PETER DRUCKER
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
  o Accountability
    ▪ Explanation and justification
    ▪ Design for values
  o Responsibility
  o Transparency

• Optimal AI is explainable AI
• Many options, not one ‘right’ choice
ETHICS IN DESIGN: AI – DOING IT RIGHT

• Principles for Responsible AI = ART
  o Accountability
    ▪ Explanation and justification
    ▪ Design for values
  o Responsibility
    ▪ Autonomy
    ▪ Chain of responsible actors
    ▪ Human-like AI
  o Transparency
ETHICS IN DESIGN: AI – DOING IT RIGHT

• Principles for Responsible AI = ART
  o Accountability
    ▪ Explanation and justification
    ▪ Design for values
  o Responsibility
    ▪ Autonomy
    ▪ Chain of responsible actors
    ▪ Human-like AI
  o Transparency
    ▪ Data and processes
    ▪ Algorithms
    ▪ Choices and decisions
ETHICS BY DESIGN – ETHICAL ARTIFICIAL AGENTS

• Can AI artefacts be built ethically?
  • What does that mean?
  • What is needed?

• Understanding ethics
• Using ethics
• Being ethical
DECISIONS MATTER!

values

interpretation

norms

concretization

functionalities

fairness

Equal resources

Equal opportunity

...
ETHICS FOR DESIGN(ERS)

• Regulation
• Certification
• Standards
• Conduct
• Ethical Guidelines
  o Guiding principles
    ▪ Respecting Fundamental Rights, Principles and Values - Ethical Purpose
    ▪ Critical concerns
  o Implementation
    ▪ Realising trustworthy AI
    ▪ Assessing Trustworthy AI

• Investment and policy strategy
  o Using AI to build an impact in Europe
    ▪ Transforming Europe's Business landscape
    ▪ Catalyzing Europe's Public Sector
    ▪ Attaining World-Class Research Capabilities
    ▪ Accomplishing Citizen's Benefits and Engagement
  o Leveraging Europe's enablers of AI
    ▪ Attracting Funding and Investments in AI
    ▪ Enabling AI with Data and Physical Infrastructure
    ▪ Generating appropriate Skills and Education for AI
    ▪ Ensuring an appropriate policy and regulatory framework
MANY MORE (AND COUNTING...)

- Initiatives
  - IEEE Ethically Aligned Design: https://ethicsinaction.ieee.org/
  - CLAIRE (and ELLIS): https://claire-ai.org/
    - Confederation of Laboratories for Artificial Intelligence Research in Europe
  - AI4EU: on demand platform
  - ALLAI (NL)

- Strategies / positions
  - Council of Europe
  - OECD
    - ...

- Declarations
  - Asilomar
  - Montreal
    - ...

UMEÅ UNIVERSITY
TAKE AWAY MESSAGE

- AI influences and is influenced by our social systems
- Design in never value-neutral
- Openness and explicitness are key!
  - Accountability, Responsibility, Transparency
- Optimal AI is explainable AI
- AI systems are artefacts built by us for our own purposes
- We set the limits
RESPONSIBLE ARTIFICIAL INTELLIGENCE

WE ARE RESPONSIBLE

Email: virginia@cs.umu.se
Twitter: @vdignum