Shaping the ethical dimensions of information technologies – a European perspective (SHERPA) - Recommendations

AI, Ethics and Human Rights - Designing a Better World

Prof. Bernd Carsten STAHL, De Montfort University, Leicester, UK
26th National Ethics Council Forum, 19.11.2020
This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no. 786641

Purposes of AI

Efficiency, optimisation, profit maximisation

Social control

Human flourishing
AI Definitions

- AI is a collection of technologies that combine data, algorithms and computing power. (European Commission, White Paper)

- AI combines a variety of intelligent behaviors and various kinds of mental labor, known as mental activities, ... [to] include perception, memory, emotion, judgement, reasoning, proving, identification, understanding, communication, designing, thinking and learning, etc. (Li and Du 2007)
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 786641.
Concepts of AI in use

General AI

Nature of intelligence
Human nature
Nature of knowledge / reality

Concepts of AI

Machine learning (narrow AI)

Opacity
Unpredictability
Big data requirements

Converging socio-technical systems

Social impact
Manipulation
Systems autonomy

Nature of intelligence
Human nature
Nature of knowledge / reality
Classification of issues

Superintelligence
Singularity
Change of human nature
Philosophy / theology

Metaphysical questions

Ethics of AI

Specific issues of machine learning
- Bias
- Discrimination
- Security
- Data protection

General questions about living in a digital world
- Autonomy
- (Economic) Power
- Warfare
- Distribution of benefits
- Calls for policy and political action

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no. 786641
This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no. 786641

- Ethics frameworks
- Professionalism
- Methodologies
- ... (technical)
- Security
- Data protection
- Standardisation
- Design guidelines
- ... (organisational)
- Specific regulation
- Human rights law
- Competition/liability
- New regulator
- ... (policy and regulatory)
- Risk management
- Impact assessment
- Quality assurance
- Data governance
- CSR
- ...
This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no. 786641.
Insights so far

• There is no magic bullet
  – epistemic complexity
  – distribution of responsibility
  – technical progress
• What is required is an intelligent mix of options
• The metaphor of an ecosystem is appropriate.
Towards an Ecosystem of Artificial Intelligence for Human Flourishing
Characteristics of Ecosystems

- Complex relationships between actors
- Openness
- Evolution as driver of change
- Co-evolution and mutual learning
- Interdependence of actors
- Unclear boundaries
This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no. 786641

Requirements for interventions

Delimitation of the ecosystem

Knowledge base

Ecosystems governance

- Geographical
- Conceptual
- Technical

- Technical
- Conceptual
- Procedural
- Capacity building

- Adaptive governance
- Flexibility
- Ability to learn
- Meta-responsibility

Adaptive governance

Flexibility

Ability to learn

Meta-responsibility

Capacity building

Technical

Procedural

Conceptual

Geographical

Delimitation of the ecosystem

Knowledge base

Ecosystems governance
Recommendations

Measures to support AI ecosystems to promote human flourishing
This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no. 786641
The SHERPA Vision

Ethics as integral to excellence in AI

A world where excellent AI and smart information systems support human flourishing by

- integrating ethical concerns by design and
- safeguarding human rights
Thank You!

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no. 786641