

Fragile Humans, Powerful Machines

Digital Humanism for the AI Age

Erich Prem, eutema

The Digital Humanism manifesto, with its diagnosis from 2019, is still largely valid today. This is surprising given that it was prepared before the advent of recent generative AI systems that have significantly changed how people work, who they ask for advice, and also how they see themselves. This talk will thus revisit Digital Humanism from the perspective of AI. It will investigate what is sometimes construed as a gap between fragile humans and powerful machines from the perspective of ethics. Diving into various ethical theories we investigate the relation of ethics and digital humanism and its use for constructively approaching a digital future worth living.

Living in the post-digital infosphere

Photo optimizer

Music recommender

Health monitoring

Speech recognition

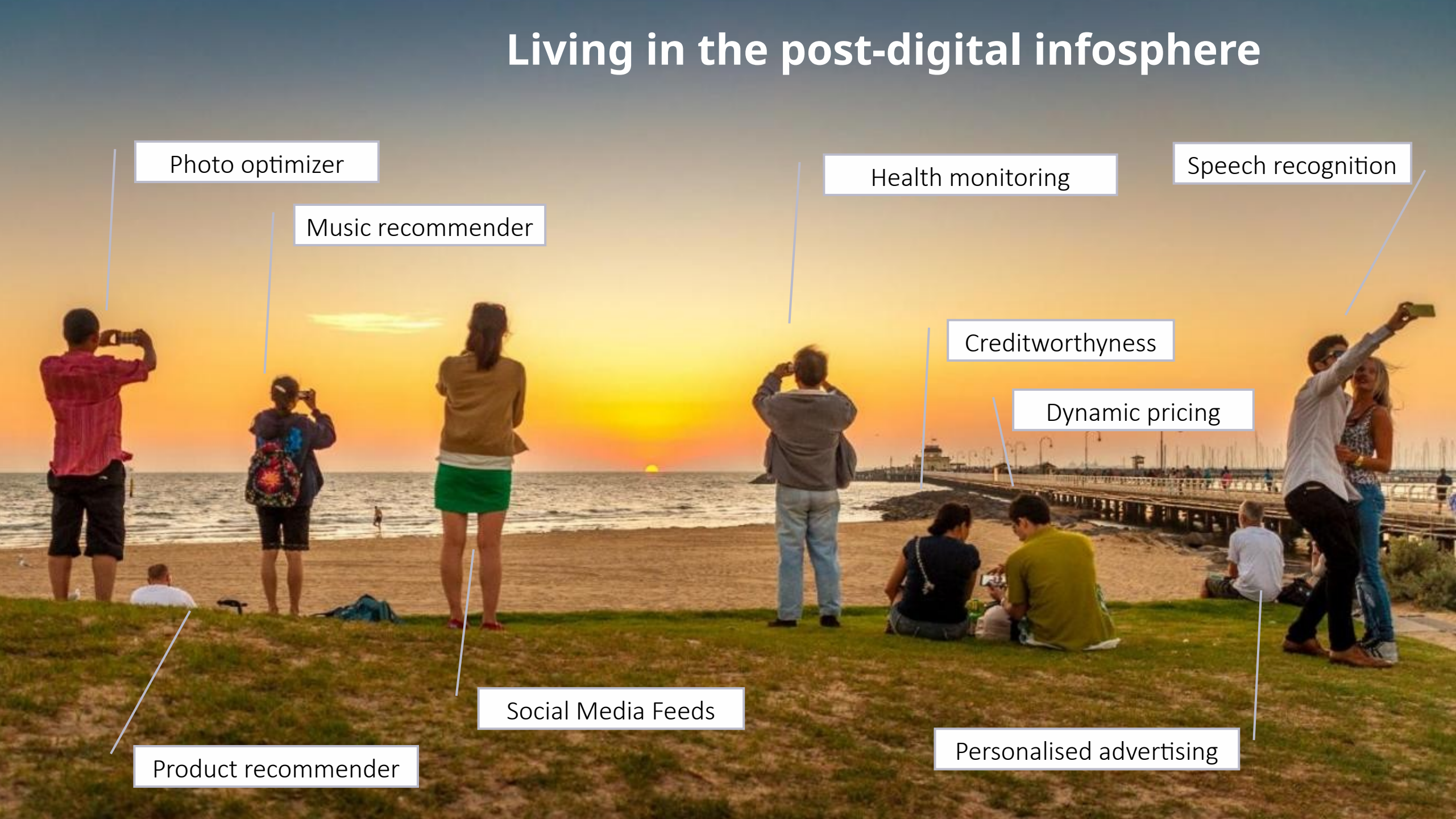
Creditworthiness

Dynamic pricing

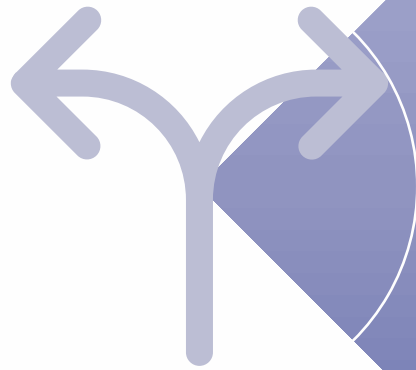
Social Media Feeds

Product recommender

Personalised advertising



Beyond information processing: AI

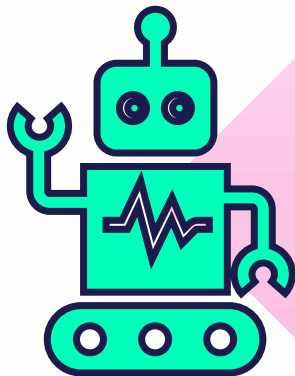


Algorithmic decision support

- Modelling
- Categorization
- Prediction



Netflix. DerStandard.



Autonomous Action – AI Agency

- Autonomous driving
- Robots
- Adaptive interfaces



<https://www.bmwgroup.com/de/innovation/automatisiertes-fahren.html>

What is a digital good life?



Silke Grabinger / Silk: SPOTSHOTBEUYS

Foto: © Meinrad Hofer

<https://www.tanzschrift.at/buehne/kritisch-gesehen/impulstanz/1734-die-ki-ist-ein-hund>

Homo mensura and digital humanism

πάντων χρημάτων μέτρον ἐστὶν ἄνθρωπος

Protagoras, ca. 490-420 v.Chr. (DK 80 B 1)

Digital technology must be designed to empower
people and advance our democratic societies.

Digital Humanism, ca. 2019

- **Digital humanism** is an approach to technology that puts human values, rights, and well-being at the center of how digital systems are designed, built, and used. Instead of asking only “What can technology do?”, it asks “What *should* technology do for people and society?”
- At its core, digital humanism sits at the intersection of Computer Science, Philosophy, and Social Sciences.
- As technologies like Artificial Intelligence, big data, and social media shape everyday life, they influence everything from elections to mental health. Digital humanism tries to ensure these systems respect human dignity, strengthen society rather than divide it enhance, rather than replace, human capabilities.

—EUDHIT

Humanism



Michelangelo di Lodovico Buonarroti Simoni

Humanism as an inquiry into human nature: *conditio humana, specificum humanum*.
Digital humanism: not essentialist, focus on society, environment, and diversity (anti-colonialist)

Digital Humanism

**From abstract principles to lived
practice**

Values, Fairness and Power by Tech?



From the door knob to the interface: Technology represents our values.

It can be fair and unfair and exert coercive power.

IT systems are political.

Five messages of digital humanism



Digital technology changes us and our lives

- Humans and their behaviour
- Societies and culture
- Economy and labour
- Power



Digital technology must protect humans and the environment

- Human dignity
- Autonomy
- Fairness
- Environment
- Diversity



Digital technology must strengthen democracy and society

- Democracy
- Human rights
- Freedom
- Solidarity
- Culture
- Education



Technology is a choice

- Design
- Cooperation of industry and science
- International Cooperation
- Democratic regulation



Humans are not machines

- and vice versa
- Serving mankind
- Care
- Limits of technology

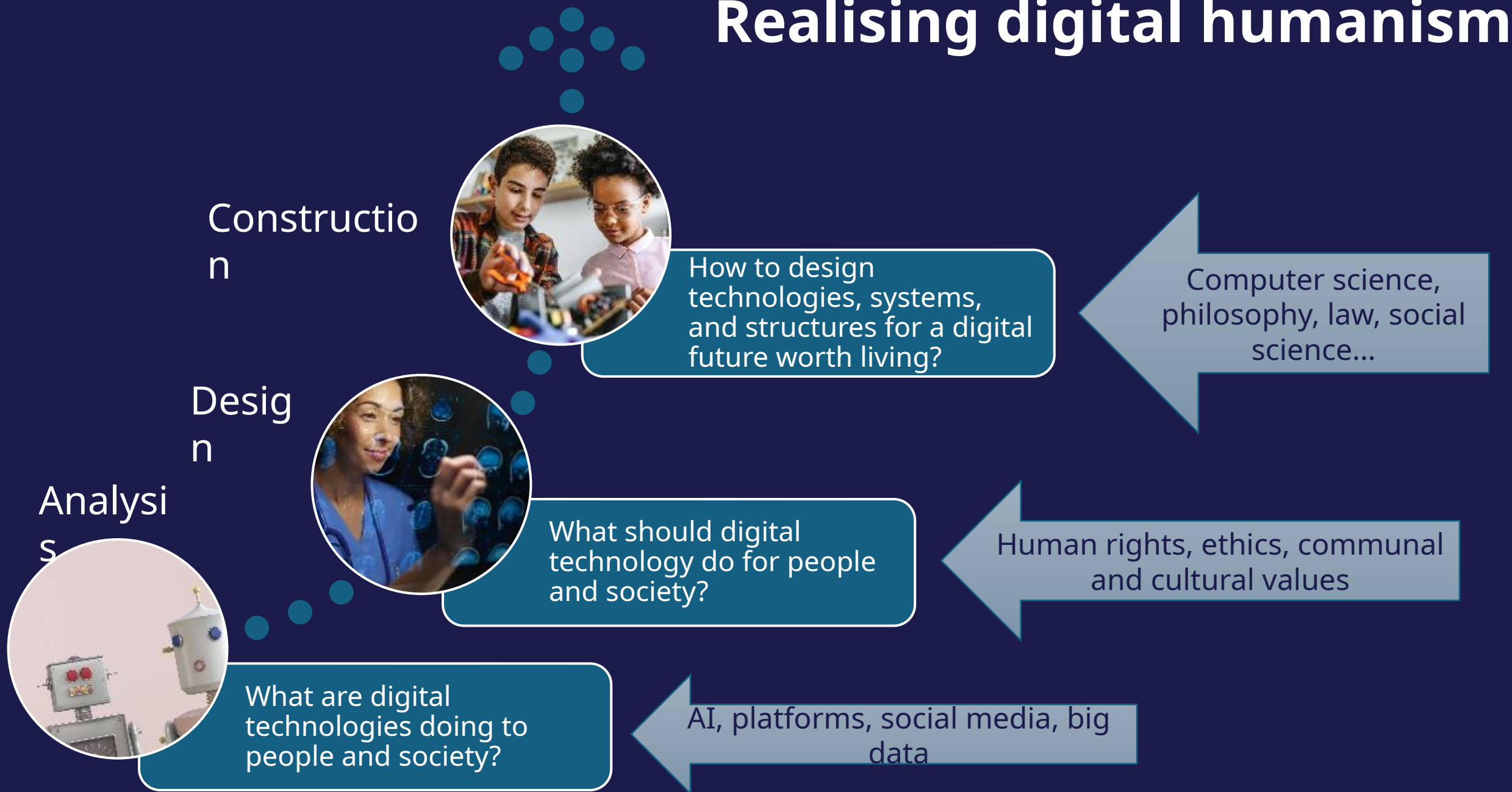
Technology must protect humans...



Third hand.
Stelarc
<http://stelarc.org/?catID=20265>

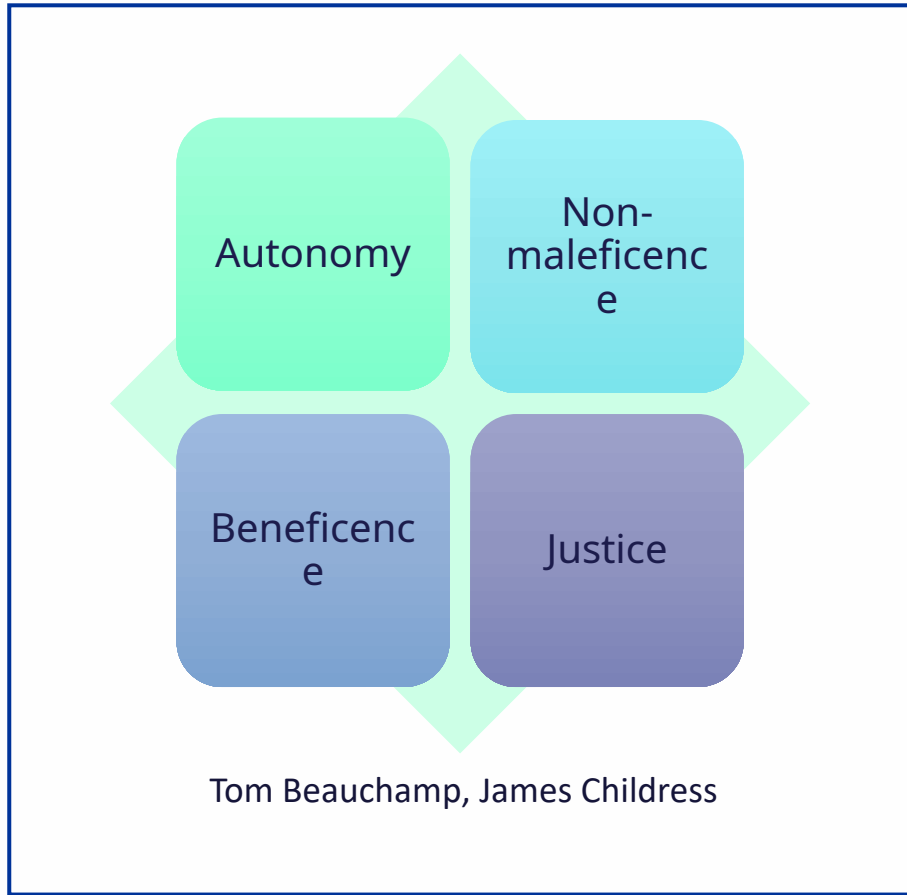
Digital humanism is not transhumanism.

Realising digital humanism



Principlism

Roots in medical history, in unethical and criminal experiments“.



Belmont report (April 18, 1979)

<https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html>

Principle	Example consequence
Respect for the person	Informed consent
Beneficence	Risk-benefit analysis
Fairness	Selection of test subjects

AI Principlism



Explainability

Consequence: XAI

Consequentialism

- **Moral worth of an action is determined solely by its consequences:** focus only on the outcome
- **Utilitarianism:** The most famous form, which aims to maximize pleasure, happiness, or well-being for the greatest number of people.



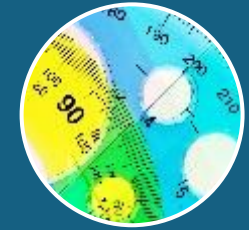
Unpredictability:

Impossible to know all the future consequences of an action before making a decision.



No Red Lines:

May justify harming a minority or violating rights if overall outcome benefits a majority.



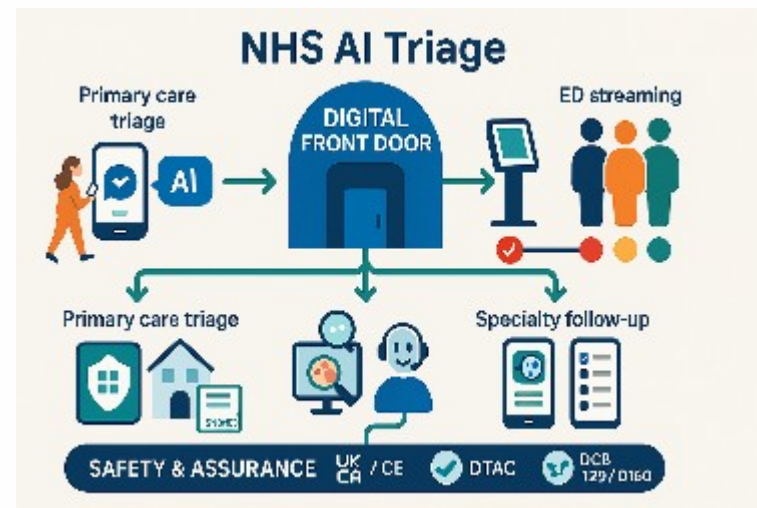
Too Demanding:

Not practically achievable or measurable (utilitarianism).



E.g. self-check-in system for emergency departments: tablets to capture structured patient history and prioritise them by clinical need.

eConsult – eTriage (ED/UTC)



<https://www.iatrox.com/blog/nhs-ai-triage-tools-2025-anima-patches-klinik-econsult-ada-symptoma-ufonia>

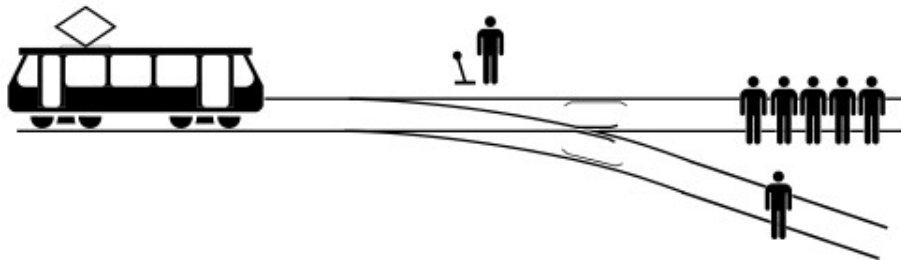
Duty ethics

- Actions based on respect for the moral law. Presuppose adherence to certain norms and obligations.
- Focus on the motivation for a behaviour: AI?
- Your rule of conduct should always apply to everyone. "Act as you would like to be treated."
- People are never means to reach ends.

All human beings are born free and equal in dignity and rights.

They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood.

Universal Declaration of Human Rights



Genuine dilemmatic decisions, such as a decision between one human life and another, depend on the actual specific situation, incorporating unpredictable" behaviour by parties affected. They can thus not be clearly standardized, nor can they be programmed such that they are ethically unquestionable.

Ethics Commission, Federal Ministry of Transport and Digital Infrastructure, June 2017

Virtue ethics

- Let's take inspiration from the character of virtuous people.
- Try to learn how "good" people behave. Virtues can be acquired and must be practiced.
- Behaviour of entrepreneurs, politicians, programmers, etc.



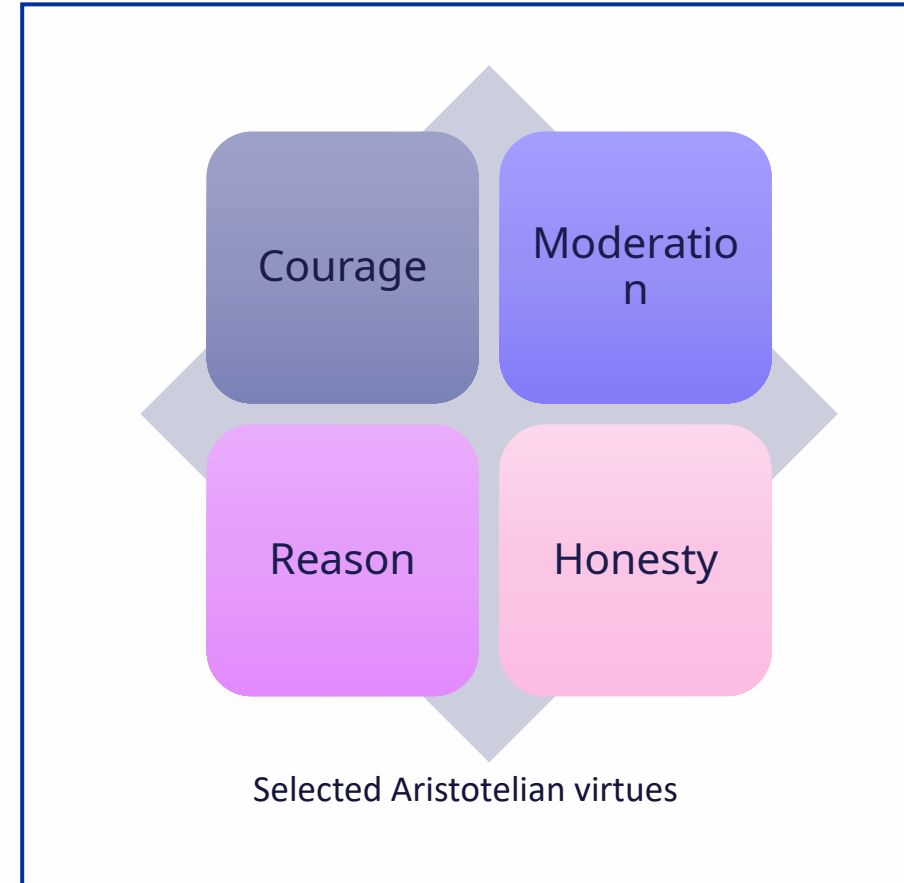
Dame Stephanie
"Steve" Shirley – the
IT woman for IT
women



René Carmille –
the statistician
who saved
thousands



Katie
Moussouris –
ethical hacking



Our wonderful fragility

Humans are fragile beings.

From our vulnerability emerges dependability, hence, a duty of care as well as society and cooperation.

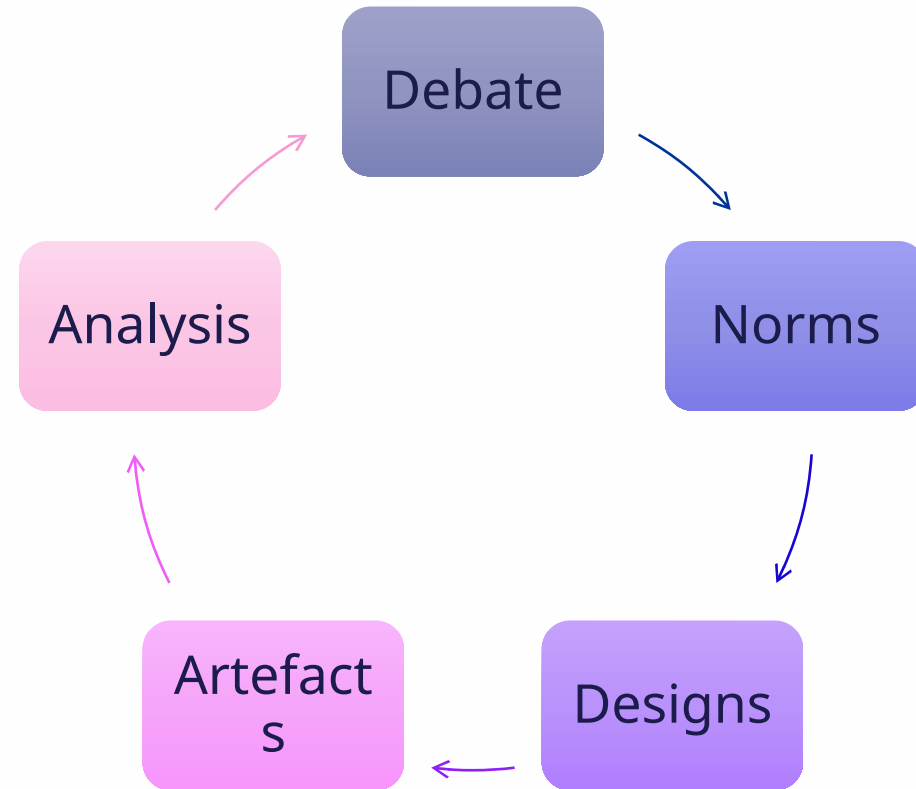
Fragility and vulnerability are not an error; they are part of how we are as humans.



Digital technology has to protect people how they are - not how they might be.

Not one simple answer

- Ethics provides guidance and helps understand consequences and implications of choices.
- Focus on major considerations.
- Not a path to one correct design.



Democracy



Ethics



Engineering

Digital designs as continuous democratic processes.

A normative effort



Dignity

- Humans and their dignity, freedom and self-determination as a basic guideline for digital processes.



Autonomy

- Processes and tools that empower people, support their autonomy and respect their differences.



Fairness

- Fair processes, that further democracy and societal cohesion and solidarity with others.

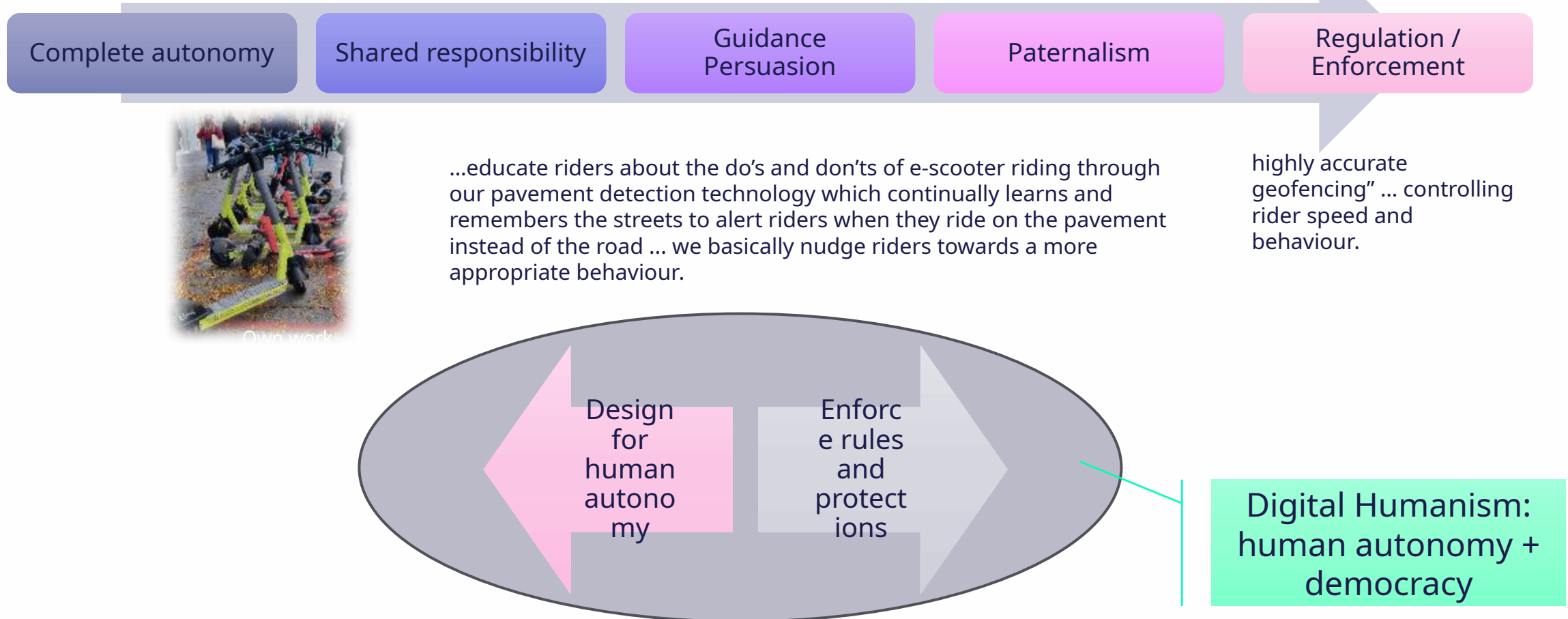


Care

- Digital tools that preserve and develop cultural and social achievements, support those in need, regenerate resources, and preserve nature.

Autonomy, Care or Enforcement?

Digital systems can prescribe and enforce or leave decisions to people.



Digital Technology changes power relations



Designer



Organisation



Hacker



State

Digitalisation shifts power relations from the person to organisations
(and other individuals)

Open design questions and directions



How to design ethically aligned systems?

- Context-dependent, accepted standards, prototype solutions



How to create democracy-compatible designs?

- Information, education, debate, analysis, monitoring



How to limit the power of IT systems and their owners?

- Regulation, user empowerment, divide and conquer

What is the European digital humanism initiative?

www.eudhit.eu





Training

- **Introduction** to digital humanism (simple manifesto, video)
- Online **seminar** series



Practice

- Digital Humanism **success stories**
- Generative AI and **critical thinking** (EUDHIT report)



Tools

- Measuring digital humanism: **indicators** for states and organisations
- A Digital Humanism **investment framework**



Policy

- EU **policy gap** analysis
- Digital humanist **positions** at EU level

The European Digital Humanism Initiative (EUDHIT)

eudhit.eu office@eudhit.eu

Verein zur Förderung des digitalen Humanismus

digitalhumanism.at office@digitalhumanism.at

eutema GmbH

office@eutema.com

eutema.com

Erich Prem

prem@eutema.com

erichprem.at



<https://link.springer.com/book/10.1007/978-3-031-45304-5>

<https://link.springer.com/book/10.1007/978-3-030-86144-5>