

# AI Ethics as Phoenix: Long Live the Good Life

**Dr. Nadezhda Gotcheva**  
**Senior Scientist**

14/10/2019 VTT – beyond the obvious

## *Beyond the obvious*

Established in 1942, VTT is one of the leading research, development and innovation organizations in Europe.

**Mission:** VTT helps customers and society to grow and renew through applied research.

**Strategy:** Make an impact through scientific and technological excellence.

VTT is owned by Ministry of Economic Affairs and Employment in Finland

# Contents

- Opening words
- AI + ethics = AI ethics: "the hottest service/product of our time"?
- Ethics and technology design
- Some examples of ethical AI research in Finland
- Closing words



Credit: Guardian

*“What I find so particularly interesting is the sense of calm that pervades all things Japanese. I am struck by how in Japan modernity mingles seamlessly with history. The Japanese manage to convey the sense that they are able to look both backwards to tradition and forwards to the unknown future with equal ease.”*

Giorgio Armani

“Future and the Arts:  
AI, Robotics, Cities, Life –  
How Humanity Will Live Tomorrow”

What is true affluence,  
What is it to be human,  
What is life?

## Bridging arts and technology

- An upcoming exhibition at Mori Art Museum in Japan (2019.11.19 - 2020.3.29)
- The title was chosen from over 15,000 AI-generated options
  - Mori Art Museum has collaborated with the “IBM Watson” suite of AI technology and services developed by IBM to come up with the title for this exhibition.

*“When human beings will be entrusting many of their decisions to AI, what would it mean to be a human, and what constitutes life?”*

# What constitutes good life?



# Technological scaffolding

- We, humans, have always used technological scaffolding to push ourselves beyond our biological and other limitations. We developed and used the telescope to extend our line of sight, and we can use AI to bridge the borders, to stretch our minds, expand our creativity and reach to the next level of humanity.

SingularityHub







**Technology should be harnessed to support  
good life for all**

# Good life for all

- To paint the picture of a good life that is supported by technology, we need a rich color palette to highlight self-efficacy, meaningfulness, independency, social relations, joy and happiness...
- AI has the potential to cut deep into the established fabric of our societies and the power to change fundamentally our everyday life.
- How we can harness AI for good and how do we anticipate the future impact of the current decisions? What are the important questions to ask?

# The important questions

- When it is asked “whom an autonomous car should be allowed to run over”, it is the wrong question (Leikas, 2019)
- The important questions to ask are those which focus on ensuring peace, safety and security of citizens, trust in society, equal availability of services, possibility to be heard, and justifying technology decisions from the perspectives of human dignity, welfare and sustainability.

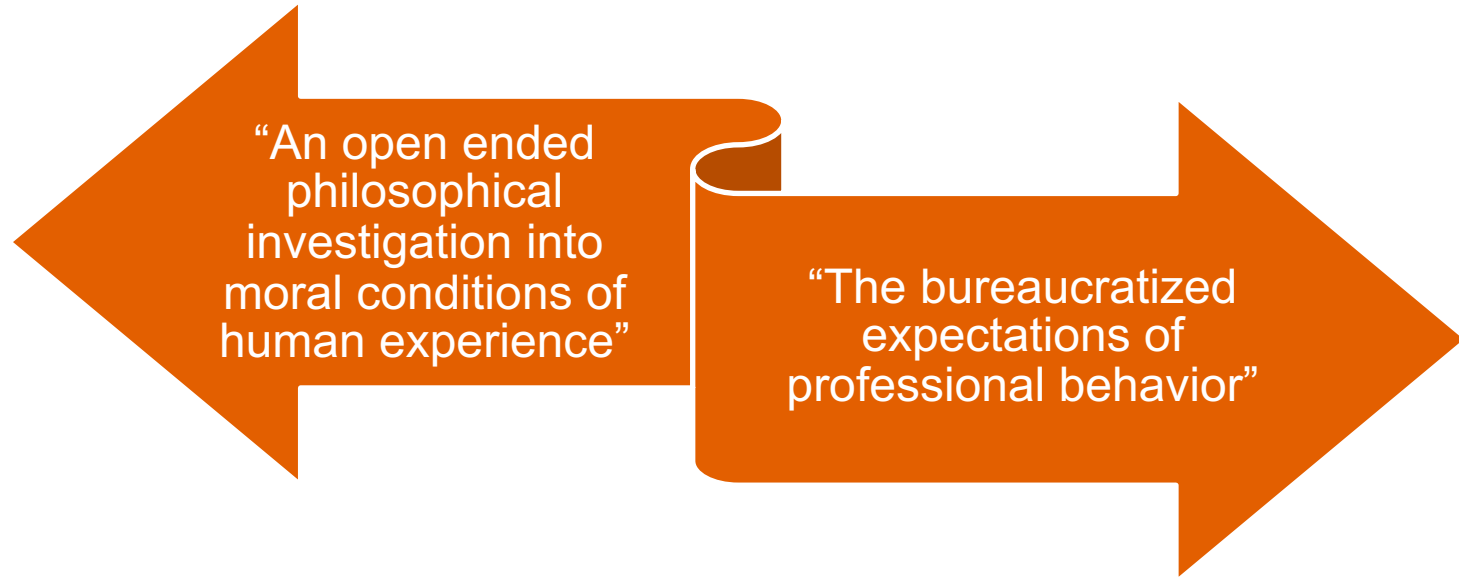


# To be treated with dignity

the state or quality  
of being worthy of  
honour or respect

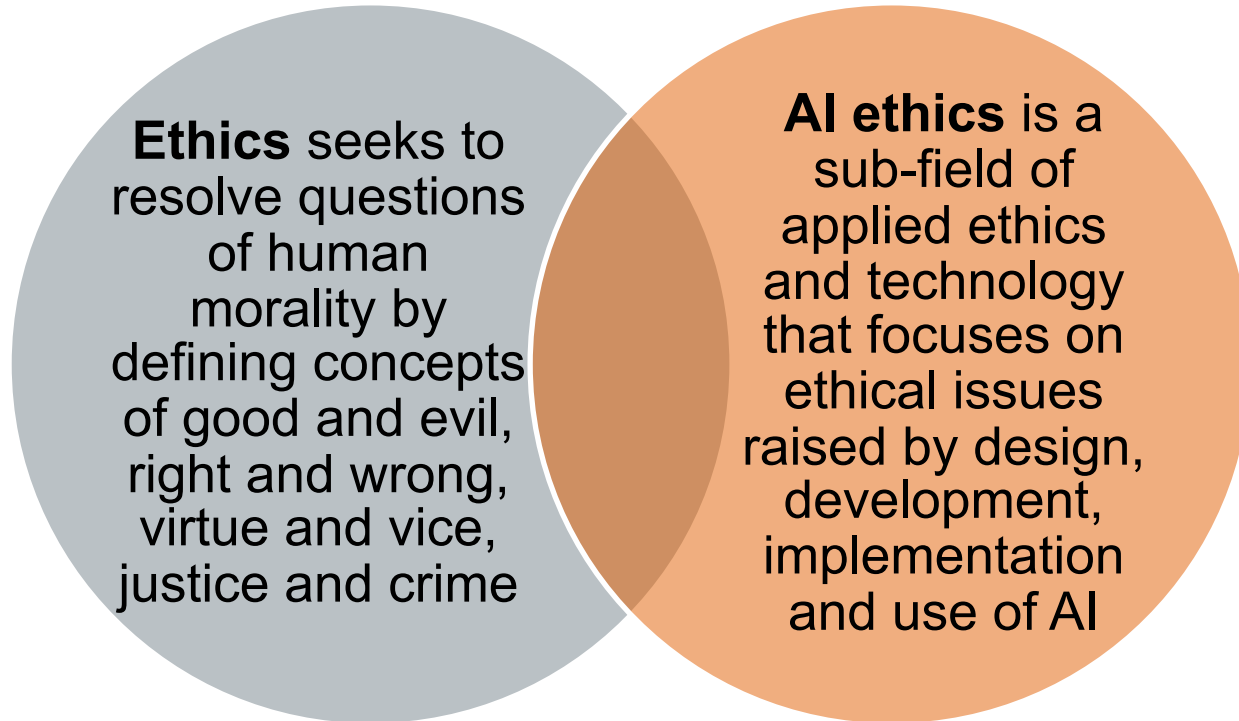


# Ethics: The “hottest product” of our time”, “an organizational responsibility”, “a capacious term”



Metcalf et al., 2019

# Ethics → AI ethics



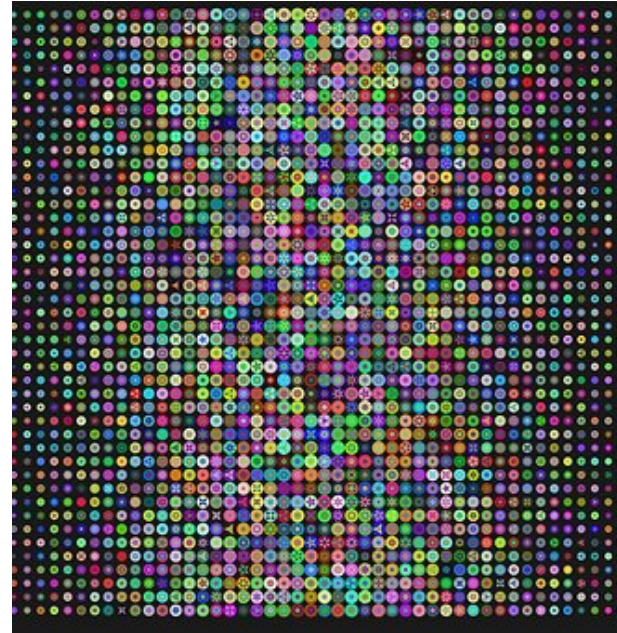
# Global competition for AI supremacy

- Businesses invest heavily in projects that utilize the capabilities of AI software and platforms
  - Worldwide spending on AI Systems: \$98 Billion in 2023 vs. \$37.5 billion spent in 2019 (IDC, 2019)
  - In Europe, spending for AI-based technologies in 2019 has increased 49% for a year, to reach \$5.2 billion (IDC, 2019).



# Who is designing the algorithms of tomorrow?

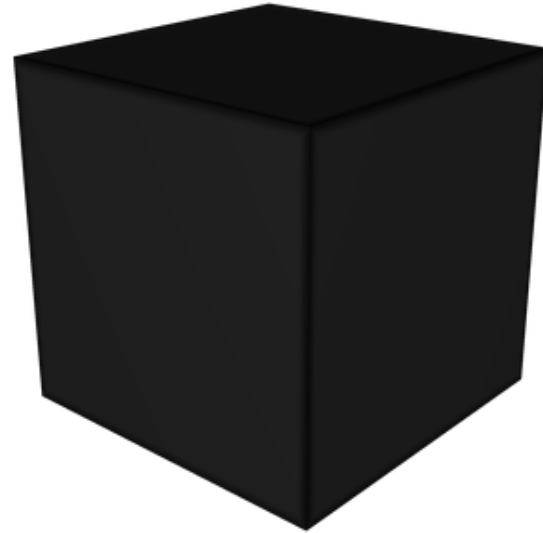
- In the current Finnish government, women account for 11 of the cabinet's ministers, while 8 are men.
  - Studies indicate that women tend to identify with moral traits more strongly than men (Kennedy et al., 2017)
- Low numbers of women in maths and science careers, as well as programmers and tech developers
  - “90% of coders are men designing the algorithms of tomorrow” (Pannier-Runacher, 2019)





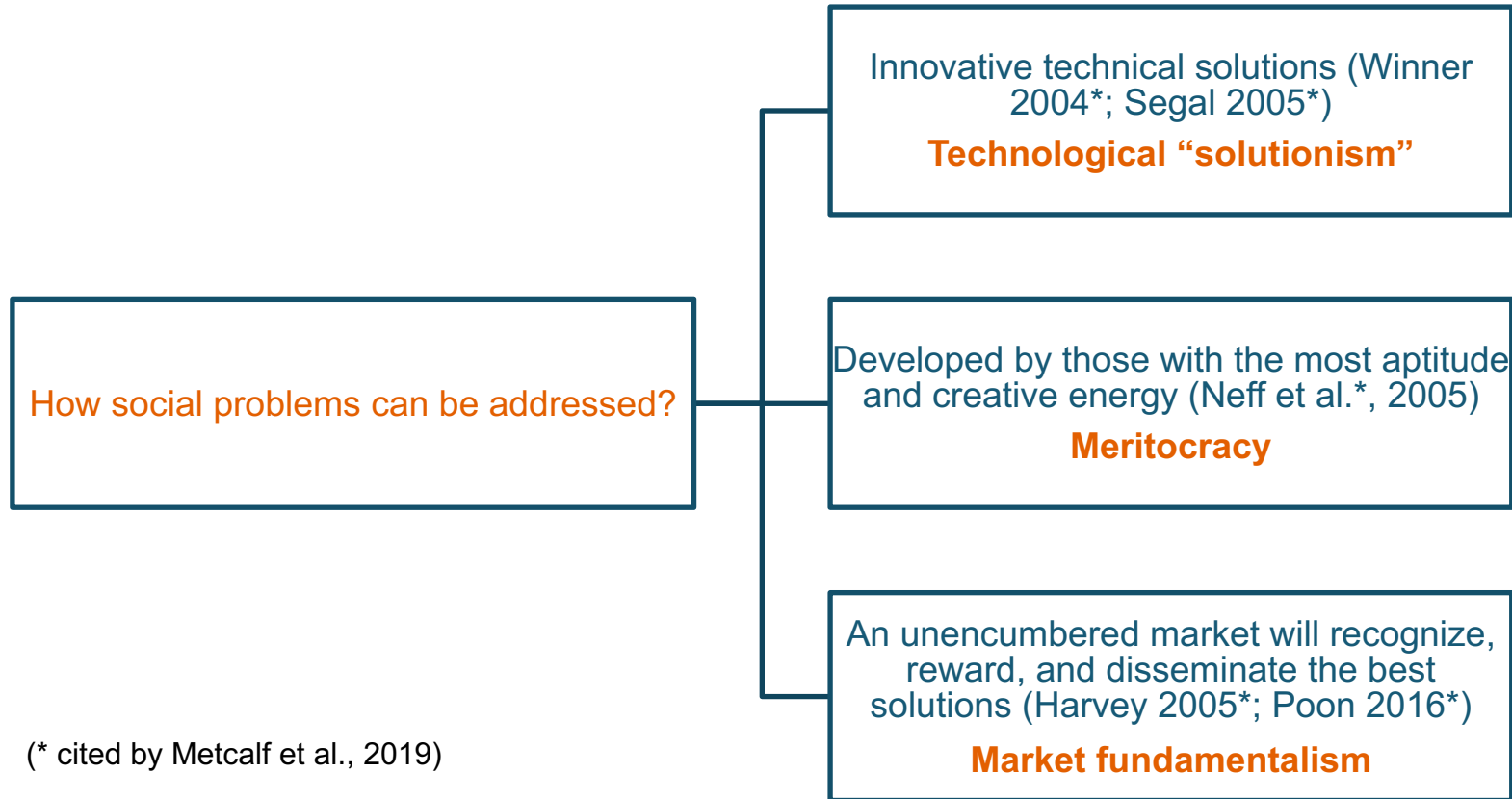
# AI as a black box

- AI may be portrayed as a **black box** for a reason, and we may not be able to understand the codes but we need to understand the **designers' and implementors assumptions, reasoning and the factors that are utilized in a technology solution.**



Kathy Baxter, cited by Vohra (2018)

# Silicon valley: Dominant tech industry logics



(\* cited by Metcalf et al., 2019)

# What the future holds and what we can do?

Global transition towards ethical AI society?



# Humble leadership (Schein and Schein, 2018)

The power of relationships, openness and trust

- **Humble inquiry, genuine curiosity towards the others** (not as roles but as whole persons), **and the ‘art of asking instead of telling’** (Schein and Schein, 2018).
- Advocates open trusting communication and relationship building for collaboration.

- *It is the responsibility of the leader to build openness and trust*
- *It is not just being polite and nice – one needs to genuinely like people and be interested in them, to invest his/her personal time and energy to build and maintain relationships.*

# AI Ethics as Phoenix

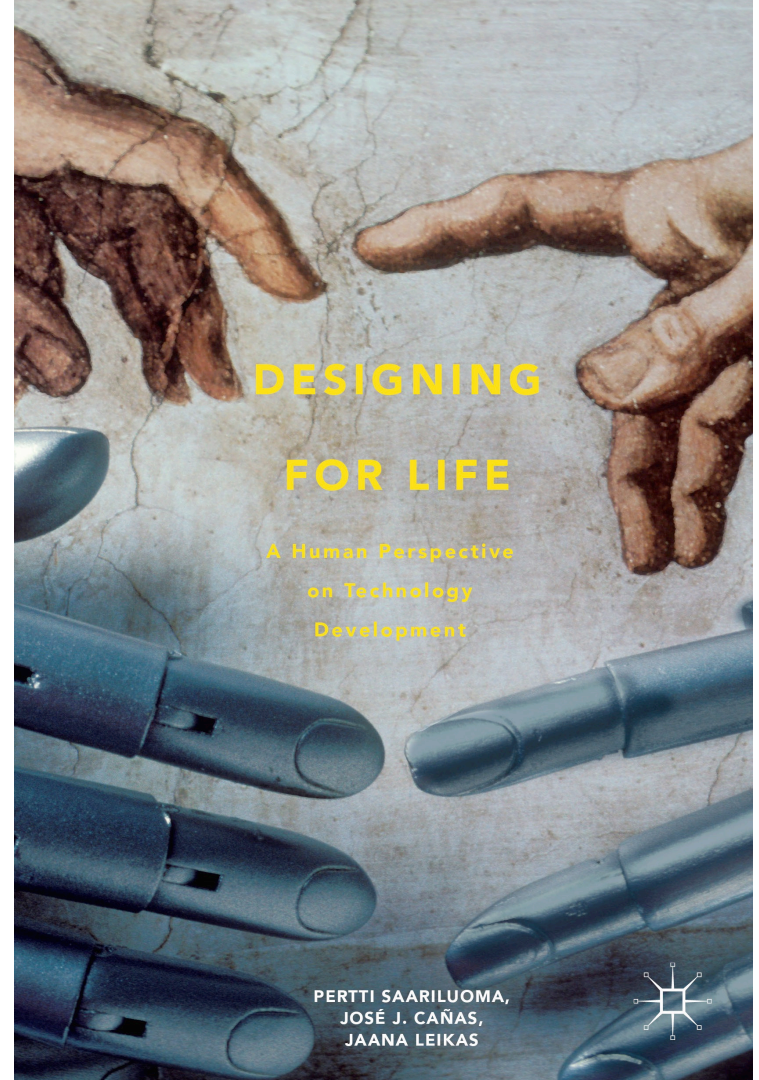
- Today, a renewal of the ethical discussion is on the agenda.
- Paired with the pervasive use of new technologies and especially AI, technology ethics is born again, like a Phoenix bird, across cultures and throughout time, to illuminate the way to live a good life.

In many cultures, the mythical Phoenix bird symbolized immortality, loyalty, honesty and doing good deeds.



# Technology development: A better life for all of us?

- What is technology used for?
- Who benefits from it and how?
- Does it harm?
- Are we ready to adopt it?
- Is our environment ready?



# Ethics and design: Three levels

## Ethics by design

- Positive and sustainable impact

## Ethics in design

- New design processes, methods and tools

## Ethics for design(ers)

- Professional ethics



## **Ethics by design**

Which values can take us forward?

What creates trust in society?

What kind of fringe areas does digitalization create,  
and who lives there?

To what kind of new humanity would new technology  
liberate us?



# Ethics by Design: From principlism to design thinking



- From acceptability to desirability:  
Meaningfulness instead of negations
- Co-design with experts and stakeholders
- Create shared value
- Multidisciplinary and multi-actor  
ecosystems

Principlism: A system of ethics based on the four moral principles of autonomy, beneficence, nonmaleficence, and justice

# Justification of technology

- ***”Technology is neither good nor bad – nor is it neutral”***  
(Kranzberg 1986)
- **Ethical questions are always present, consciously or unconsciously**
- **Ethical codes and legislation far behind technology development**

# Ethically sustainable welfare technology

- It shall not simply respond to the everyday needs of people but we expect it to contribute significantly to quality of life and generate new, unexpectedly beautiful, positive needs.



# Ethical values and principles

- **Dignity**
- **Respect for rights**
- **Do no harm**
- **Justice**
- **Autonomy**
- **Informed consent**
- **Privacy**
- **Equality...**



# Dialogue and understanding of the context

## Temporal value of decisions

What is the value of today's decisions in the future?

Will the decisions stand the test of time?

## How to steer the change in society?

What kind of marginals evolve?

"Hollows" in the welfare society? Who are they?



## ”bttn – the IoT button”

- By the Finnish designer Harri Koskinen
- Not a panic button or an emergency device
- A tactile, simple use
- No need for smart phone – has own mobile data or Wi-Fi connectivity
- Elderly people can press the bttn to contact relatives to indicate greeting or a need for assistance

<https://bt.tn>

## Small things matter



# How to promote ethics in technology development? What dimensions should be taken into account?

## Ethics in Design process & tools

How to implement ethics in product/service development processes

How to guarantee ethical outcome

How to manage knowledge accumulation from one design process/design version to another

## Ethical assessment of use cases

What are case-specific ethical issues

What are context-dependent issues

How to guarantee citizen involvement

## Ethical impact in society

How to ensure acceptability in society

How to exploit technology optimally in society

How to justify technology in terms of Quality of Life of citizens

How to govern positive impact

- Integrity
- Autonomy
- Informed consent
- Participation
  
- Transparency
- Privacy
- Trust
- Accountability
  
- Discourse
- Responsibility
- Fairness
- Sustainability

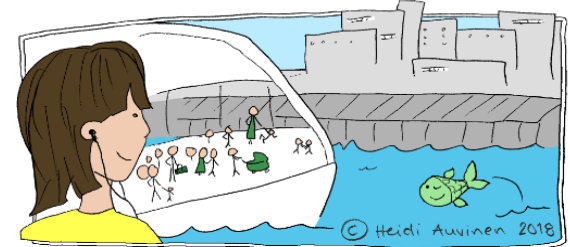
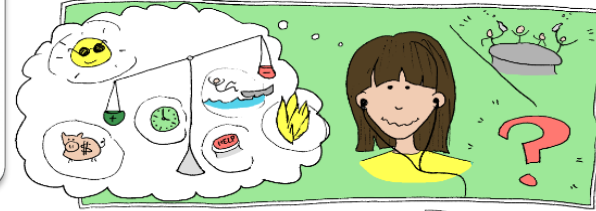
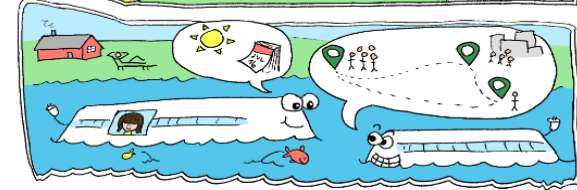
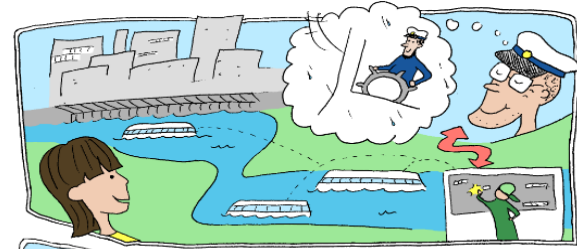
# An example

How to address ethical issues in design of autonomous systems?

1. Identification of ethical values in the given context

2. Examination of the design goals against the identified values

3. Analysis and understanding of ethical issues based on the examination



(Leikas, Koivisto & Gotcheva, 2018)



# Finland: the first EU country to write a national AI strategy

In October 2017, Finland was the first EU country to put a national AI strategy in writing



In a second report released in June 2018, the government estimated that around 1 million of its population would need to update their AI skills.

# Demistifying AI: Finland launched free AI education for all

## Welcome to the Elements of AI free online course!

Join over 205,000 other people learning about the  
basics of AI.

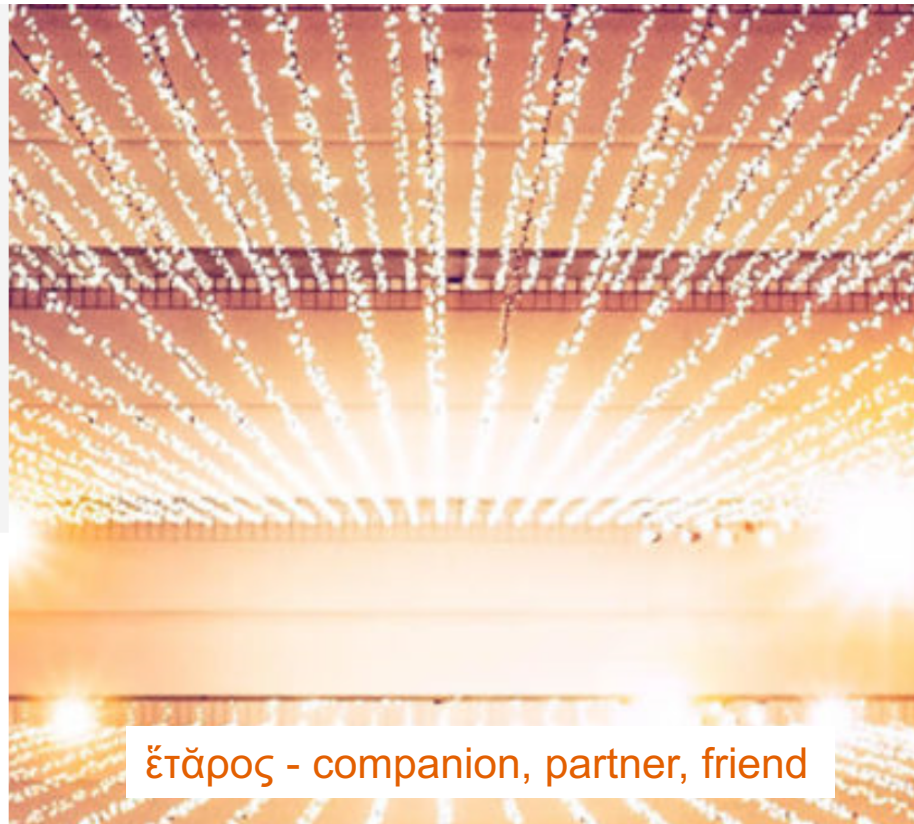
<https://www.elementsofai.com/>

ETAIROS

# Towards ethical use of AI

The aim of the ETAIROS project is to study and develop practical processes and frameworks that help public, private and third-sector organizations enhance the ethical sustainability of applying Artificial Intelligence (AI) and associated technologies.

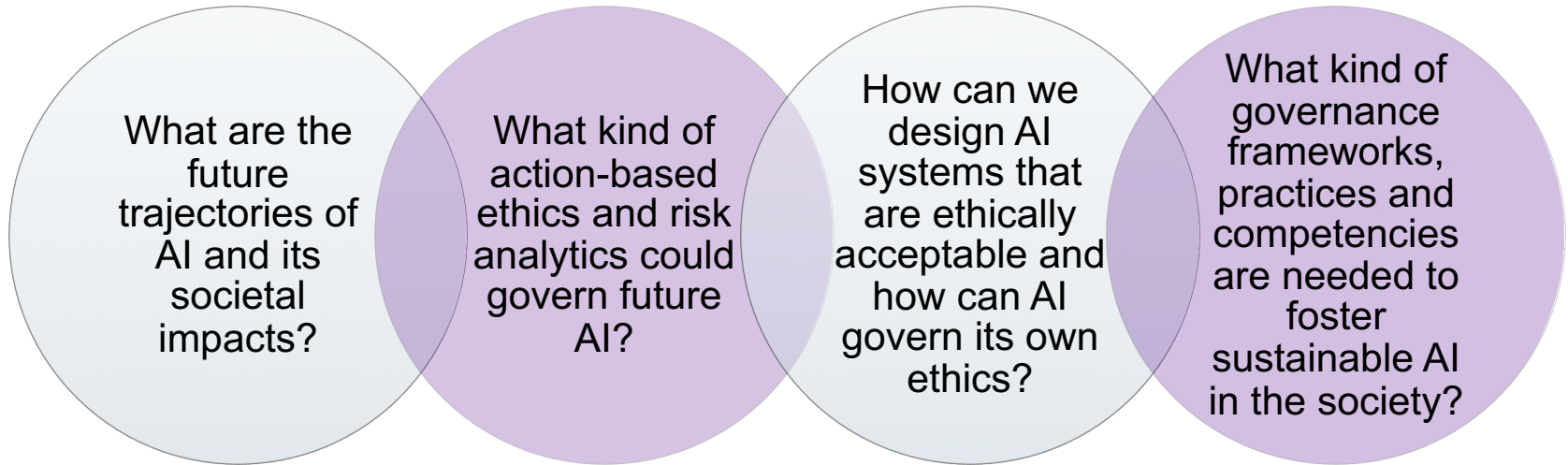
“Ethical AI for the Governance of the Society” is funded by Strategic Research Council at the Academy of Finland.



ἑταῖρος - companion, partner, friend

<https://www.etairos.fi/en/>

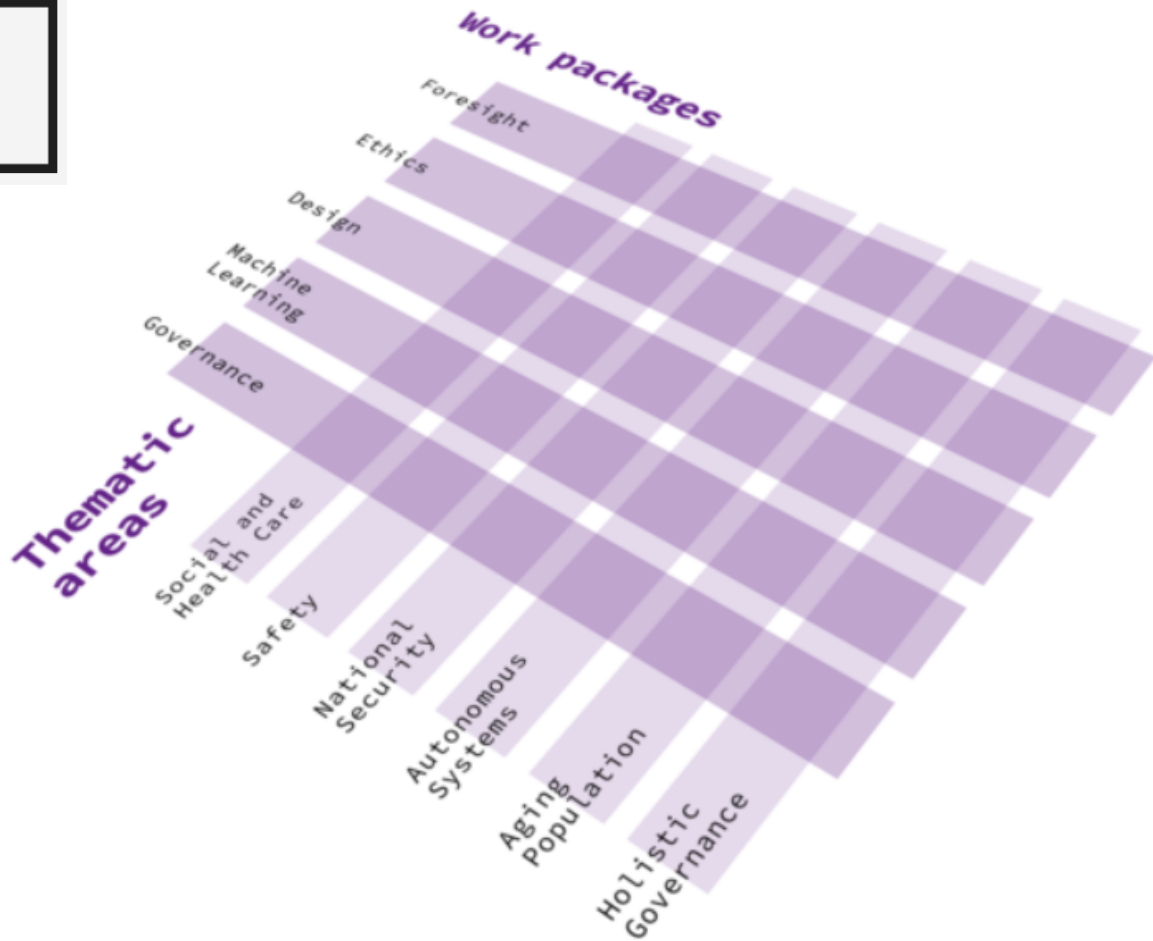
# ETAIROS project: Main research questions



The research project is implemented jointly by University of Tampere, University of Helsinki, University of Jyväskylä, University of Turku and VTT Technical Research Centre of Finland Ltd. 4FRONT Pvt. Ltd. is responsible for the communications and interaction.

ETAIROS

## Structure and thematic areas



## Closing words

*“...sense of calm”*

Daily life is increasingly hectic and we suffer from information overload – we need to take time to reflect, to allow time for dialogue and listening to ensure positive impact rather than wishful thinking by ethical discussion.

*“...modernity mingles seamlessly with history”*

Ethics is ancient yet AI ethics is modern

*“...able to look both backwards to tradition and forwards to the unknown future with equal ease”*

We need futures thinking and action yet we need to learn from the past to co-create a desirable future and good life for all.

# References

- Guardian (2019) Giorgio Armani: 'Japan always keeps its soul' <https://www.theguardian.com/lifeandstyle/2019/sep/28/giorgio-armani-japan-always-keeps-its-soul>
- Guardian (2019) Barbie, the astronaut: French toymakers sign up to fight stereotypes, Pannier-Runacher, A., French Minister of State, attached to the Minister of Economy and Finance (cited by Guardian), <https://www.theguardian.com/world/2019/sep/24/barbie-the-robotics-engineer-french-toymakers-pledge-to-fight-stereotypes>
- IDC (2019) Worldwide Spending on Artificial Intelligence Systems Will Be Nearly \$98 Billion in 2023, According to New IDC Spending Guide, International Data Corporation <https://www.idc.com/getdoc.jsp?containerId=prUS45481219>
- Kennedy, J. et al. (2017) Social-cognitive approach to understanding gender differences in negotiator ethics: the role of moral identity, *Organizational Behavior and Human Decision Processes*, Volume 138, January 2017, Pages 28-44.
- Leikas, J., Koivisto, R., & Gotcheva, N. (2019). Ethical framework for designing autonomous intelligent systems. *Journal of Open Innovation: Technology, Market, and Complexity* , 5(1), [18]. <https://doi.org/10.3390/joitmc5010018>
- Leikas, J. (2019) The ethics of AI – what are we even talking about?, VTT Blog <https://vttblog.com/2019/01/16/the-ethics-of-ai-what-are-we-even-talking-about/>
- Metcalf, J., Moss, E. and Boyd, D. (2019). Owing Ethics: Corporate Logics, Silicon Valley, and the Institutionalization of Ethics, *Social Research: An International Quarterly*, Volume 82, Issue 2, Summer,, pp. 449-476.
- Mori Art Museum (2019) [https://www.mori.art.museum/en/exhibitions/future\\_art/index.html](https://www.mori.art.museum/en/exhibitions/future_art/index.html)
- SingularityHub (2019) <https://singularityhub.com/2019/06/17/the-rise-of-ai-art-and-what-it-means-for-human-creativity/>
- Vohra, A. (2018) AI and Ethics: What designers need to know? A conversation with Kathy Baxter <https://medium.com/net-solutions-ux/ethics-and-ai-what-designers-need-to-know-67f81506f5f6>

When not marked otherwise, photos credit: Pixabay or VTT

# bey<sup>0</sup>nd

## the obvious

Dr. Nadezhda Gotcheva  
Nadezhda.gotcheva@vtt.fi  
+358 40 132 6030

@VTTFinland  
@

[www.vtt.fi](http://www.vtt.fi)