AI Ethics in Practice: Designing for Ecosystems

Nitin Sawhney, Ph.D.

Professor of Practice

August 5, 2020

Department of Computer Science

Aalto University
School of Science

Human-Centred Research & Design in Crisis

Ethics in AI: Three-Part Discussion

Ethics & Politics of AI in Society

AI Ethics in Practice: Designing for Ecosystems

Decolonizing AI & Rethinking Resistance

Aalto University
School of Science

Ethics of AI in Society

Recapping the Key Concepts

1. Bias and Fairness
2. Accountability and Remediability
3. Transparency, Explainability and Trust
4. Safety and Privacy
5. Value-Alignment

Aalto University
School of Science

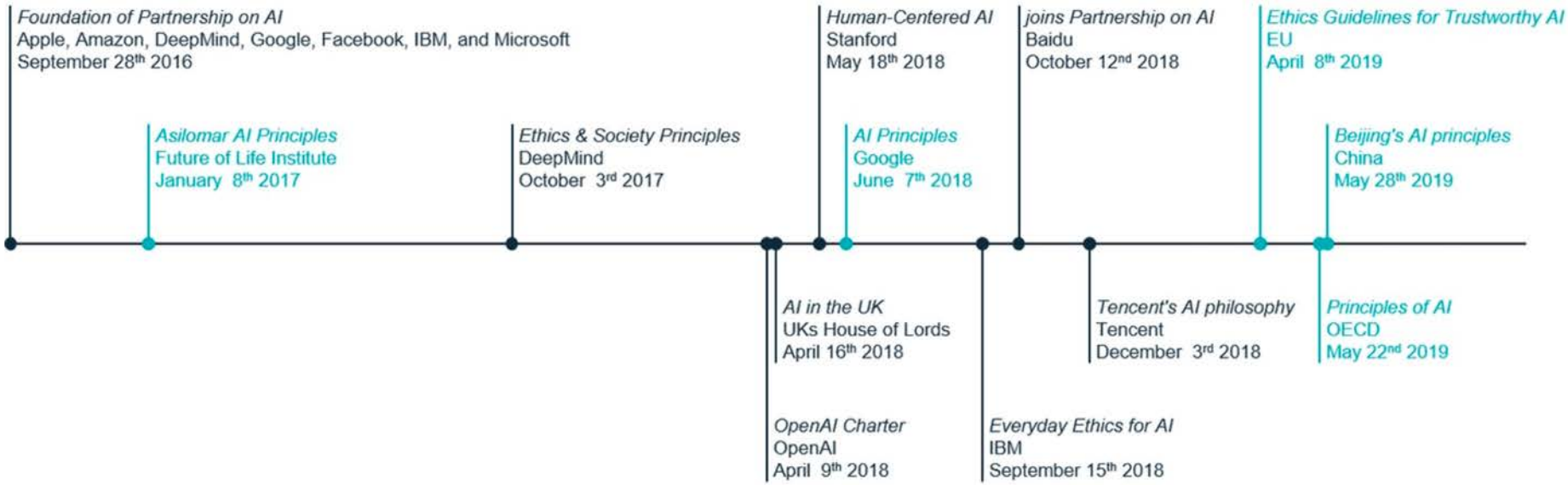
AI FOR GOOD

Solving Humanity's Challenges With Artificial Intelligence

Engaging Ethical AI Principles in Practice

Aalto University
School of Science

Ethics in AI: Timeline of Deliberations



Markus Schmitz, *Artificial Intelligence and Data Ethics II*, CIONET, Nov 5, 2019.

Aalto University
School of Science

PARTNERSHIP ON AI

ABOUT PARTNERS RESOURCES CAREERS

VIEW ALL | NEWS | RESEARCH | WORKING GROUPS | EVENTS

www.partnershiponai.org

NEWS

What's the responsibility of the AI industry in ensuring that AI serves to create an inclusive global economy?

NEWS

It matters how platforms label manipulated media. Here are 12 principles designers should follow.

PARTNERSHIP ON AI

NEWS

Statement by PAI's Executive Director on Black Lives Matter and Systemic Racism

NEWS

Building Responsible AI with New International Partners

Values AI needs to respect

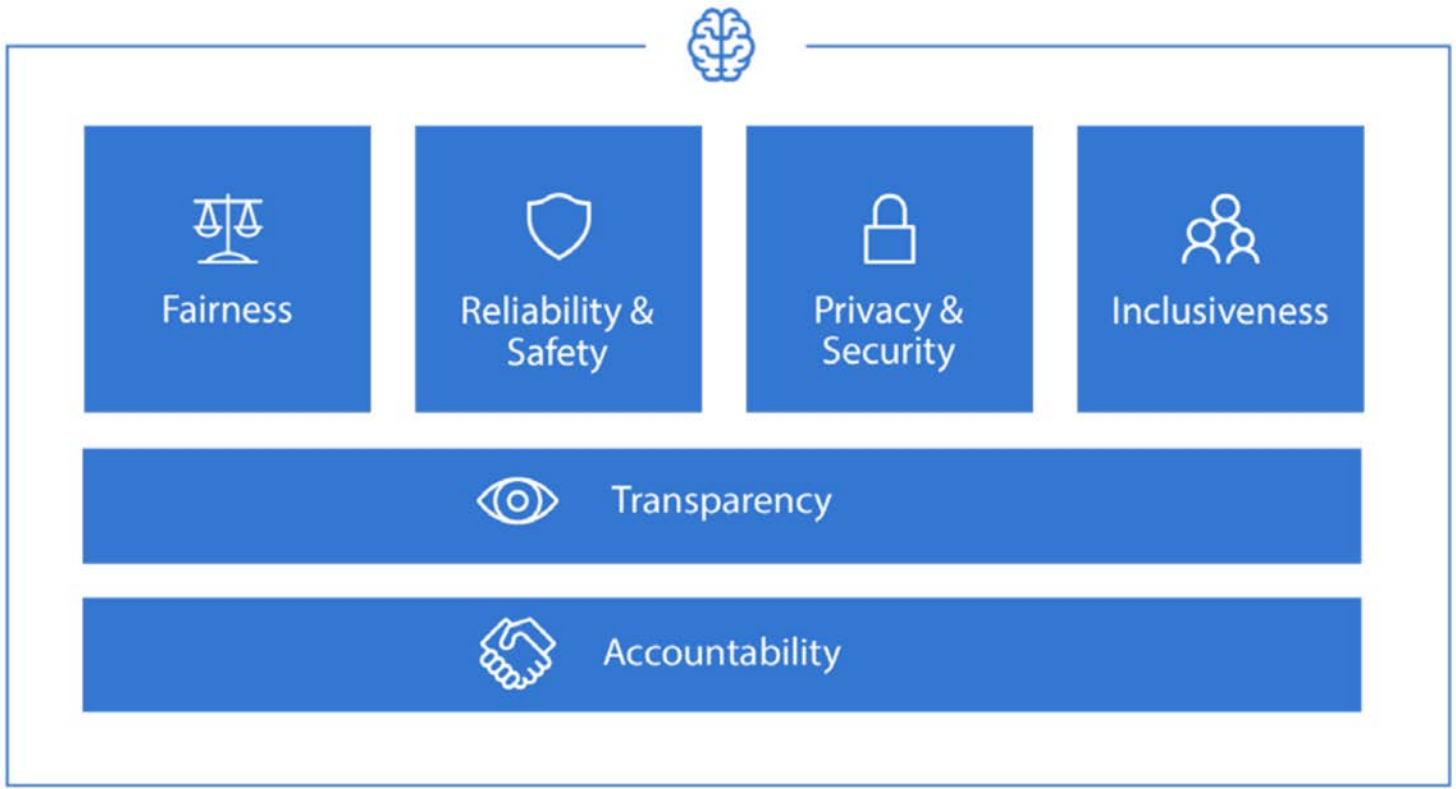
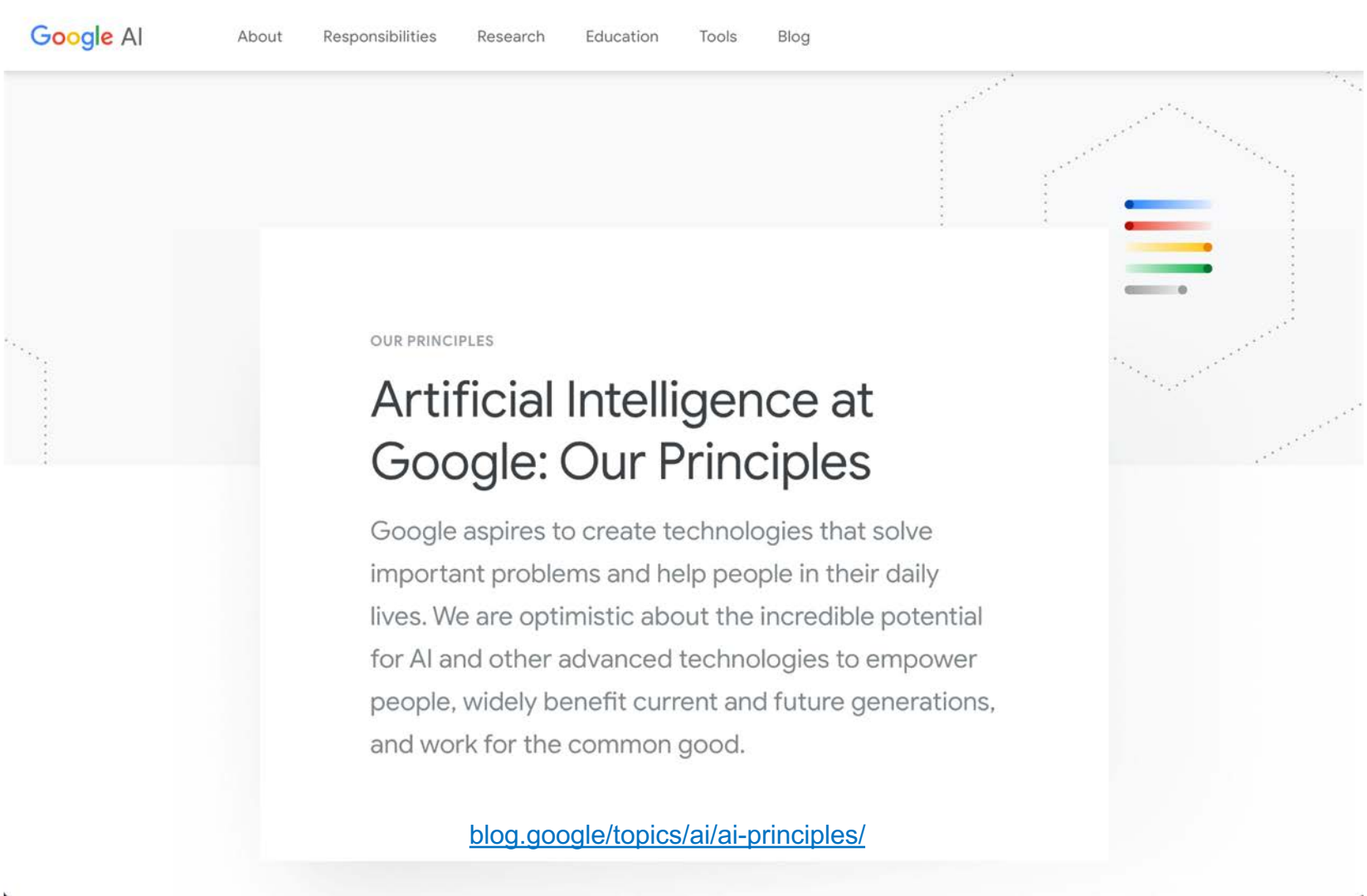


Chart 5.
Source: Microsoft Corporation

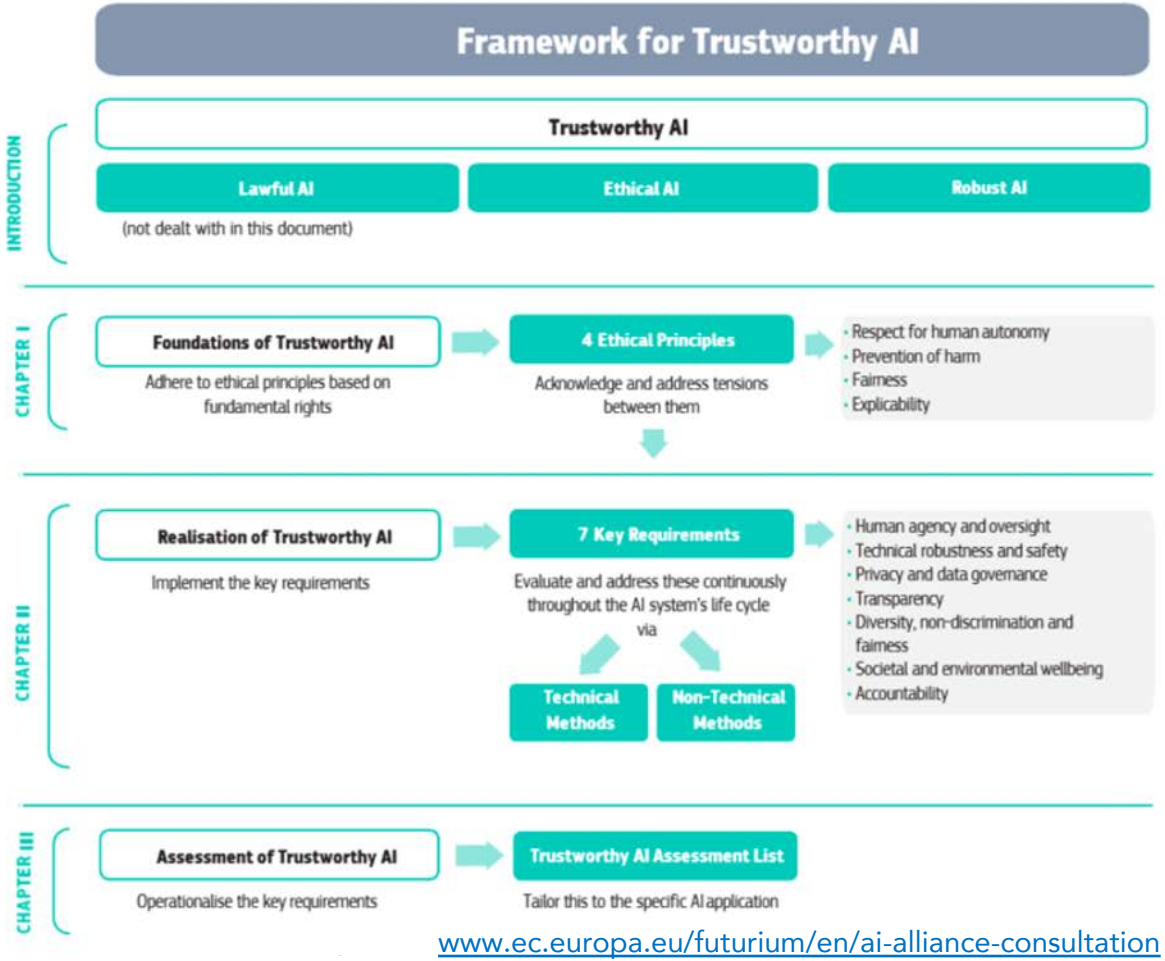
www.microsoft.com/en-us/ai/responsible-ai

Aalto University
School of Science



Aalto University
School of Science

EU Trustworthy AI Ethics Framework

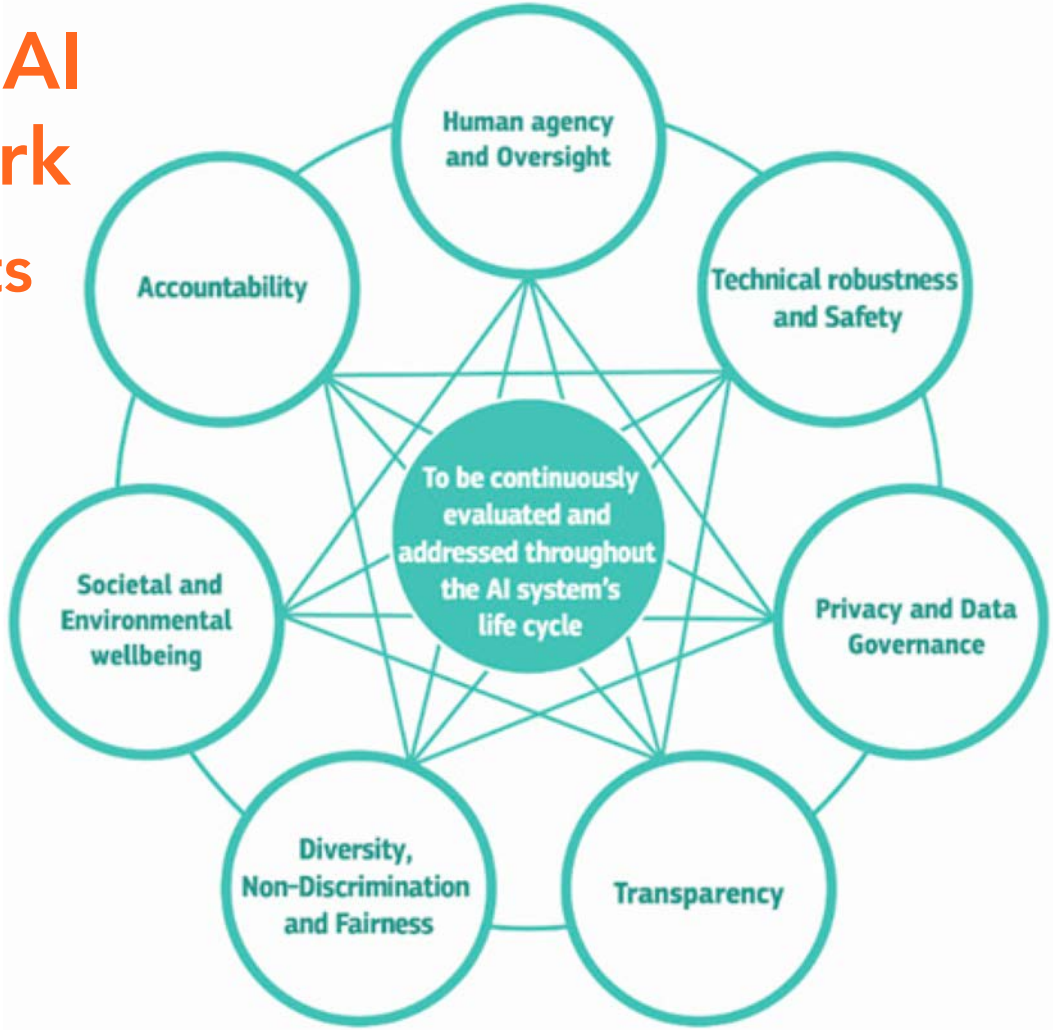


Aalto University
School of Scien

www.ec.europa.eu/futurium/en/ai-alliance-consultation

EU Trustworthy AI Ethics Framework

7 Key Requirements



Aalto University
School of Science

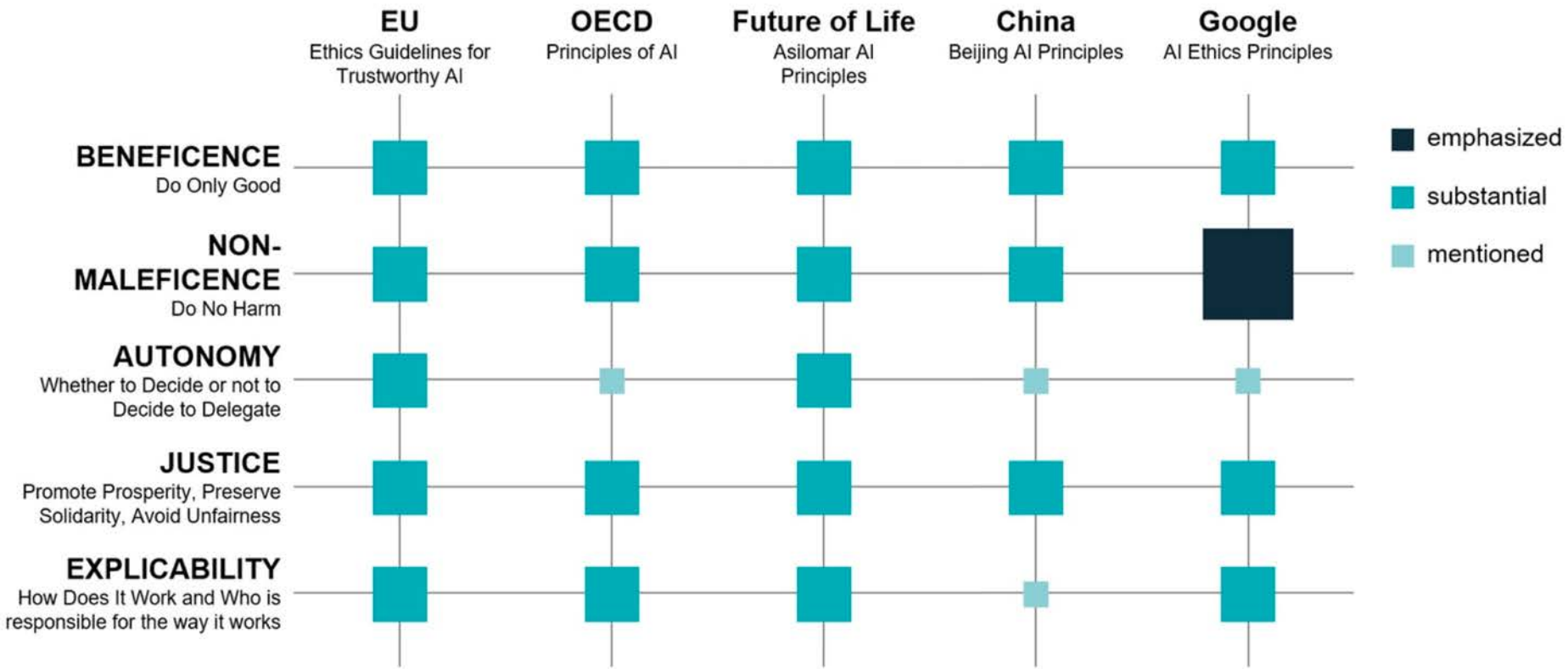


Figure 2: Five organizations in alignment with Floridi's AI principles.
The darker and larger the square, the more focus on the respective principle
Markus Schmitz, *Artificial Intelligence and Data Ethics II*, CIONET, Nov 5, 2019.

Ethics in AI

Comparing Principles

The principles published by different organizations are not contradictory and **follow similar themes**.

They are all highly general and need to be more specifically **tailored to be actionable**.

The principles themselves are **inherently in conflict**, requiring critical tradeoffs for decision-making.

There is **no enforcement mechanism** for the principles; none of them are legally-binding.

Aalto University
School of Science


THE ROYAL SOCIETY
DeepMind
You and AI
THE ROYAL SOCIETY

Enabling AI Ethics in Practice:
An Action Agenda Moving Forward

Aalto University
School of Science

Enabling AI Ethics in Practice

Action Agenda Moving Forward

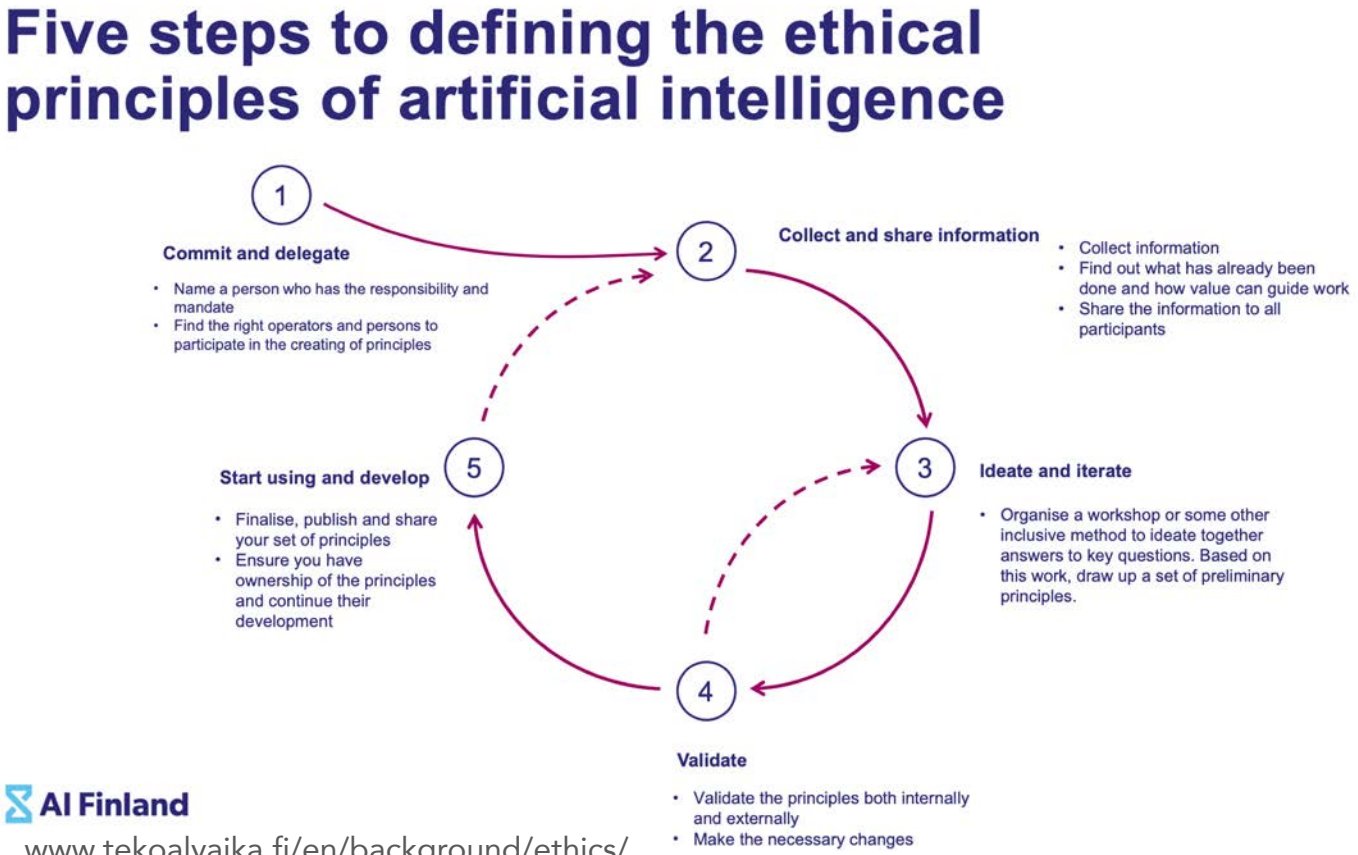
1. Devising Ethical AI Principles
2. Ensuring Oversight & Governance of AI
3. Promoting Diversity and Inclusion in AI
4. Democratizing AI though Open Education
5. Participatory, Value-Sensitive & Speculative Design for AI

Suggest other actions …

Aalto University
School of Science

Devising Ethical AI Principles

Action #1: Shaping Ethical AI Principles Collectively



Aalto University
School of Science

AI Finland

www.tekoalyaika.fi/en/background/ethics/

Statement on AI, Robotics and Autonomous Systems

European Group on Ethics in Science and New Tech (EGE)

1) **Human dignity**:
   Limits to classifications & awareness whether we are interacting with a machine or human
2) **Autonomy**:
   Human ability to choose whether to delegate decisions and actions to AI or not
3) **Responsibility:**
   AI should be developed in ways serving social good as determined by democratic processes
4) **Justice, equity, and solidarity**:
   No discriminatory bias in datasets & equal access to AI tech & fair distribution of benefits
5) **Democracy:**
   Key decisions on regulation and application results of democratic and public debate
6) **Rule of law and accountability**:
   Protection agains risks stemming from AI that infringes human rights eg safety and privacy
7) **Security, safety, bodily and mental integrity**:
   All safety dimensions taken into account in development and tested before release
8) **Data protection and privacy:**
   Also limit for tech influencing personal opinions
9) **Sustainability:** Priority for environmental protection

Ethical guidelines for the use of AI

OP Group

1) **People-oriented approach**

   We deploy data and AI responsibly to promote the wellbeing of our customers. We define clearly the goals of our AI work and refine them when necessary to respond to changes in data, technical possibilities and our work environment.

2) **Openness and transparency**

   We operate openly in relation to our customers, partners and stakeholders and ensure the transparency of our AI applications and their evaluation. We are open about the ways we use AI, and we subject our work to review.

3) **Impact assessment**

   We examine carefully how our choices affect our customers and their environments and strive always to make responsible choices when we apply AI.

4) **Ownership**

   We assign owners to all the principles guiding our work and all the algorithms we develop. We ensure that the AI we use is ethical throughout its life cycle.

5) **Privacy protection**

   We safeguard the protection of privacy and personal data in the data we use in accordance with our data protection policies.

The Futurice Principles for Ethical AI

Ethics is an integral part of our way of working. We will always uphold our responsibility to identify and raise ethical implications and concerns related to our work and help our clients deal with ethical questions related to autonomous systems in a responsible way.

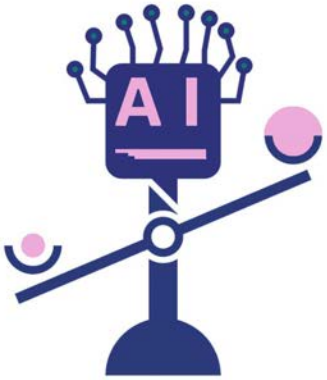
Autonomous systems do what they do as a result of countless technological, economic, ethical and political decisions by human beings. As designers and builders of autonomous systems, we must never relinquish our responsibility for the greater good in the pursuit of business, governmental or political outcomes by us or our clients. We remain committed to retaining human control and the greatest possible degree of transparency in the systems we build.

The following ethical principles are meant to support and guide our decision-making when creating autonomous systems and dealing with data and algorithms.

www.futurice.com/blog/introducing-the-futurice-principles-for-ethical-ai

futurice

futurice



Futurice Principles for Ethical AI

01 Purpose & Impact

Focus on the purpose and impact

- Respect and be mindful about the impact on people affected by the system.
- Ensure that the systems we design and build have a clear purpose and can be trusted to behave as expected and anticipated.
- Consider the impact of the system beyond the user and consider any positive and negative consequences the system might have.

futurice

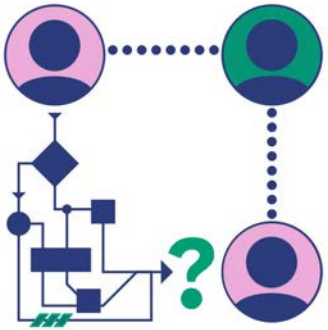
Futurice Principles for Ethical AI

02 Transparency & Trust

Prioritize transparency in the systems we design and build, and strive to increase trust in all of them

- Go for maximum transparency and openness in the systems whenever possible.
- Be mindful about how the system impacts people's behavior.
- When being able to justify the system's working principles and outcomes is paramount, make sure to design and build in explainability from the beginning.
- Build systems that are ready for auditing.

futurice



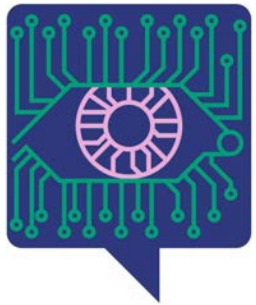
Futurice Principles for Ethical AI

03 Inclusion & fairness

Aim for inclusion by striving to understand whom the system we are designing and building will impact.

- Design the system carefully from the beginning with input from as diverse a group of people as possible.
- Avoid creating or reinforcing bias that can lead to unfair outcomes.
- Use diverse/inclusive training and test data to ensure fairness and inclusivity.
- Make sure to create use cases that represent all impacted people.

futurice



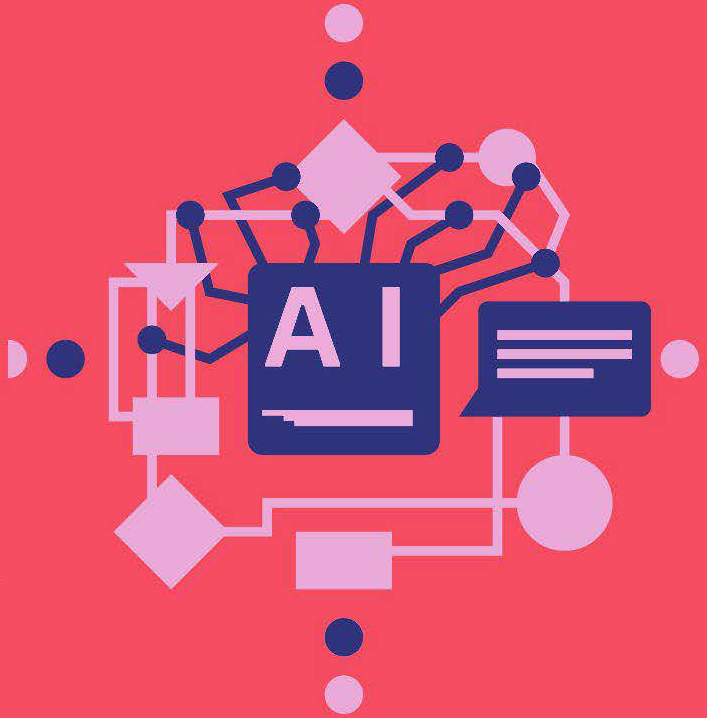
Futurice Principles for Ethical AI

04 Privacy & safety

Collect, store and use personal data safely and default to high privacy

- Make it explicit to users what kind of personal data is being used and how.
- Collect and store as little sensitive data as possible.
- Make it as easy as possible for users to exercise their rights for data privacy (GDPR)
- Anonymise data as much as possible.

futurice



05 Don’ts

Don’t work on systems that go against human rights

Don’t manipulate

- Do not use private data to promote ideas or actions that impacted people might consider unwanted or harmful.
- Do not use manipulative features or design, or exploit human biases – instead, design for understanding.

Don’t harm humans or the environment

- The systems we build should never raise a direct threat towards people or the environment. The systems webuild must always guarantee the protection of the physical, psychological as well as social safety of individuals.

Don’t incite violence

- Violence is sparked by disrespect and distrust between individuals and groups. The systems we build should never promote the division of societies or social groups.

www.theodi.org/article/data-ethics-canvas

Aalto University
School of Science

Oversight of AI

Action #2: Establishing AI Ethics Councils for Oversight & Governance

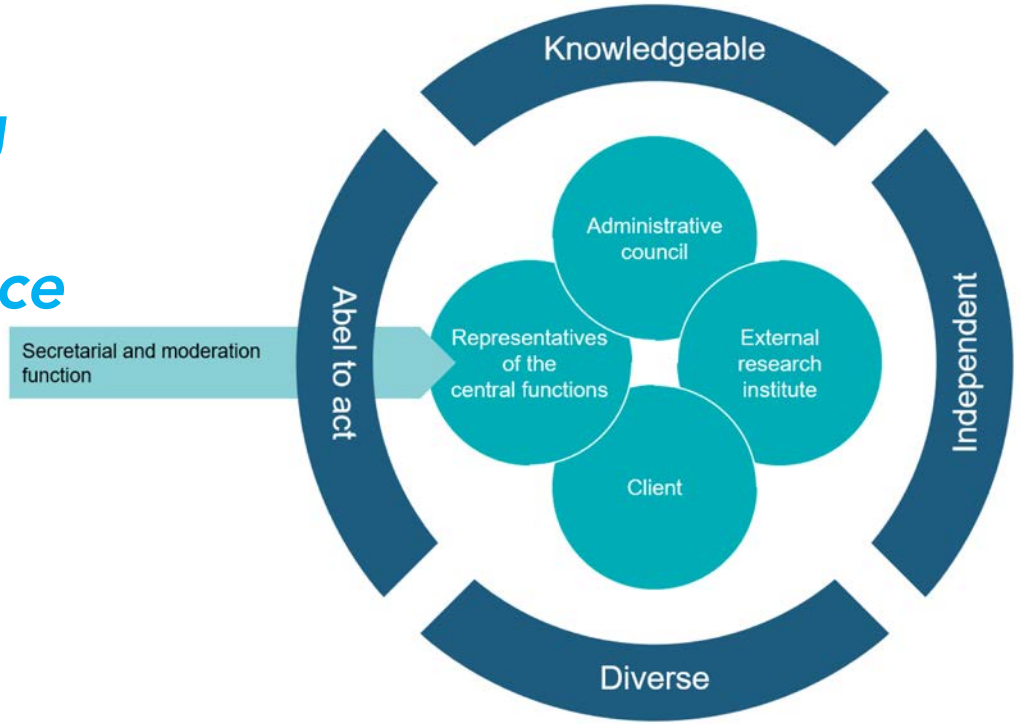


Figure 1: A potential setup of an AI ethics council

Aalto University
School of Science

Markus Schmitz, *Artificial Intelligence and Data Ethics IV*, CIONET, January 14, 2020.

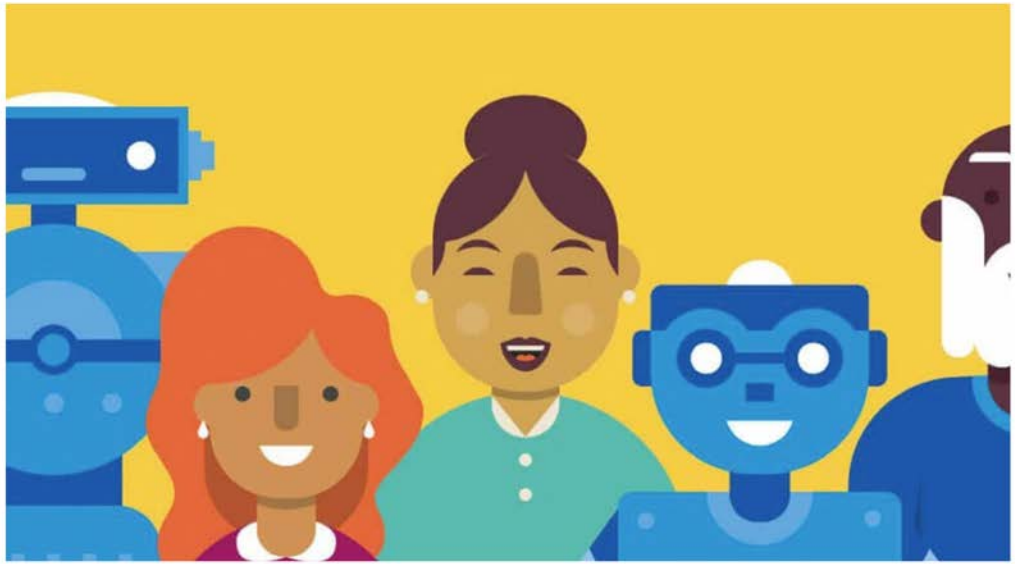
Diversity and Inclusion in AI

Action #3: Creating initiatives for diversity, inclusion & expanded participation in AI

Diversity is the "what"; inclusion is the "how." Inclusion is a measure of culture that enables diversity to thrive.

Diversity & Ethics In AI: A Reflection Of Its Designer

by **Chandana Madaka** on Jul 4, 2019



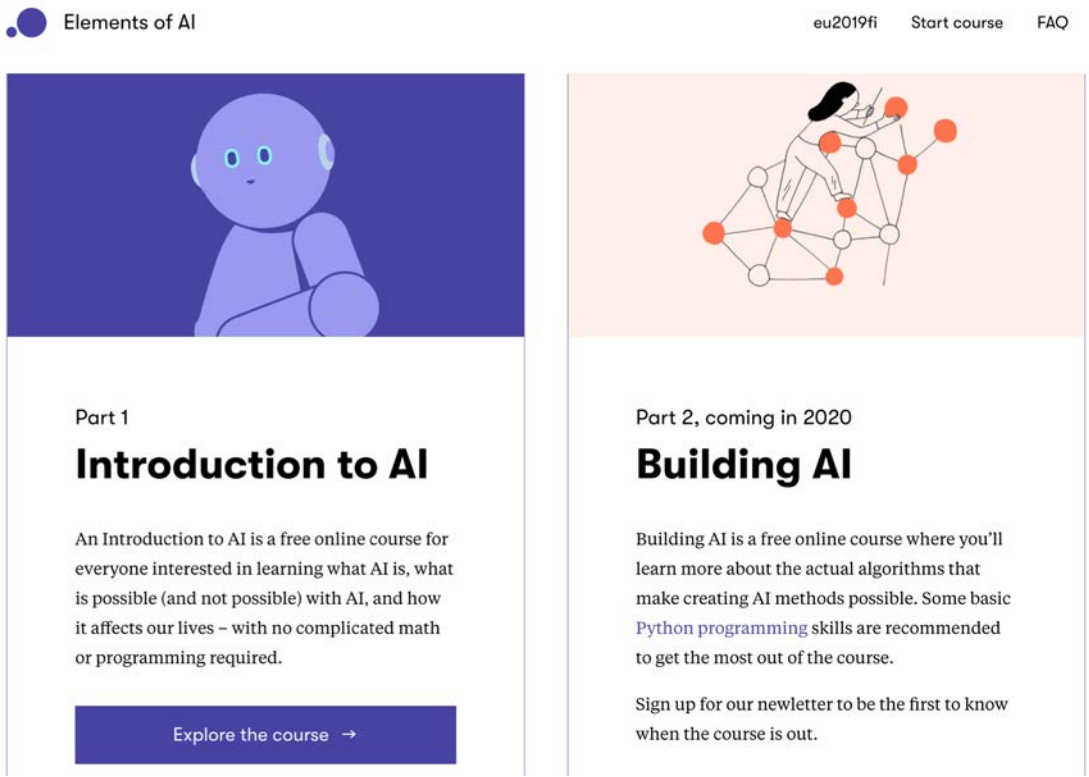
https://blog.strands.com/ethics-diversity-in-ai

Aalto University
School of Science

Democratizing AI

Action #4: Expanding Awareness, Open Education & Positions in AI & Ethics

Founder Teemu Roos, University of Helsinki – to raise awareness about the opportunities and risks of AI among people who are strangers to science, so that they can decide for themselves what uses of AI would be beneficial and where they want their government to invest.

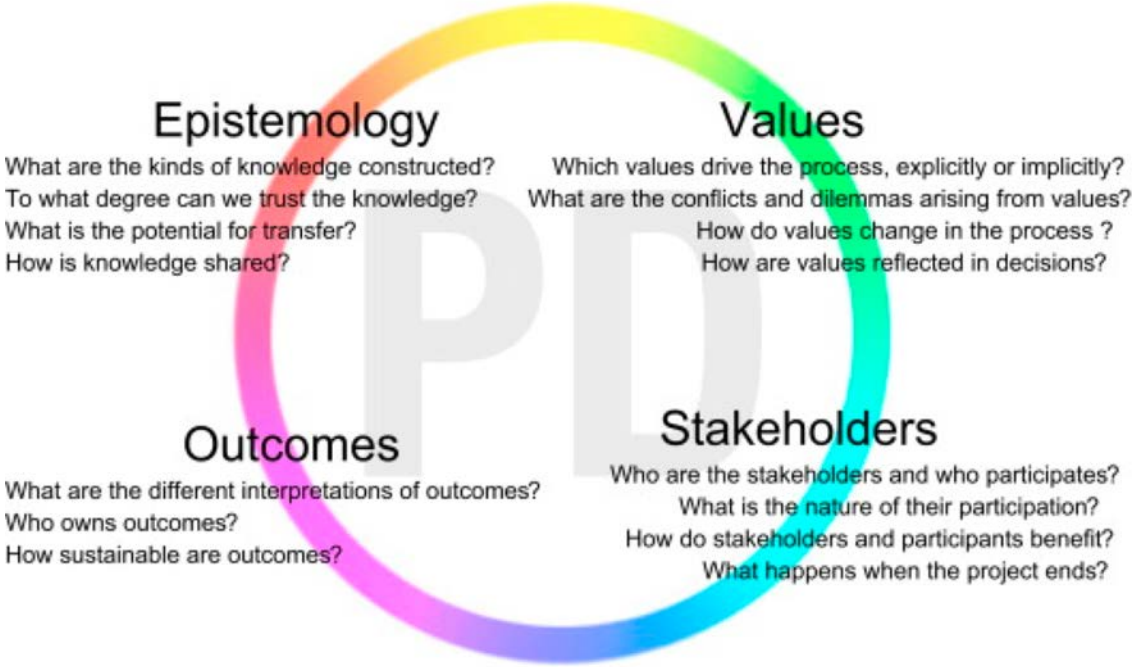


https://www.elementsofai.com

Aalto University
School of Science

Participatory, Value-Sensitive & Speculative Design for AI

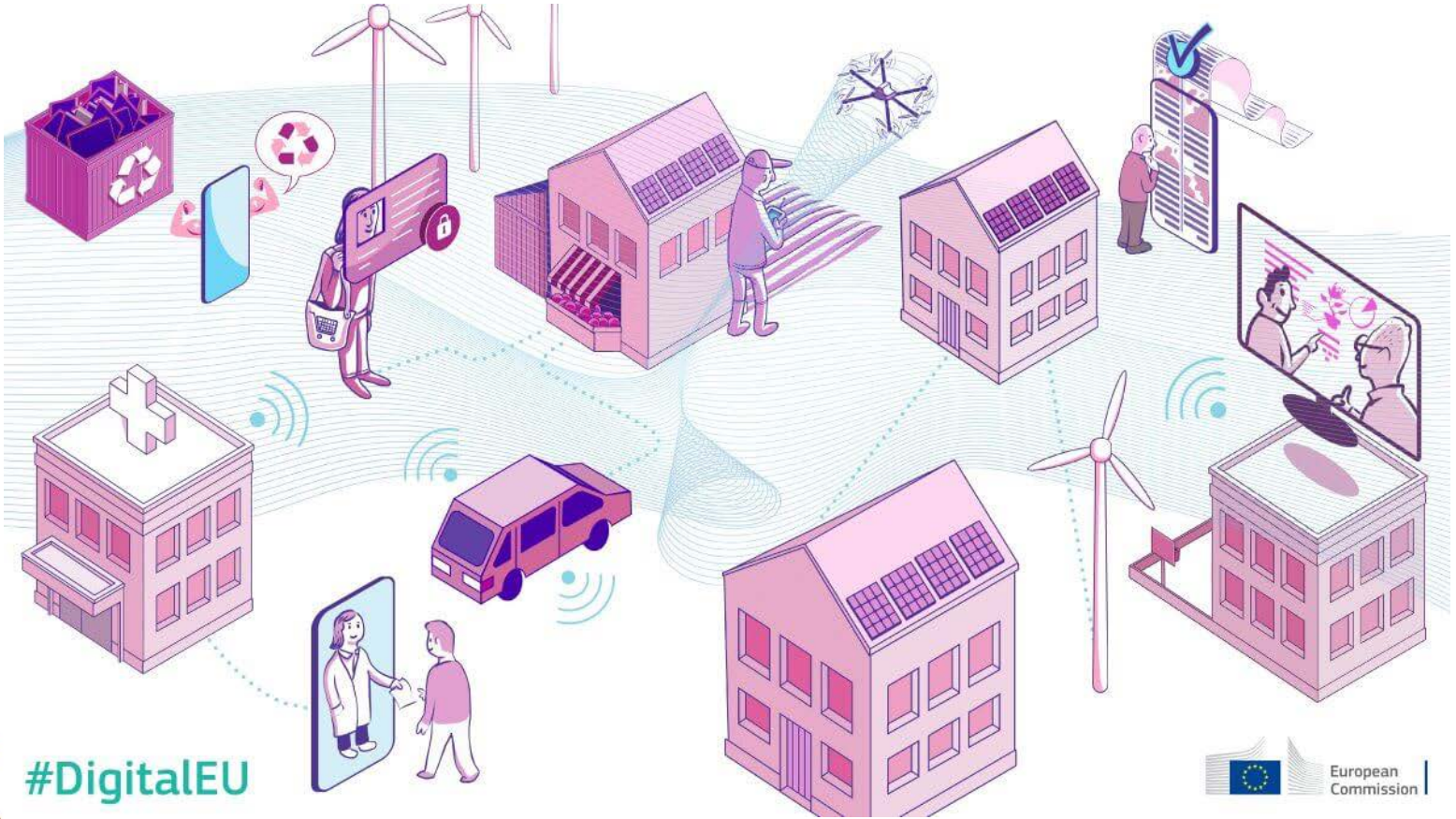
Action #5: Emphasizing the Design Process



Frauenberger, C. et al. In pursuit of rigour and accountability in participatory design. International Journal of Human Computer Studies, Feb 2015.

Aalto University
School of Science

Mapping AI Ethics & Designing for Ecosystems

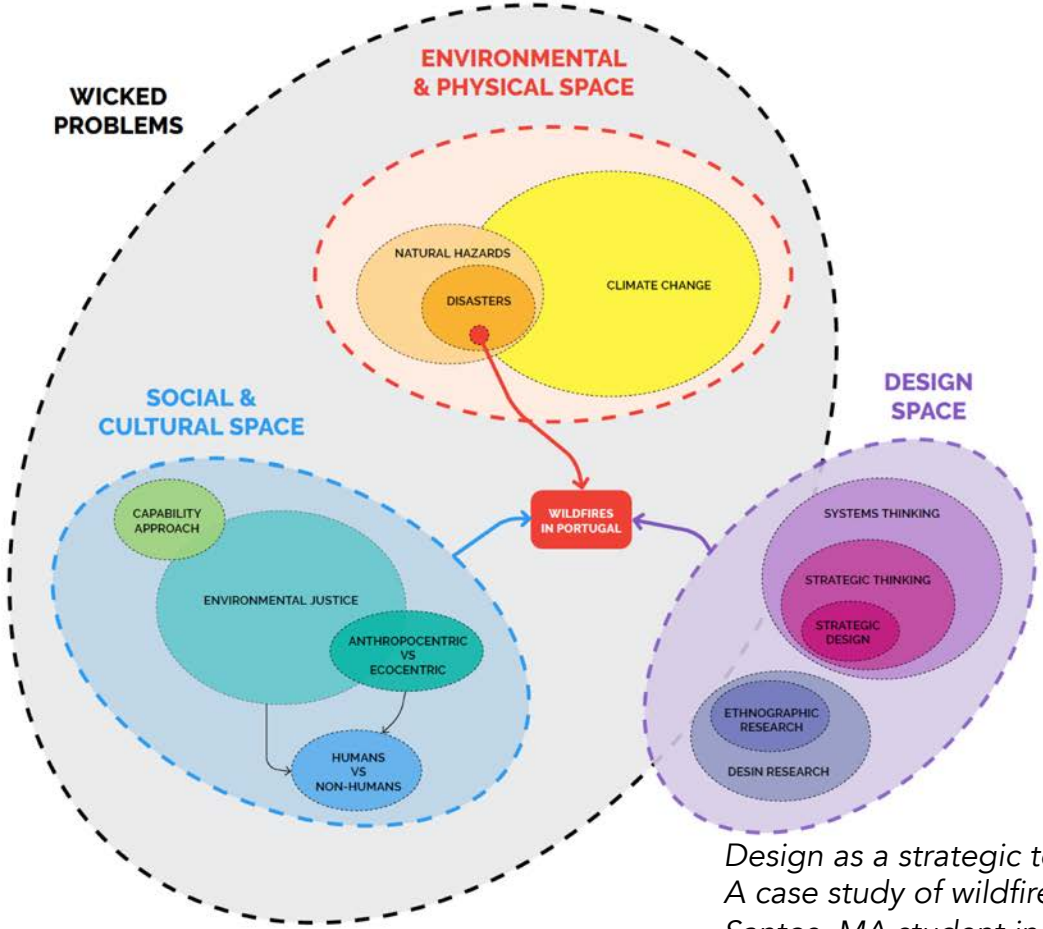


Aalto University
School of Science

#DigitalEU

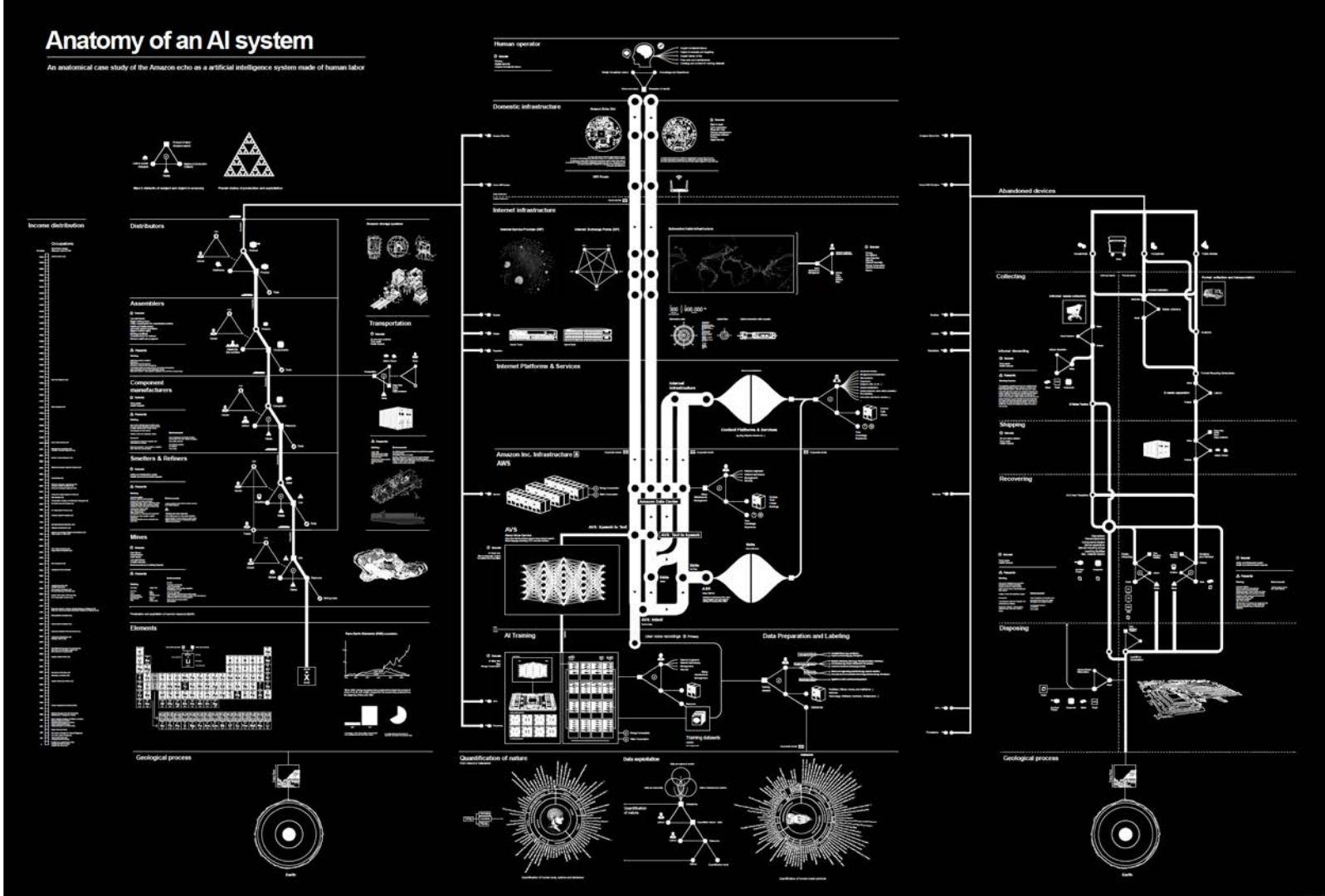
European Commission

Mapping Ecosystems



Design as a strategic tool in disaster prevention: A case study of wildfires in Portugal by André Santos, MA student in Collaborative and Industrial Design, Aalto University, 2020.

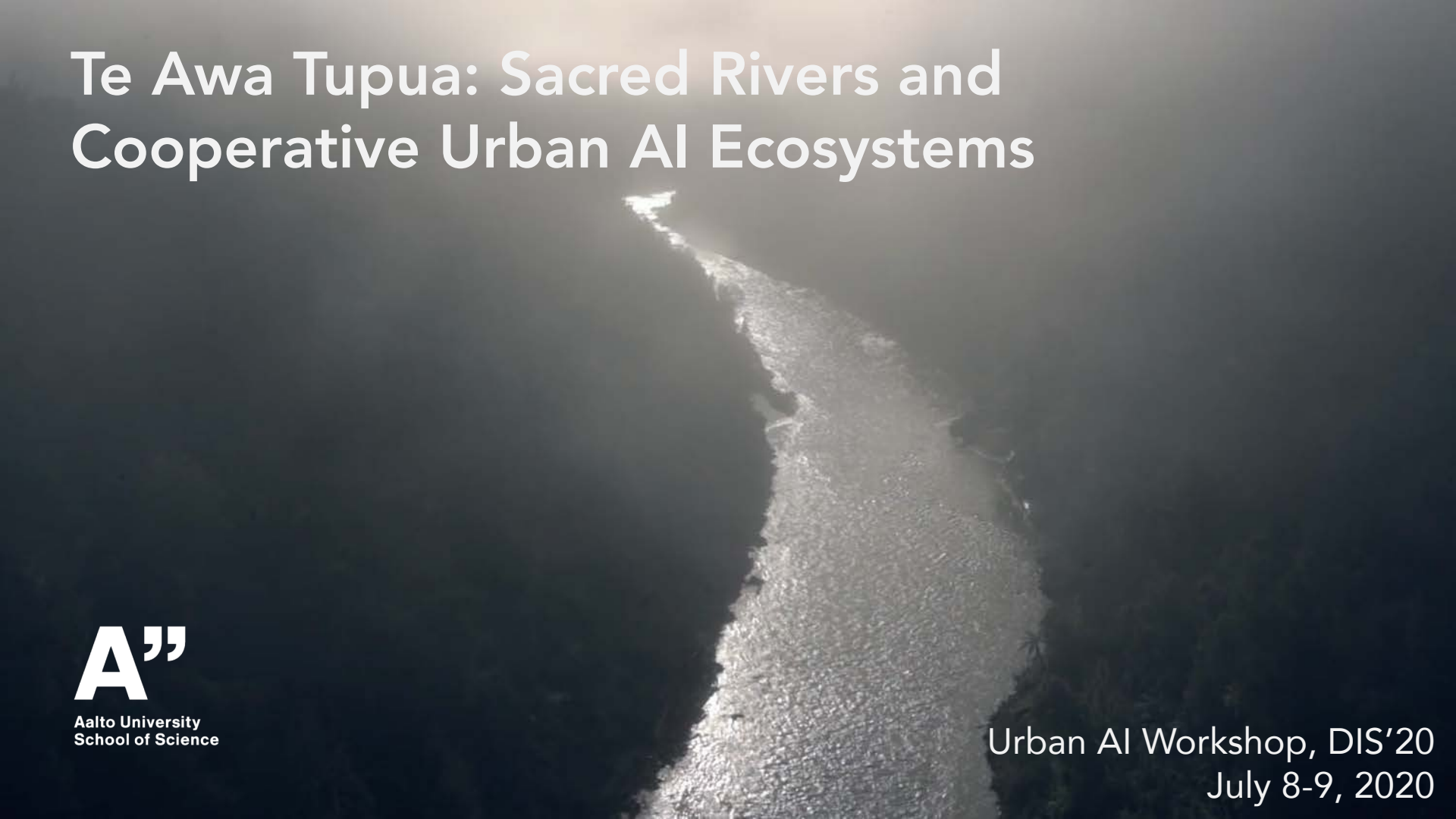
Aalto University
School of Science



Vladan Joler & Kate Crawford, AI Now Institute, 2018.
www.anatomyof.ai

Aalto University
School of Science

Te Awa Tupua: Sacred Rivers and Cooperative Urban AI Ecosystems



Aalto University
School of Science

Urban AI Workshop, DIS'20
July 8-9, 2020

Urban AI Ecosystems in Crisis

Multiple intersecting crises in cities

1. Climate change, inequity, affordable housing and inclusive employment for poor marginalized residents and migrant communities.

2. Unprecedented adverse health and economic effects of the emerging COVID-19 pandemic.

3. Systemic discrimination and violence against Blacks and other historically marginalized groups due to structural racism.

Aalto University
School of Science

Responsible Urban AI in Crisis

Rights-based Discourse

1. **Right to the City, Right to Livelihoods**: Sustaining Urban Ecosystems in Mediating Climate Crises.

2. **Right to Information, Right to Privacy**: Contestations in the COVID 19 Pandemic.

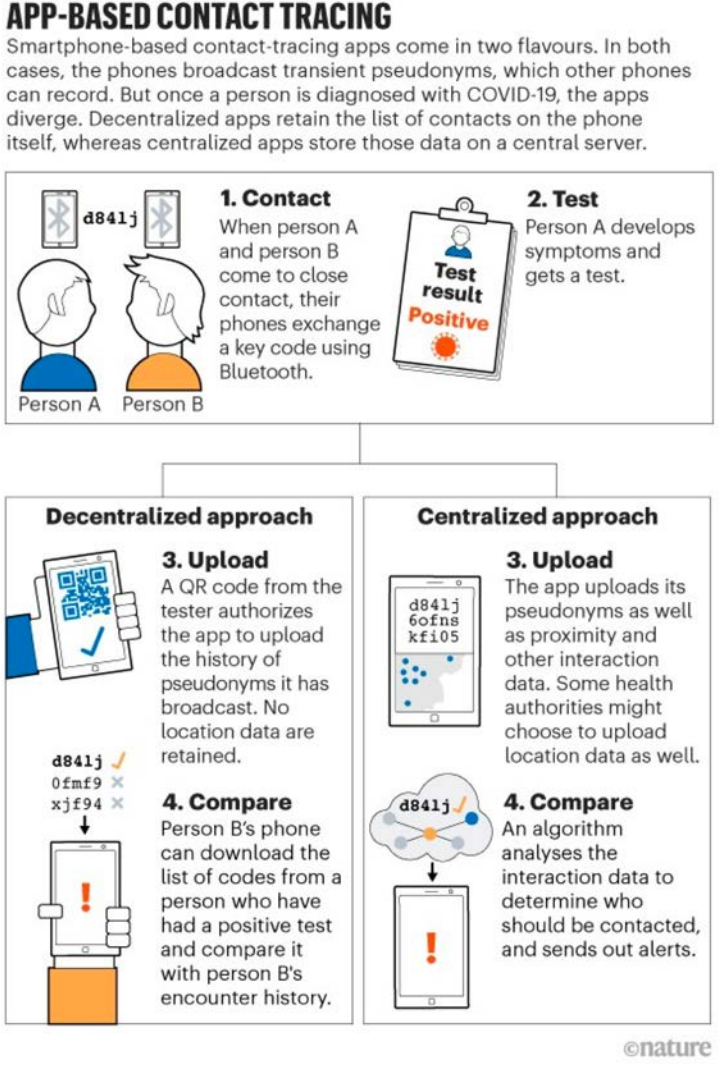
3. **Right to Equality and Non-Discrimination, Right to Free Assembly**: Racial Profiling, Predictive Policing, and Urban Protests.

Nitin Sawhney. *Te Awa Tupua: Sacred Rivers and Cooperative Urban AI Ecosystems*. DIS 2020.

Aalto University
School of Science

Contact Tracing

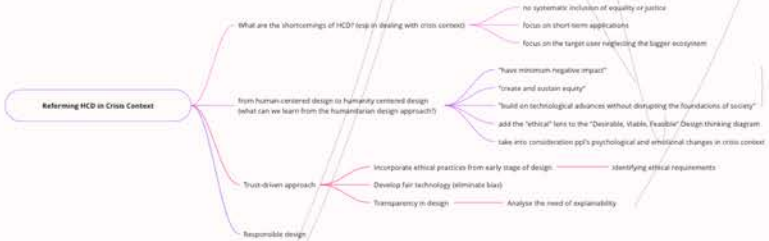
www.cdc.gov/coronavirus



Aalto University
School of Science

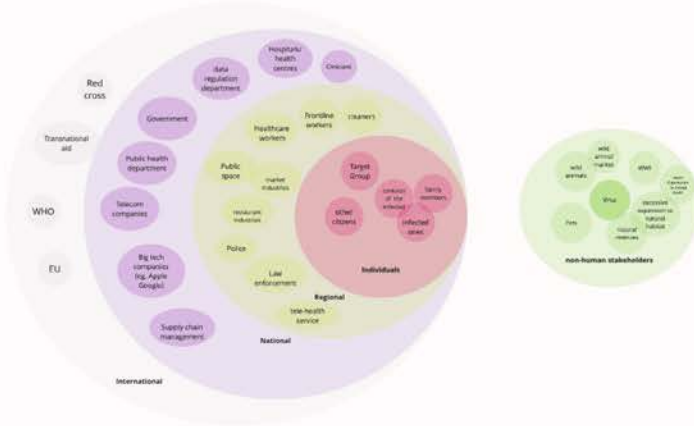
Reforming HCD in Relation to Challenges in Contact Tracing Technologies

Refraring the Topic



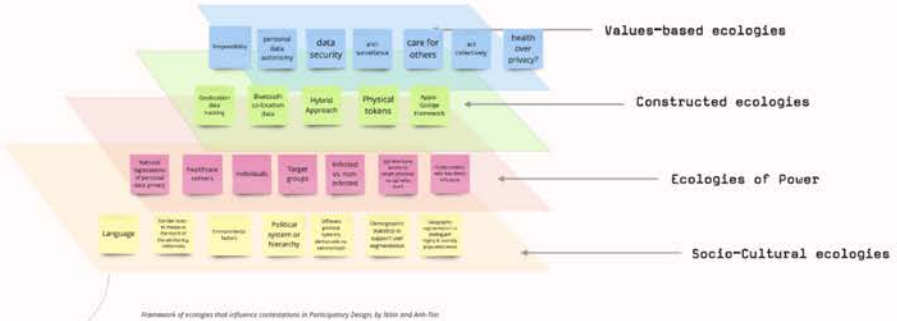
Mapping The Key Stakeholders

Key stakeholders of the Pandemic



Overall Ecology of The Crisis Domain

in the context of dealing with the challenges in CT tech



Framework of ecologies that influence contestations in Participatory Design, by Stein and Ann-Tim

Concerns/Challenges

Opportunities

Emerging Outcomes

Enabling a CT system that ...

- Build trust and empower individuals (Instead of top-down and centralized)
- Convey the proactive data protection method
- Be transparent about data-collection
- Highlight the explainability of the technology
- Be human-centered, bringing up the physiological aspect in crisis contect

Beyond Contact Tracing....

- Is contact tracing the ultimate way?
- Think about larger questions of what pandemics means to the society
- Onto building a resilient society, how to live with virus in a long term?

miro.com/app/board/o9J_kpEI43U=

A VOICE FOR NATURE

The Whanganui River in New Zealand is a legal person. A nearby forest is too. Soon, the government will grant a mountain legal personhood as well. Here's how it happened, and what it may mean.

BY **KENNEDY WARNE**

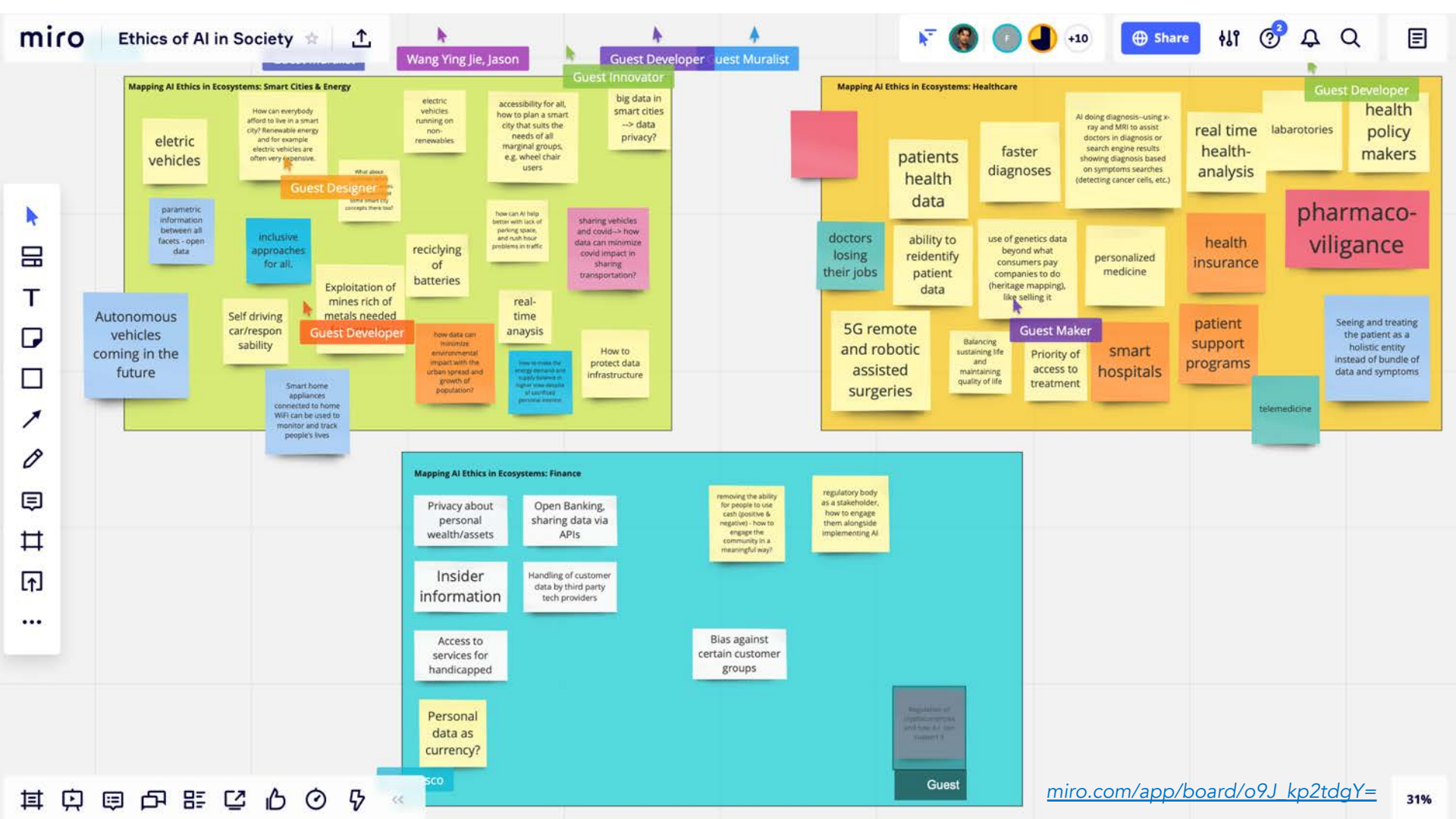
PHOTOGRAPHS BY **MATHIAS SVOLD**

Responsible Urban AI in Crisis

Ways Forward

1. Anchor Ethics in AI using a rights-based discourse.
2. Conduct participatory, value-sensitive and ethically-engaged research and design practices.
3. Develop Cooperative Urban AI Ecosystems embracing human, AI and non-human agency.
4. Devise inclusive policies for oversight, open access, and governance that support ecological justice (social & environmental) in urban sphere.

Aalto University
School of Science

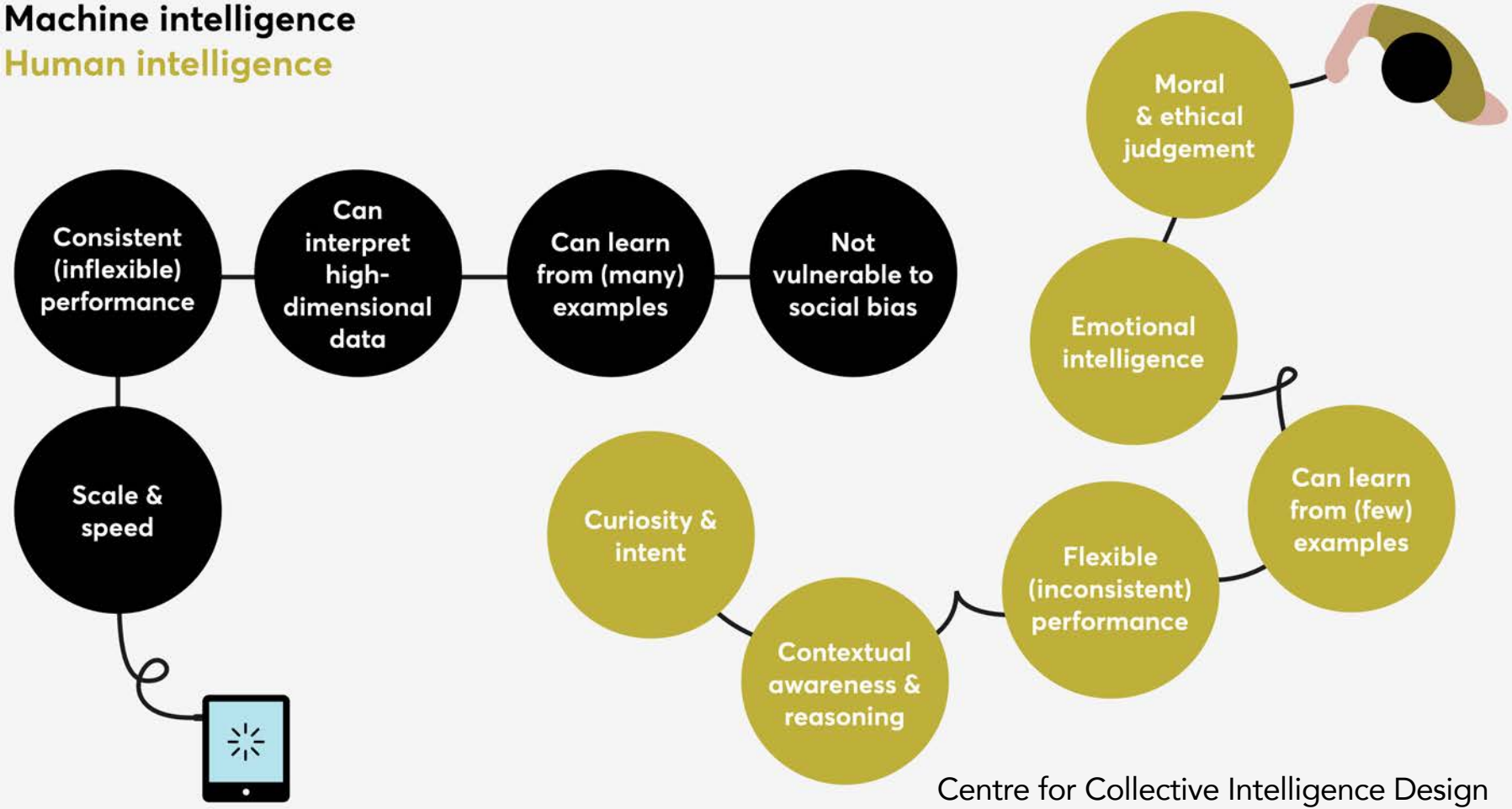

miro
Ethics of AI in Society
Share
Wang Ying Jie, Jason
Guest Developer
Guest Innovator
Guest Designer
Guest Developer
Guest Maker
Guest
Mapping AI Ethics in Ecosystems: Smart Cities & Energy
eletric vehicles
How can everybody afford to live in a smart city? Renewable energy and for example electric vehicles are often very expensive.
electric vehicles running on non-renewables
accessibility for all, how to plan a smart city that suits the needs of all marginal groups, e.g. wheel chair users
big data in smart cities --> data privacy?
parametric information between all facets - open data
inclusive approaches for all.
reciclying of batteries
how can AI help better with lack of parking space, and rush hour problems in traffic
sharing vehicles and covid--> how data can minimize covid impact in sharing transportation?
Autonomous vehicles coming in the future
Self driving car/respon sability
Exploitation of mines rich of metals needed
how data can minimize environmental impact with the urban spread and growth of population?
real-time anaysis
How to protect data infrastructure
Smart home appliances connected to home WiFi can be used to monitor and track people's lives
Mapping AI Ethics in Ecosystems: Healthcare
patients health data
faster diagnoses
AI doing diagnosis--using x-ray and MRI to assist doctors in diagnosis or search engine results showing diagnosis based on symptoms searches (detecting cancer cells, etc.)
real time health-analysis
labarotories
health policy makers
pharmaco-viligance
doctors losing their jobs
ability to reidentify patient data
use of genetics data beyond what consumers pay companies to do (heritage mapping), like selling it
personalized medicine
health insurance
5G remote and robotic assisted surgeries
Balancing sustaining life and maintaining quality of life
Priority of access to treatment
smart hospitals
patient support programs
telemedicine
Seeing and treating the patient as a holistic entity instead of bundle of data and symptoms
Mapping AI Ethics in Ecosystems: Finance
Privacy about personal wealth/assets
Open Banking, sharing data via APIs
removing the ability for people to use cash (positive & negative) - how to engage the community in a meaningful way?
regulatory body as a stakeholder, how to engage them alongside implementing AI
Insider information
Handling of customer data by third party tech providers
Access to services for handicapped
Bias against certain customer groups
Personal data as currency?
miro.com/app/board/o9J_kp2tdgY=
31%

Civic AI using Collective Intelligence

How can AI and collective intelligence enable new forms of community response to the climate crisis?

Centre for Collective Intelligence Design

https://www.nesta.org.uk/project/centre-collective-intelligence-design/



Centre for Collective Intelligence Design

https://www.nesta.org.uk/project/centre-collective-intelligence-design/

MIT CENTER FOR COLLECTIVE INTELLIGENCE

Designing Collectively Intelligent Systems

Collective Intelligence Design Lab

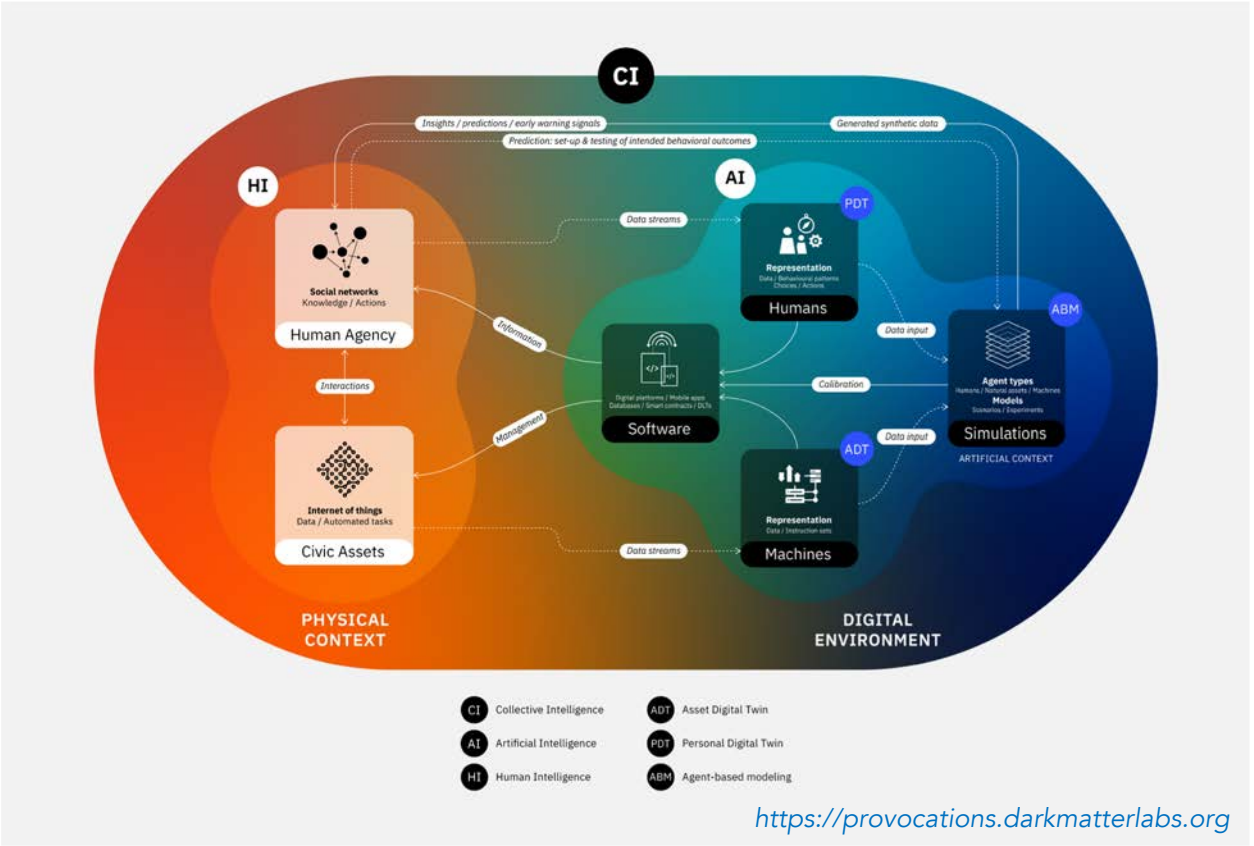
https://cci.mit.edu/research/

MIT's Collective Intelligence Design Lab (CIDL) helps groups design innovative new kinds of collectively intelligent systems (superminds) to solve important problems.

Aalto University School of Scien

Collective Intelligence

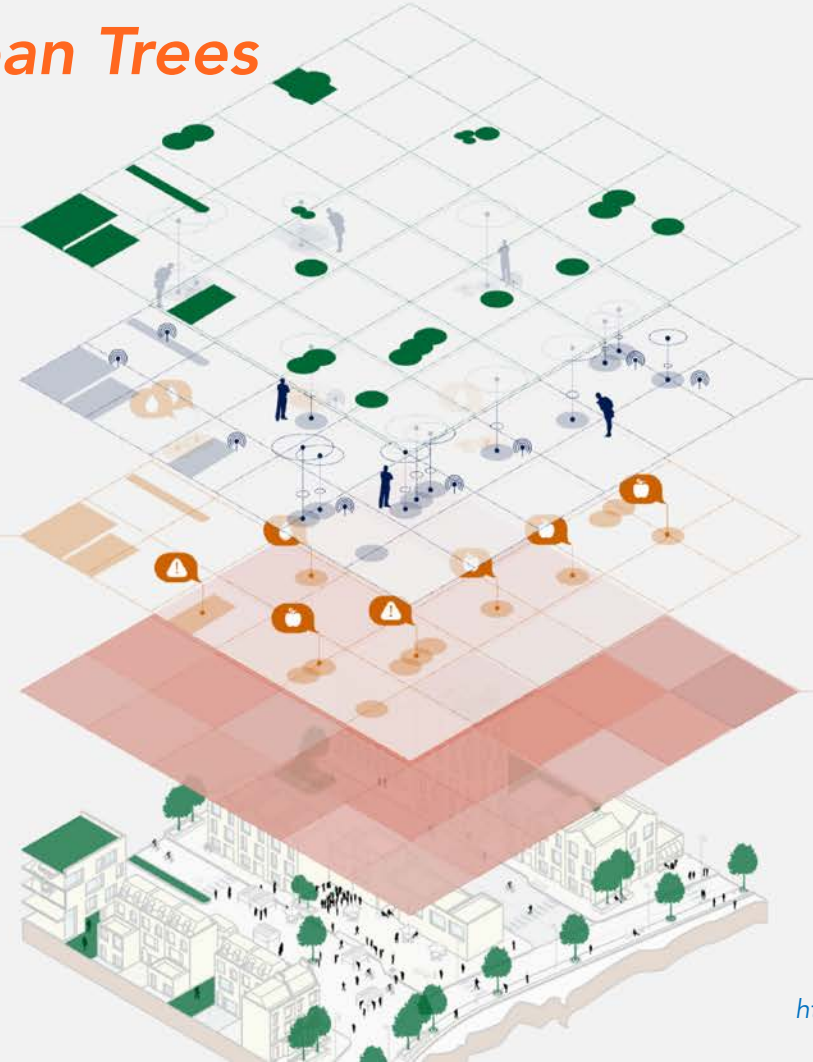
Augmented Collective Intelligence Framework



https://provocations.darkmatterlabs.org

Aalto University
School of Science

Use Case: Urban Trees



MAPPING

Using machine learning algorithms to automatically identify location & type of green infrastructure from satellite imagery

Citizens verify data to help train learning algorithm

MAINTAINING

Distributed sensors used to help track soil nutrients and maintenance needs

AI agents notify citizens if nearby trees need care or if fruit is ready to pick

MONITORING

Distributed sensors track ground conditions: tree health, storm runoff & air temperature

Satellite imagery & historic weather forecasts used to monitor impact on local temperature

AI agents prompt citizen to verify data & provide local measurements: tree size, species biodiversity etc.

MODELLING

Large-scale analysis of outcomes and impact data can be used to provide evidence for outcomes-based investment

Agent based simulations used to model outcomes of alternative greening strategies

https://provocations.darkmatterlabs.org

Use Case: Collective Climate Action

SENSE-MAKING & IDENTIFYING

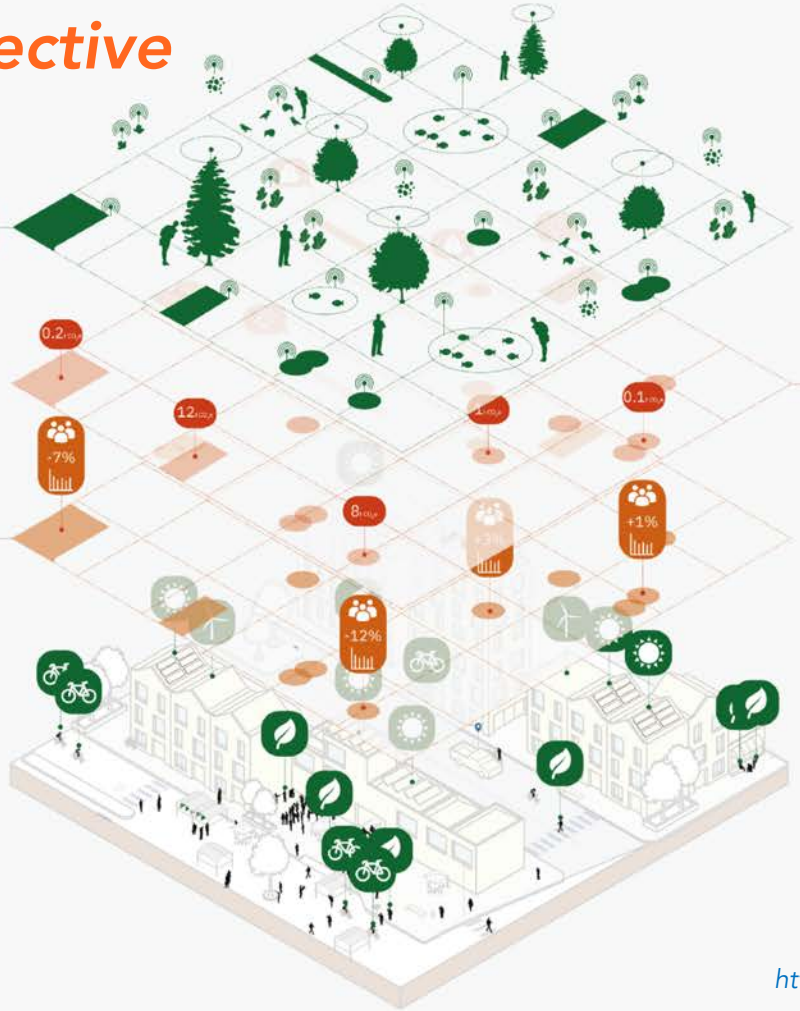
Building more-than human shared understanding & list of potential interventions

- Machines as agents of non-human perspectives and civic assets helping leverage data & information to augment shared understanding
- Citizens collect and verify data to help train the learning algorithm; AI and citizens recommend technically appropriate interventions

FEEDBACK

Provide feedback to sustain behavioural change

- Edge AI is used to locally process data that reflects a citizen's lifestyle, then provides personalised feedback and means for social influence on climate-positive actions
- A dashboard is used to provide transparency of citizens' collective effort
- Community peer-to-peer learning and influence ensure prolonged behavioural change



SIMULATING POTENTIAL IMPACT

Calculating potential impact of choices to encourage behavioural change

- Simulations used to explore what-if scenarios to estimate the potential collective impact of choices, identify the gap between potential impact versus targets, and provide alternative decision pathways
- Citizens explore various pathways and make informed commitments
- Machine learning is applied to improve forecasting based on real-world monitoring

News — 11 June, 2020

HOT and iLab Liberia Exploring the Potential of Machine Learning to Augment Human Mappers in Monrovia

Countries

Liberia

Associated Projects

Scaling Machine-Learning Workflows for Participatory Mapping: Monrovia ML Challenge

Aalto University School of Scienc

https://www.hotosm.org

The Future of Minds and Machines: How artificial intelligence can enhance collective intelligence

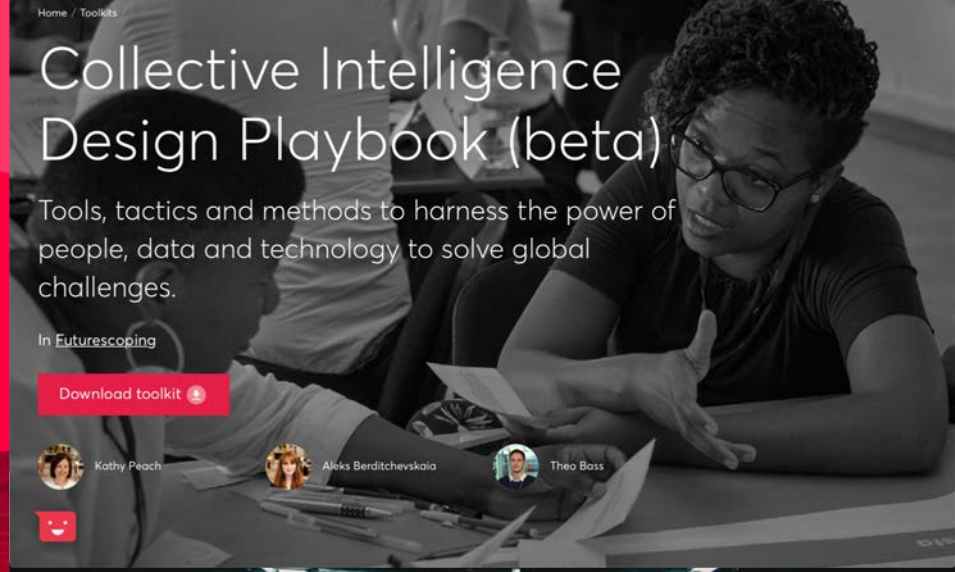
Home / Blogs /

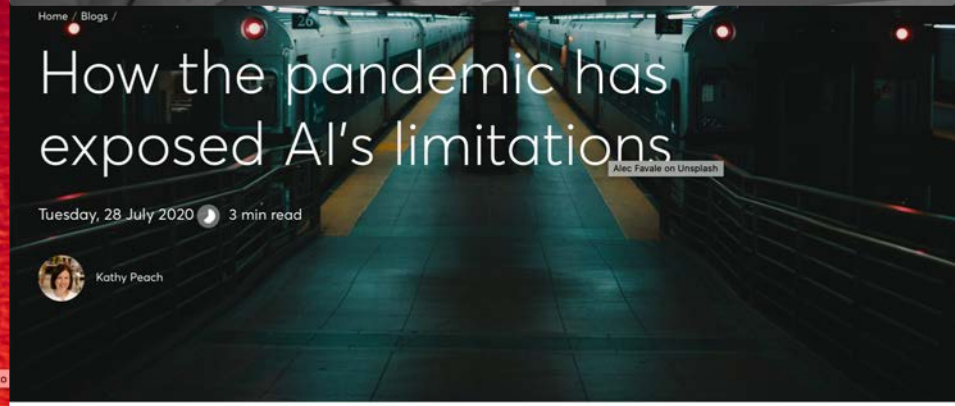
Five ways collective intelligence can help developing countries face COVID-19

As developing nations gear up to tackle COVID-19, these ideas for collective intelligence projects might provide some support

Photo

Tuesday, 21 April 2020 | In Health 4 min read





https://www.nesta.org.uk/project/centre-collective-intelligence-design/

It should have been artificial intelligence's moment in the sun. With billions of dollars of investment in recent years, AI has been touted as a solution to every

Decolonizing AI & Rethinking Resistance

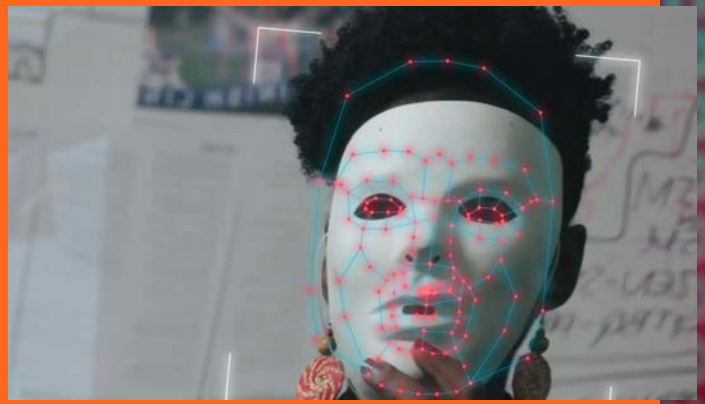
Nitin Sawhney, Ph.D.

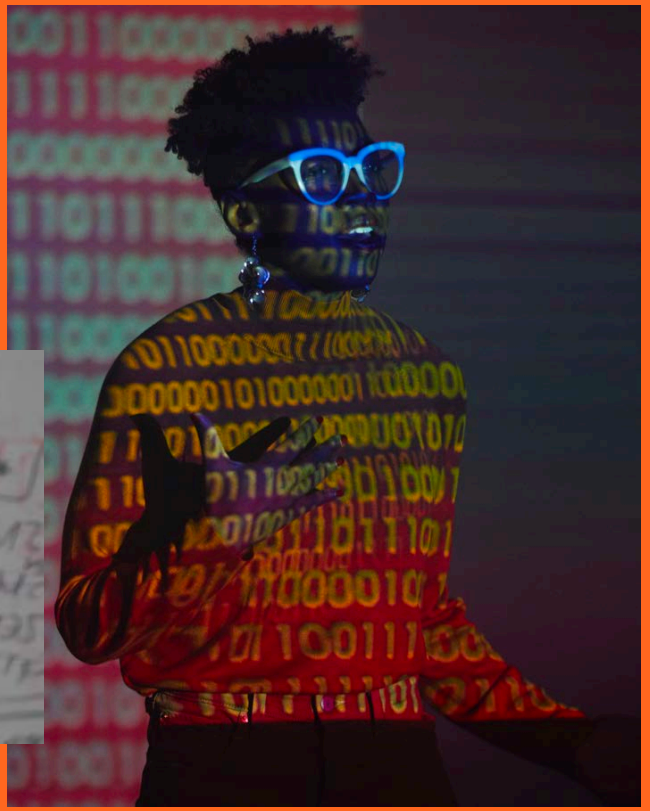
Professor of Practice

August 7, 2020

Department of Computer Science

Aalto University
School of Science





Human-Centred Research & Design in Crisis