AI: RESPONSIBILITY IN A CHANGING WORLD

Prof. Dr. Virginia Dignum

Chair Responsible AI – Director AI Policy Lab

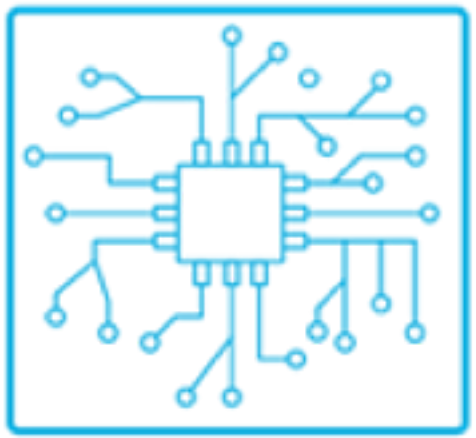
Email: virginia@cs.umu.se

UMEÅ UNIVERSITY

AI Policy Lab

at Umeå University

AI – MORE IS BETTER?

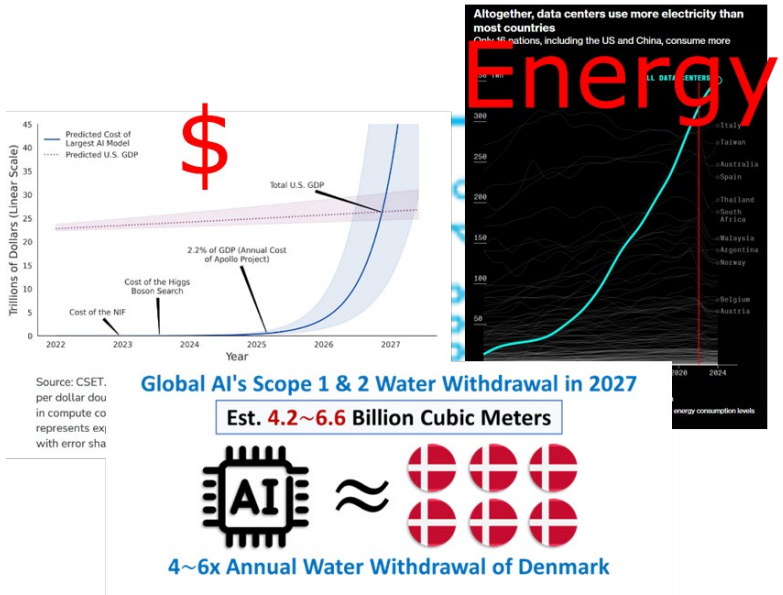


Computing Power

Algorithm Power

Data Availability

AI – MORE IS BETTER?



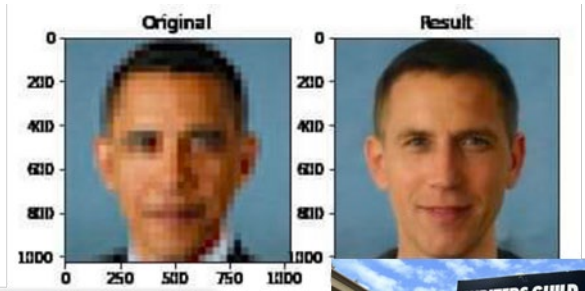
Global AI's Scope 1 & 2 Water Withdrawal in 2027

Est. 4.2~6.6 Billion Cubic Meters

AI ≈

4~6x Annual Water Withdrawal of Denmark

Computing Power



Google exploited homeless black people to develop the Pixel 4's facial recognition AI

Russia Tests New Disinfor… Tactics in Africa to Expand

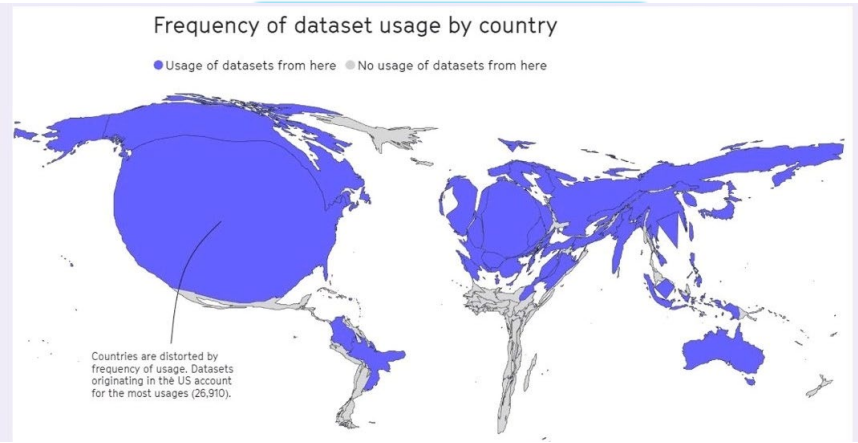
Amazon's facial recognition matched 28 members of Congress to criminal mugshots

Flawed Algorithms Are… Millions of Students' Es…

WHAT HAPPENS WHEN AN ALGORITHM CUTS YOUR HEALTH CARE

Indigenous elder slams 'hollow a… tokenistic' consultation by Sidew…

Algorithm Power



Data Availability

AI Policy Lab
at Umeå University

Overcoming Racial Bias In AI Systems And Startlingly Even In AI Self-Driving Cars

Racial patien

AI expert calls for end to UK use of 'racially biased' algorithms

AI E Live Reg

Gender bias in AI: building fairer algorithms

Bias in AI: A problem recogn still unresolved

Amazon, Apple transcribing bl AI voice recogn

Millions of black people affected by racial bias in health-care algorithms

When It Comes to Gori

Google exploited homeless black people to develop the Pixel 4's facial recognition AI

Russia Tests New Disinformation Tactics in Africa to Expand Influen

Amazon's facial recognition matched 28 members of Congress to criminal mugshots

Flawed Algori

AI IMPACT MEANS HUMAN RESPONSIBILITY

AI
at Umeå University

AI IS NOT INTELLIGENT

DIFFERENT

- Correlation rather than causal mechanisms
- A language 'Frankenstein monster'
  - Incapable of moral thinking, to determine what ought and ought not to be
  - Incapable of distinguishing the possible from the impossible

AI IS NOT ARTIFICIAL

JUST

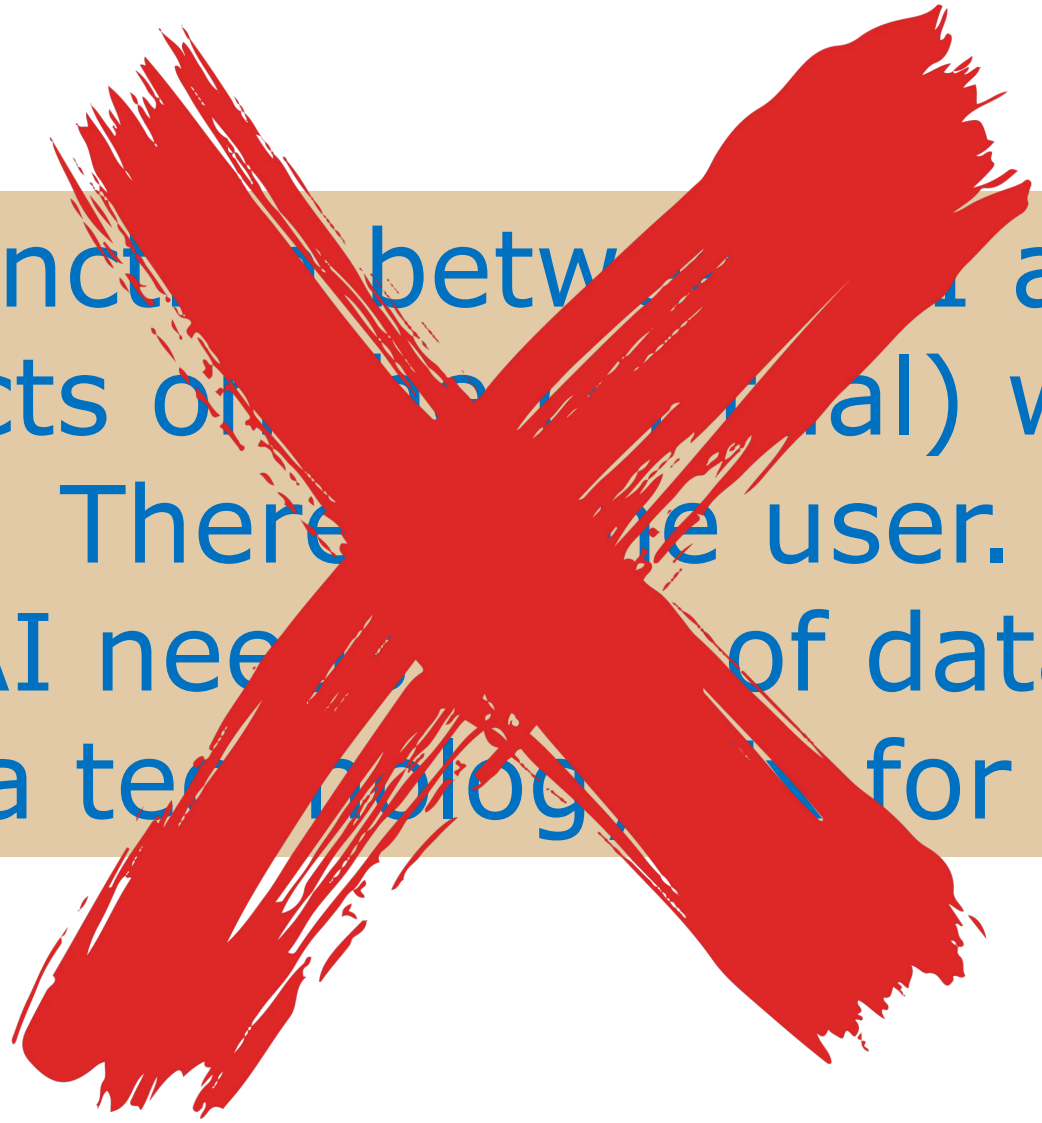
- Built by people for a given purpose
- Dependent on the labor of many
- Using natural resources

Manipulation of language is not a proxy for intelligence!

https://fof.se/artikel/overtygande-sprak-ar-inget-belagg-for-intelligens/

AI Policy Lab
at Umeå University

WHAT IS AI?

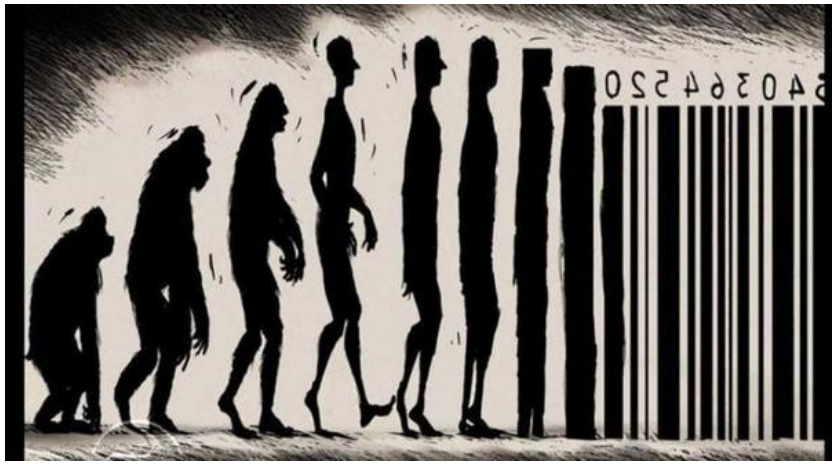


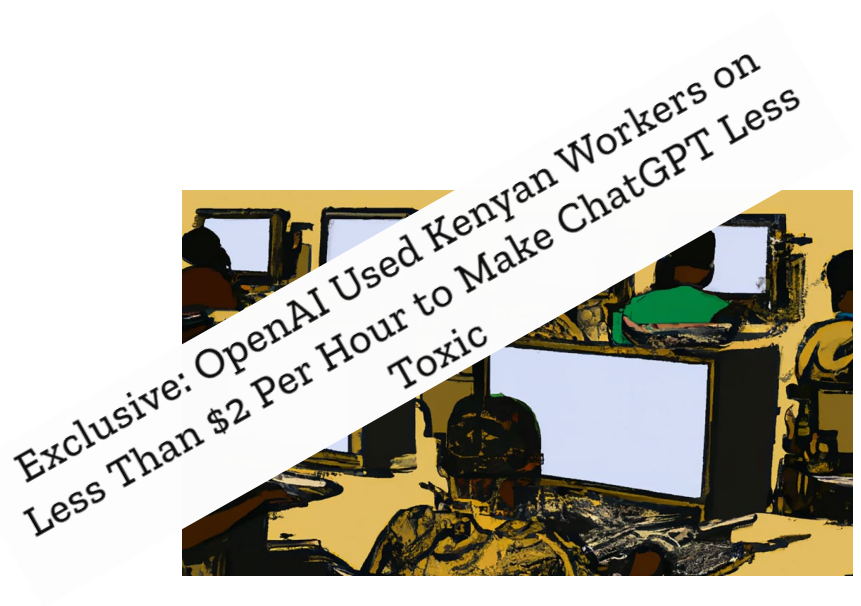
Clear distinct[illegible] betw[illegible] and not AI.
AI acts o[illegible]al) world.
There[illegible]e user.
AI nee[illegible] of data.
There is a te[illegible]olog[illegible] for ill effects!

AI Policy Lab
at Umeå University

RESPONSIBLE AI – WHY CARE?

- Datification
  - Reality is more than data
  - Data is constructed
  - Data is biased
  - Data rich, insight poor
- Power
- Sustainability



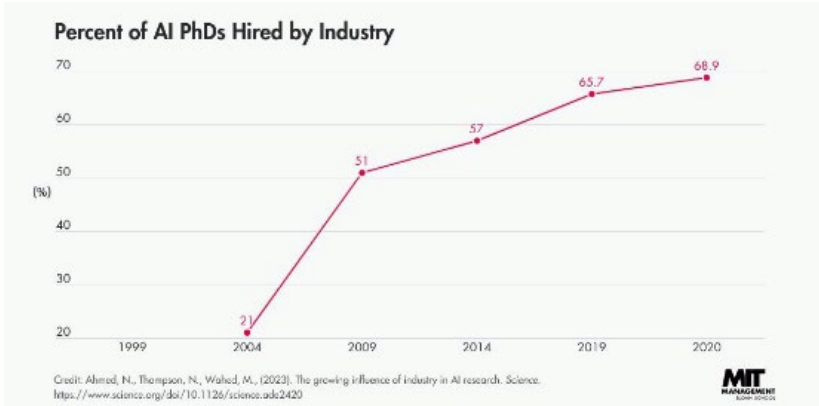


Time magazine, Jan 2023

RESPONSIBLE AI – WHY CARE?

- Datification
- Power
  - Who is developing AI?
  - Who is deciding?
- Sustainability

At Tech's Leading Edge, Worry About a Concentration of Power



- **18% researchers at conferences are women**
- **80% professors are men**
- **Workforce**
  - **Google: 2,5% black, 3,6% Latino, 10% women**
  - **Facebook: 3,8% black, 5% Latino, 15% women**

RESPONSIBLE AI – WHY CARE?

- Datification
- Power
- Sustainability
  - Computational cost of AI
  - Human and social costs

INDEPENDENT PREMIUM

THE LONGER READ

Behind the AI boom, the armies of overseas workers in 'digital sweatshops'

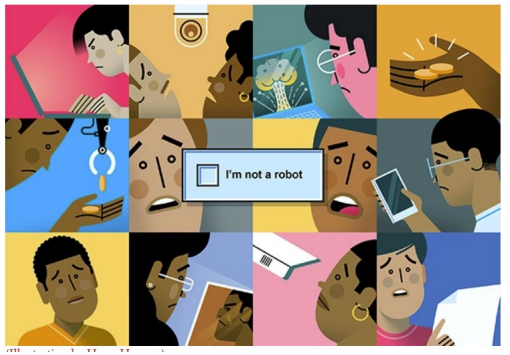
Technology

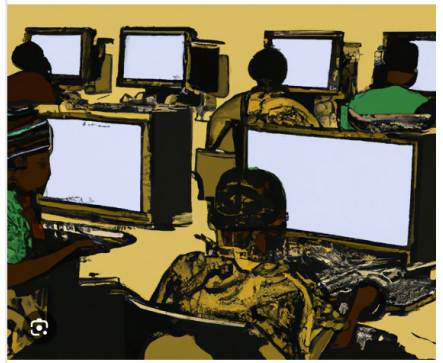
Turk Wars: How AI Threatens the Workers Who Fuel It

The much-hyped AI tools of the future are being built by a globally dispersed army of data workers.

CITE SHARE COMMENT PRINT ORDER REPRINTS

By Krystal Kauffman & Adrienne Williams | Oct. 11, 2023

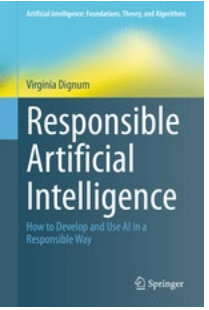




OpenAI Used Kenyan Workers on Less Than $2 Per Hour: Exclusive | TIME

AI Policy Lab
at Umeå University

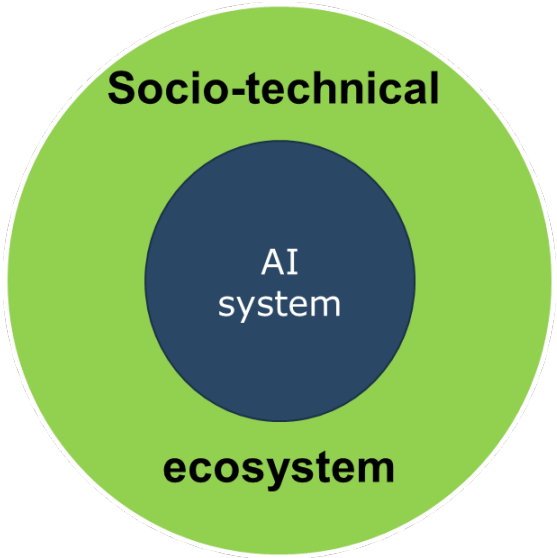
CAN WE BE RESPONSIBLE?



RESPONSIBLE ABOUT AI

AI does not happen to us!

- AI is designed. We make the choices
  - Who should decide?
  - Which values should be considered?
  - Whose values? How to prioritize?
- Ask Question Zero!
  - Is AI the best option here?
  - Ask why before asking how
- AI does not exist in a vacuum
  - There is no technology fix for ill effects!
  - Ethics, regulation, governance concern the ecosystem



Responsible AI solutions need to be social rather than technical!

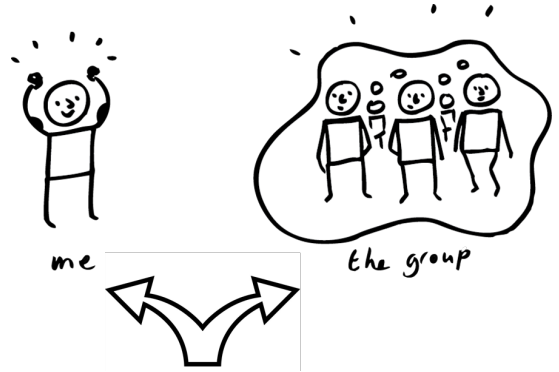
AI Policy Lab
at Umeå University

WHAT IS AI?

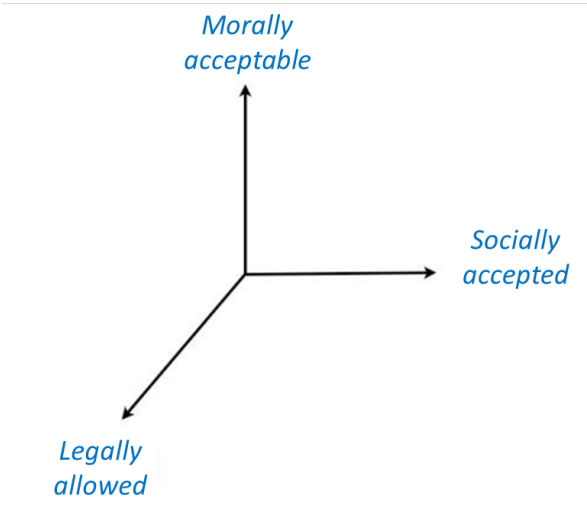
WHAT DO WE WANT AI TO BE?

- Human-like?
  - Why?
  - What does this mean?
- Tool?
  - For what? For who?
- Simulation or operation?
  - Understand intelligence by building intelligence, or
  - Active intervention in real world
- Normative or descriptive?
  - Do as we say or do as we do?

WHICH DECISIONS SHOULD AI MAKE?



HOW SHOULD AI TAKE DECISIONS?



AI Policy Lab
at Umeå University

RAI IS MULTIDISCIPLINARY

understanding and critiquing the intended and unforeseen, positive and negative, socio-political consequences of AI for society in terms of equality, democracy and human rights

- **governance**, not only in terms of competences and responsibilities, but also in terms of **power, trust and accountability**;
- **societal, legal and economic** functioning of socio-technical systems;
- **value-based design** approaches and of ethical frameworks;
- **inclusion and diversity** in design, and how such strategies may inform processes and results;
- **distributed and increasingly ubiquitous nature of AI** applications and developing new scholarly perspectives on human-machine communication.

QUESTION ZERO

"Should this AI system be built at all?”

challenges developers, researchers, and organizations to reflect on the necessity, purpose, and ethical implications of building an AI system before its creation: Ask why before asking how

- Purposeful:
  - Consider societal benefits and potential harms
- Responsible:
  - beyond technical feasibility and economic viability
- Preventive:
  - Prioritizes resource allocation for socially valuable innovations
- Participatory:
  - Involves diverse stakeholders and ensures marginalized voices are heard

DESIGN CHOICES

AI Policy Lab
at Umeå University

DESIGN CHOICES

Choices

Formulation

Information

Involvement

Legitimacy

Aggregation

DESIGN IS POLITICAL

AI Policy Lab
at Umeå University

GOVERNANCE – WHY? WHAT FOR?

- Regulation as incentive for responsible innovation, sustainability, and fundamental human rights
  - powerful stepping stone for innovation with societal benefits
  - signaling expected ambitions enhancing innovation, competitive power

Cars drive faster with brakes

-

In a game without rules, no one wins

  - products, and results
  - Need for better understanding and integration of existing frameworks alongside introducing more regulation
- Avoidance of an "arms race" narrative in AI regulation

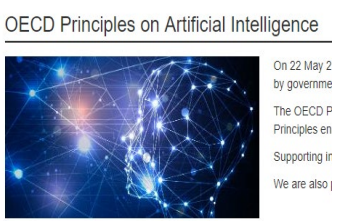
PRINCIPLES AND GUIDELINES

EU HLEG	OECD	IEEE EAD
• Human agency and oversight • **Technical robustness and safety** • Privacy and data governance • **Transparency** • **Diversity**, non-discrimination and fairness • **Societal and environmental well-being** • **Accountability**	• benefit people and the planet • respects the rule of law, **human rights**, democratic values and **diversity**, • include appropriate safeguards (e.g. human intervention) to ensure a **fair and just society**. • **transparency** and responsible disclosure • **robust, secure and safe** • Hold organisations and individuals **accountable** for proper functioning of AI	• How can we ensure that A/IS do not infringe **human rights**? • effect of A/IS technologies on **human well-being**. • How can we assure that designers, manufacturers, owners and operators of A/IS are responsible and **accountable**? • How can we ensure that A/IS are **transparent**? • How can we extend the benefits and minimize the risks of AI/AS technology being misused?

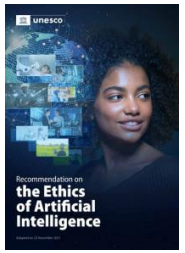
Level	Framework & Reach
Global	UNESCO (194 countries), OECD (>70 jurisdictions), GPAI (25+ members), CoE treaty (50+ countries), G7 principles
Regional	EU AI Act (27 EU states), Santiago Declaration (Latin America/Caribbean)
National	930+ initiatives in 71 countries
Others	200+ guidelines across NGOs, acader private bodies

https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence

https://ethicsinaction.ieee.org



https://www.oecd.org/going-digital/ai/principles/



https://www.unesco.org/en/artificial-intelligence/recommendation-ethics

https://www.un.org/ai-advisory-body

Well over 1000 published worldwide!

AI Policy Lab
at Umeå University

SAFEGUARDING DEMOCRACY

- Strengthening Regulation and Oversight
  - Implement clear laws for AI, facial recognition, cryptocurrency, and spyware.
  - Enhance transparency and accountability in technology deployment.
  - Strengthen democratic institutions to oversee and guide technological development.
- Promoting Public Participation
  - Democratize AI development by involving diverse societal stakeholders.
  - Align technology systems with democratic values and public interest.
- Enhancing Digital Literacy
  - Foster critical understanding of AI and digital technologies through education.
  - Empower citizens to advocate for their rights in the digital age

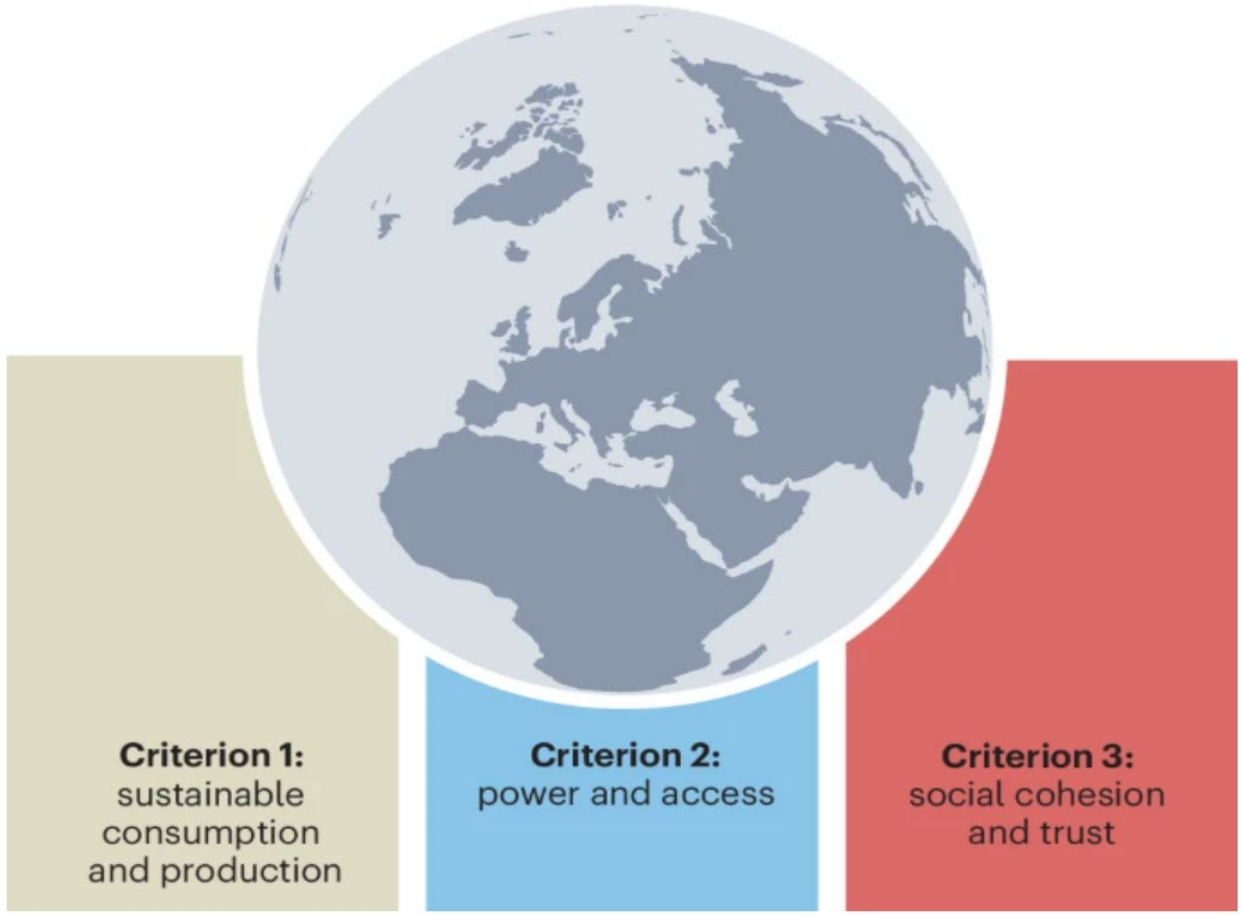
SAFEGUARDING THE PLANET

AI impact demands systemic stewardship, beyond environment impacts:

1. help to accelerate the **transition to sustainable** production and consumption in ways that respect planetary boundaries, or at least do not obstruct these objectives

2. be developed, deployed and used in ways that **ensure equitable access** to AI tools for global sustainability and avoid concentrations of power

3. be developed, deployed and used to **support greater societal cohesion**, build trust and provide access to reliable information for planetary stewardship



Owen Gaffney, Amy Luers, Franklin Carrero-Martinez, Berna Oztekin-Gunaydin, Felix Creutzig, Virginia Dignum, Victor Galaz, Naoko Ishii, Francesca Larosa, Maria Leptin & Ken Takahashi Guevara. The Earth alignment principle for artificial intelligence. *Nat Sustain* (2025). https://doi.org/10.1038/s41893-025-01536-6

OPERATIONALISATION: REGULATION AND MORE

- Regulation
  - AI Act: Human-centered, risk-based approach
- Standards
  - soft governance; non mandatory to follow
  - demonstrate due diligence and limit liability
  - user-friendly integration between products
- Organisation structures and procedures
  - Advisory boards and ethics officers
  - Set and monitor ethical guidelines
- Assessment for trustworthy AI
  - responsible AI is more than ticking boxes
  - Means to assess maturity are needed
- Awareness and Participation
  - Education and training
  - Appeal to civic duty / voluntary implementation

RAI – MORE THAN ETHICS

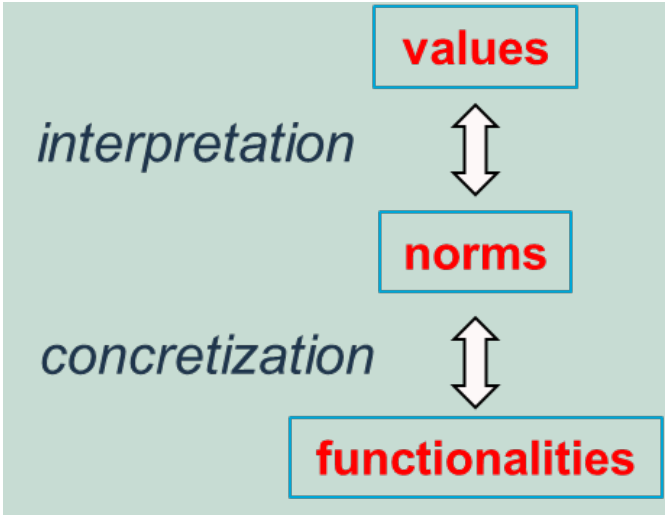
- Not philosophising about ethics
  - Ethics is not about the answer but about recognizing the issue
  - Ethics is a (social) process not a solution

- Not technification of ethics
  - Your implementation does not ‘solve’ ethics

- Fundamentally, is about choices, priorities, tradeoffs
  - Accuracy / Explanation
  - Accuracy / Computational resources
  - Security / privacy
  - Equity / equality
  - Long term benefit / Short term
  - …

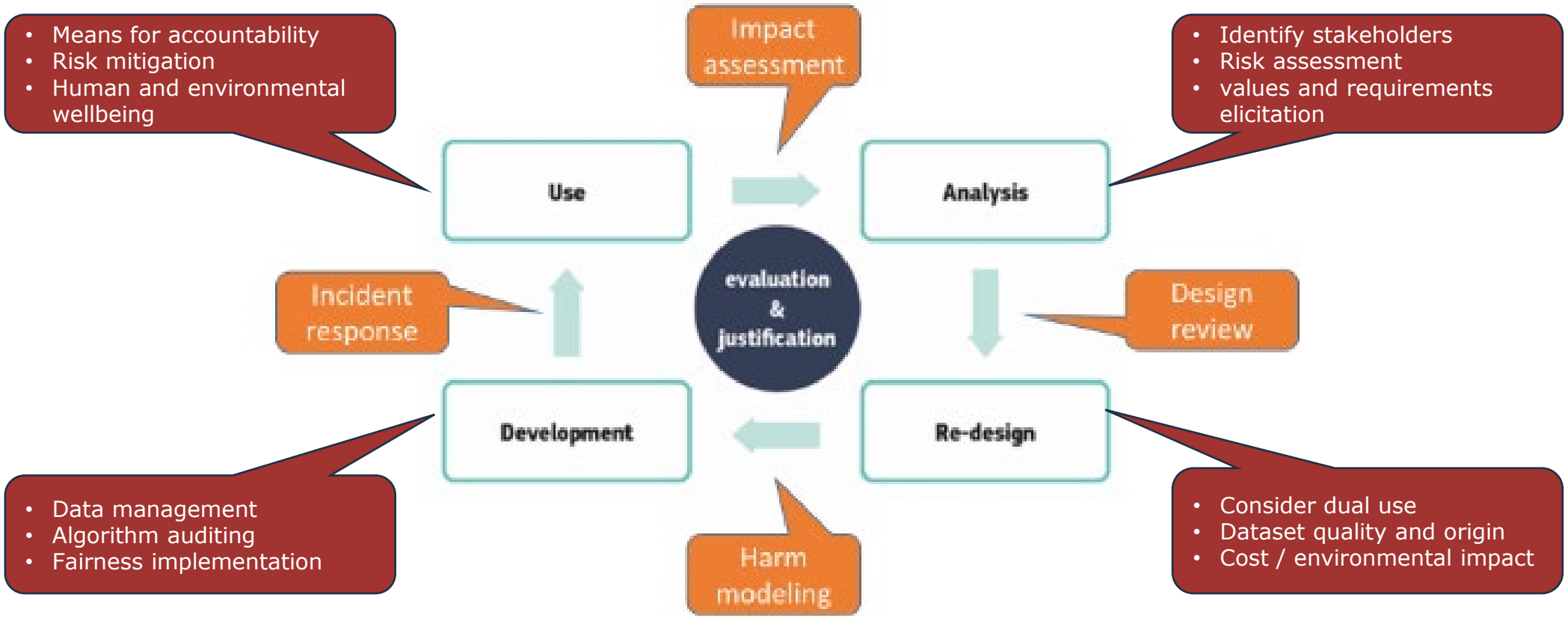
RAI IS ABOUT BEING EXPLICIT

- Design for Values
  - Legal and ethical aspects are not an add-on!
- Governance
  - External monitoring and control
  - Agreements, contracts, norms
- Design
  - Question your options and choices
  - Motivate your choices
  - Document your choices and options



https://medium.com/@virginiadignum/on-bias-black-boxes-and-the-quest-for-transparency-in-artificial-intelligence-bcde64f59f5b

RESPONSIBILITY IN THE AI LIFECYCLE



AI Policy Lab
at Umeå University

A NEW AI PARADIGM ?

- AI is not a universal solution
  - recognize its limitations in addressing complex challenges
  - Question zero: should AI be used here?
- Address inherent risks of bias and discrimination
  - improving data and algorithmic transparency
  - accommodate diverse contexts and values
- Beyond disciplines
  - integrate social and technical expertise in AI design
  - addressing systemic societal challenges
  - involving varied societal actors, not just technologists
- robust technical standards
  - Verifiable, sustainable, participatory

TRUSTWORTHY GOVERNANCE

- Development of AI
  - What? Requirements, needs, possibilities, capabilities
- Use of AI
  - How? license to operate
- Context in which AI is developed and used
  - Where/When? rules of the game

- Global efforts
  - Sustainability
  - Inclusion and participation
  - Diversity
  - Distribution of benefits and costs
  - Agenda 2030 - SDGs

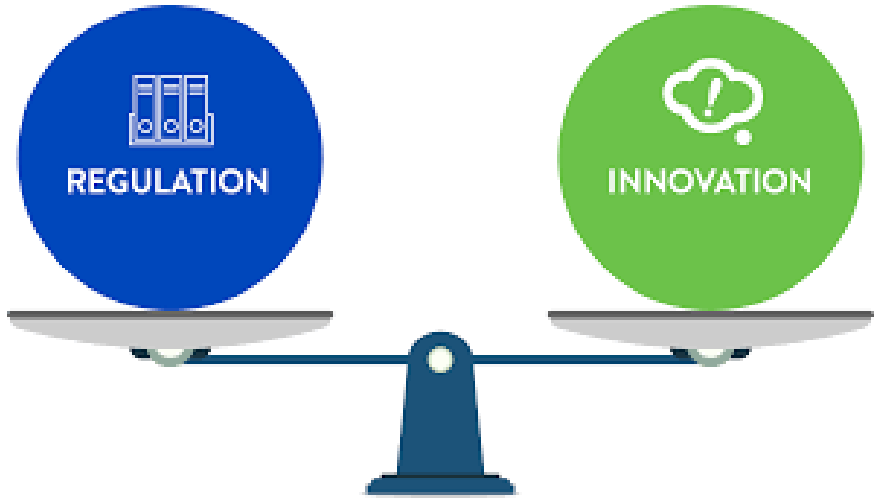
political will and innovation needed!

AI Policy Lab
at Umeå University

RESPONSIBLE AI IS INNOVATION

- Innovation needed
  - Technical requirements
  - Use requirements
  - Context of operation

- Multidisciplinary innovation!
  - Technological innovation
  - Organisational innovation
  - Regulation innovation
  - Governance innovation
  - Social innovation



RESPONSIBLE AI IS NOT A CHOICE!

Not *innovation vs governance* but

governance as stepping-stone for innovation

- Innovation is not a race!
  - There is no finish line
  - Exploration rather than one fixed direction
- Adopting responsible AI
  - Builds trust
  - Drives transformation
  - Support business differentiation

AI Policy Lab
at Umeå University

THANK YOU!

AI Policy Lab
at Umeå University