AI: RESPONSIBILITY IN A CHANGING WORLD

Prof. Dr. Virginia Dignum

Chair Responsible AI – Director AI Policy Lab

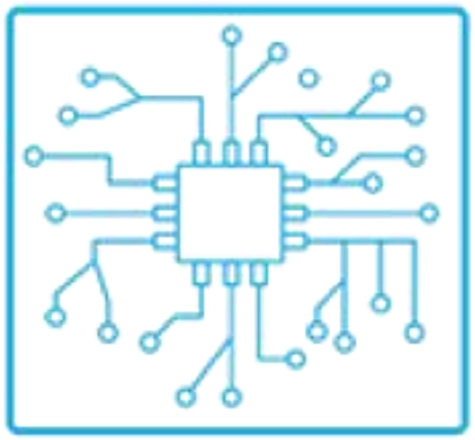
Email: virginia@cs.umu.se

UMEÅ UNIVERSITY

AI Policy Lab

at Umeå University

AI – MORE IS BETTER?



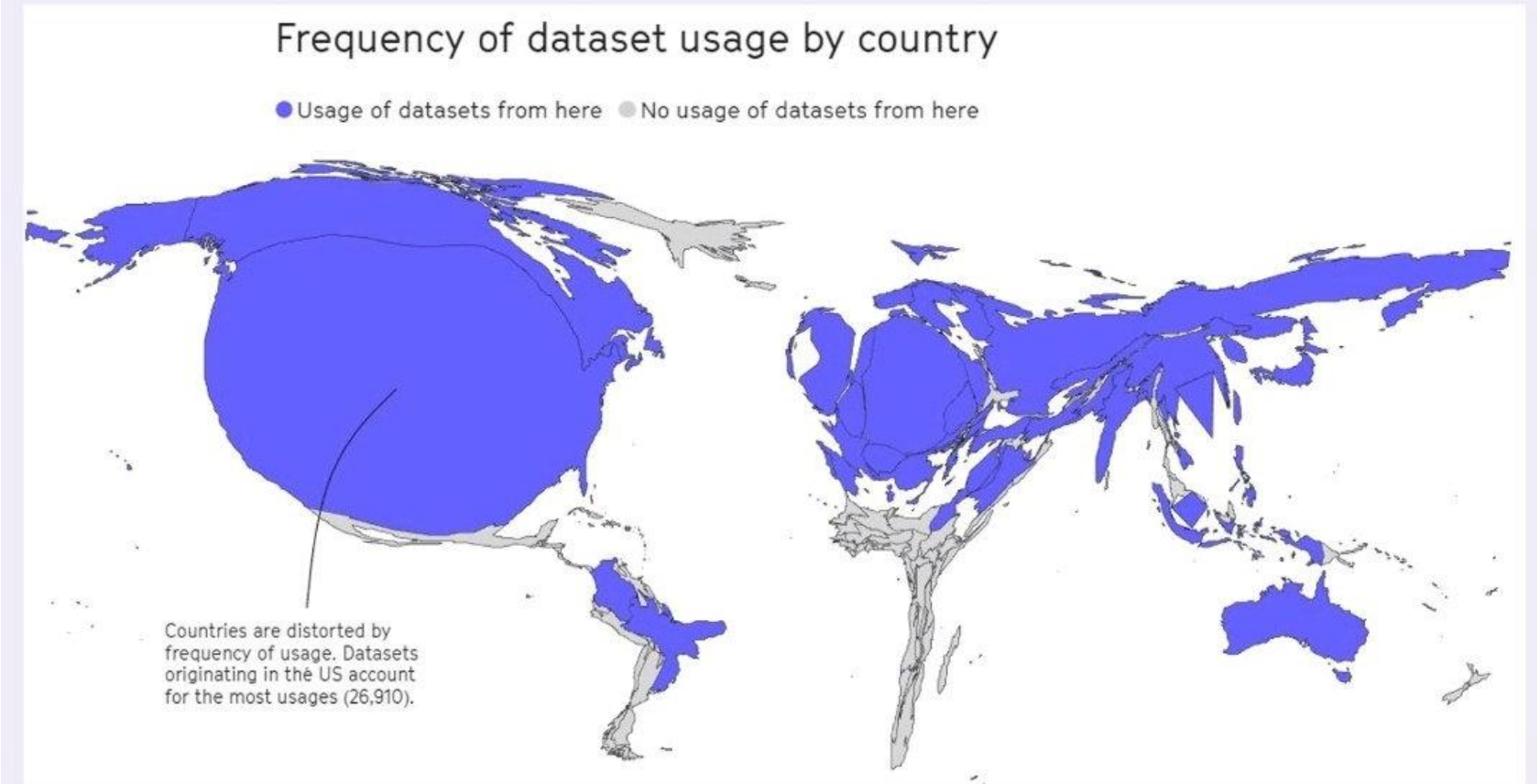
Computing Power

Algorithm Power

Data Availability

HOW AI 'SEES' THE WORLD

- 50% of datasets are connected to 12 institutions
- WEIRD demographics (Western, educated, industrialised, rich, democratic)



Mozilla internet health report 2022

HOW AI MODELS THE WORLD

- AI is an instrumental rational system
  - AI agents have preferences, or priorities, on outcomes of actions;
  - AI agents optimize actions based on those preferences.

Stuart Russell and Peter Norvig.*Artificial intelligence: a modern approach.* PrenticeHall, 2010.

- AI principles are Global North stereotypes
  - Optimisation / Efficiency / Rationality / Agency / Autonomy

Abeba Birhane. Algorithmic colonization of Africa. *Scripted*, 18(2), 1-24, 2021.
Virginia Eubanks. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press, 2018.

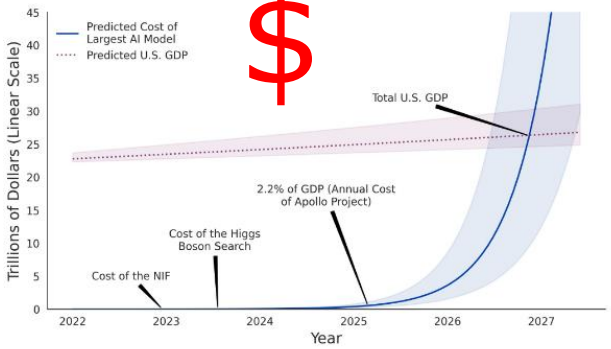
- But… People are relational rational; act in context
  - pursue seemingly incompatible goals and hold inconsistent beliefs, different motives (altruism, fairness, justice, or to prevent future regret)
  - actions are influenced by the context, including others and different situations.
  - We don't maximize forever: good is good enough -> satisficing!

Virginia Dignum. Social Agents: Bridging Simulation and Engineering. *Communications of the ACM*, November 2017, Vol. 60 No. 11, Pages 32-34

AI Policy Lab
at Umeå University

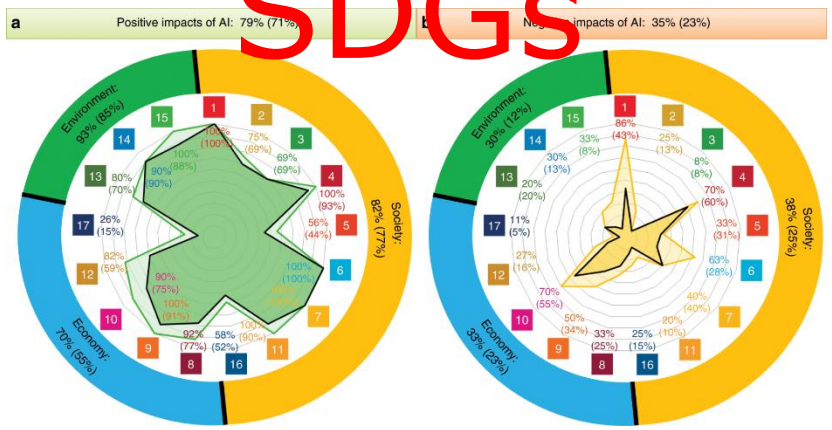
HOW AI USES THE WORLD

- Computational cost of AI
- Human and social costs

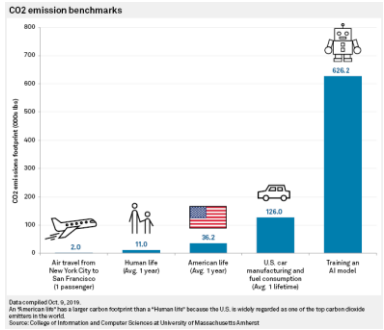


$

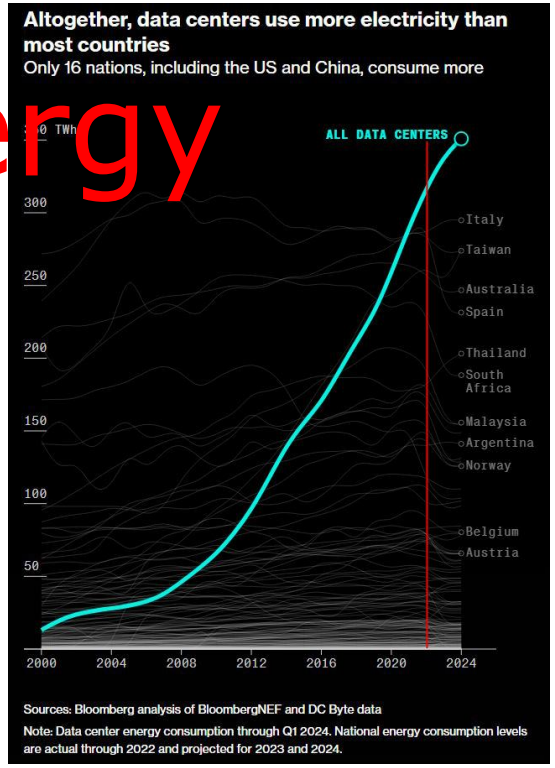
Source: CSET. Note: The blue line represents growing costs assuming compute per dollar doubles every four years, with error shading representing no change in compute costs or a doubling time as fast as every two years. The red line represents expected GDP at a growth of 3 percent per year from 2019 levels with error shading representing growth between 2 and 5 percent.



SDGs



Energy



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

Water

AI Policy Lab
at Umeå University

DEMOCRACY UNDER ATTACK?

- Erosion of Democratic Oversight
  - AI enables private companies to control surveillance and infrastructure
  - Bypassing regulatory oversight and reducing government authority.
- Corporate Power
  - AI-driven automation and data control give corporations unprecedented economic and political dominance
  - combined with inadequate regulatory oversight, this allows them to operate with minimal accountability, undermining democratic processes and societal norms
- Impact on Human Rights
  - AI-powered surveillance and facial recognition enable authoritarian control and mass privacy violations
  - endangering individual freedoms and human rights globally

Marietje Schaake, The Tech Coup, 2024.

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

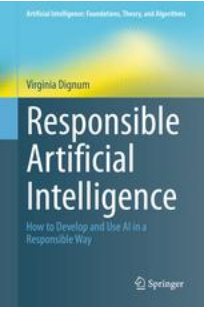
Amazon's facial recognition matched 28 members of Congress to criminal mugshots

Flawed Algori

AI IMPACT MEANS HUMAN RESPONSIBILITY

AI
at Umeå University

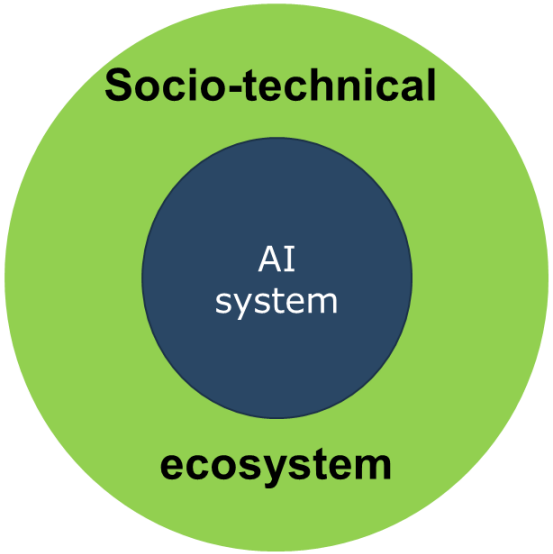
CAN WE BE RESPONSIBLE?



RESPONSIBLE 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

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

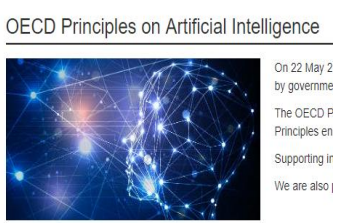
  - products, and results
  - Need for better understanding and integration of existing frameworks alongside introducing more regulation
- Avoidance of an "arms race" narrative in AI regulation

PRINCIPLES AND GUIDELINES

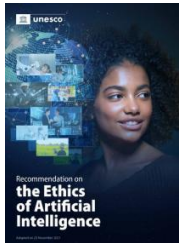
EU HLEG	OECD	IEEE EAD
• Human agency and oversight • **Technical robustness and safety** • Privacy and data governance • **Transparency** • **Diversity**, non-discrimination and fairness • **Societal and environmental well-being** • **Accountability**	• benefit people and the planet • respects the rule 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?

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

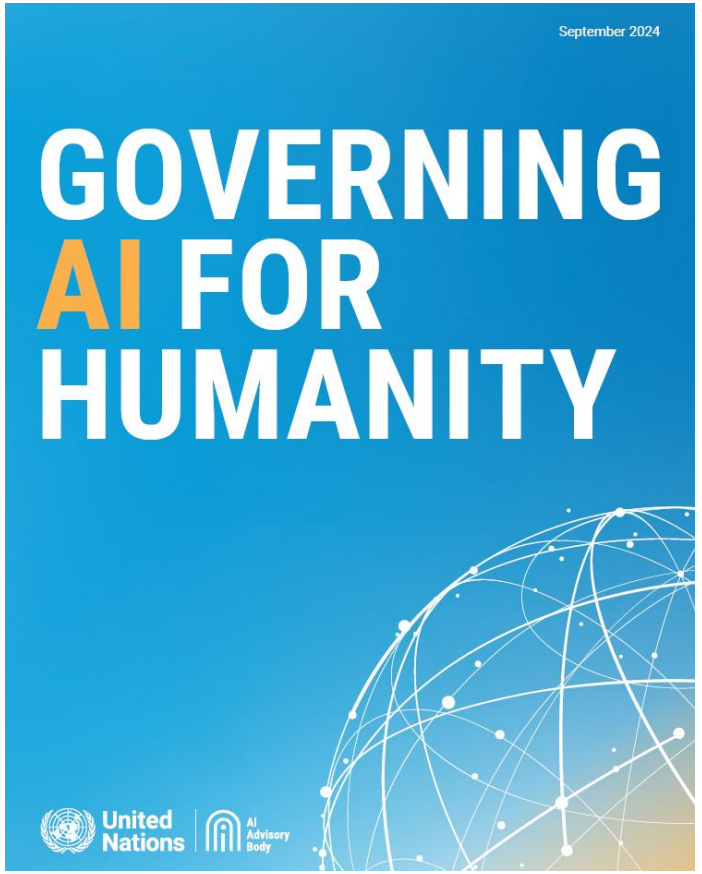
https://ethicsinaction.ieee.org



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



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



Available at
www.un.org/ai-advisory-body

AI Policy Lab
at Umeå University

REAL RISKS VS. HYPOTHETICAL RISKS?

- **Two main views on AI risks:**
- **Long term** (AI Safety)
  - Focuses on future, big risks like superintelligent AI.
  - These dangers haven’t happened yet but could be huge and irreversible.
  - Backed by longterm thinkers, tech leaders, and national security experts.
- **Now** (AI Ethics)
  - Focuses on current, real problems caused by AI.
  - Issues of bias, discrimination, privacy loss, misinformation, job loss, and too much power in a few hands.
  - Backed by ethics researchers, civil rights groups, labor experts, and privacy advocates.
  - Argue we shouldn't ignore today's urgent problems while worrying about far-off risks.

SAFEGUARDING DEMOCRACY

- **Stronger Rules and Oversight**
  - Make clear laws for things like AI, facial recognition, and cryptocurrencies.
  - Ensure companies are transparent and accountable.
  - Support democratic institutions to guide tech development.
- **Public Participation**
  - Involve people from all backgrounds in shaping AI and tech.
  - Ensure technology aligns with democratic values and the public good.
- **Digital and Ethics Education**
  - Teach everyone how AI and digital tools work.
  - Promote ethical and responsible use of technology.
  - Empower people to defend their rights in the digital world.

RESPONSIBLE AI IS NOT A CHOICE!

Not *innovation vs governance* but

governance as stepping-stone for innovation

- Innovation is moving technology forward, is not the use of existing tech ‘as is’

- Adopting responsible AI
  - Build trust
  - Drive for transformation
  - Business differation

AI Policy Lab
at Umeå University

THANK YOU!

AI Policy Lab
at Umeå University