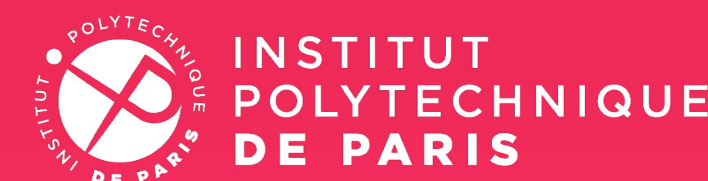


AI as a sociotechnical system

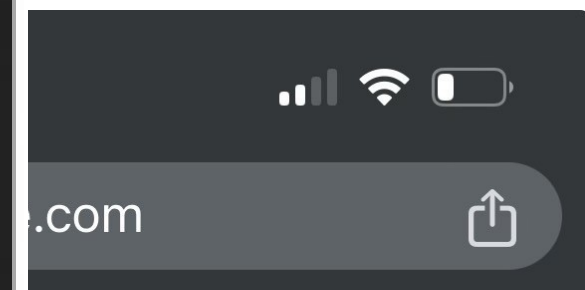
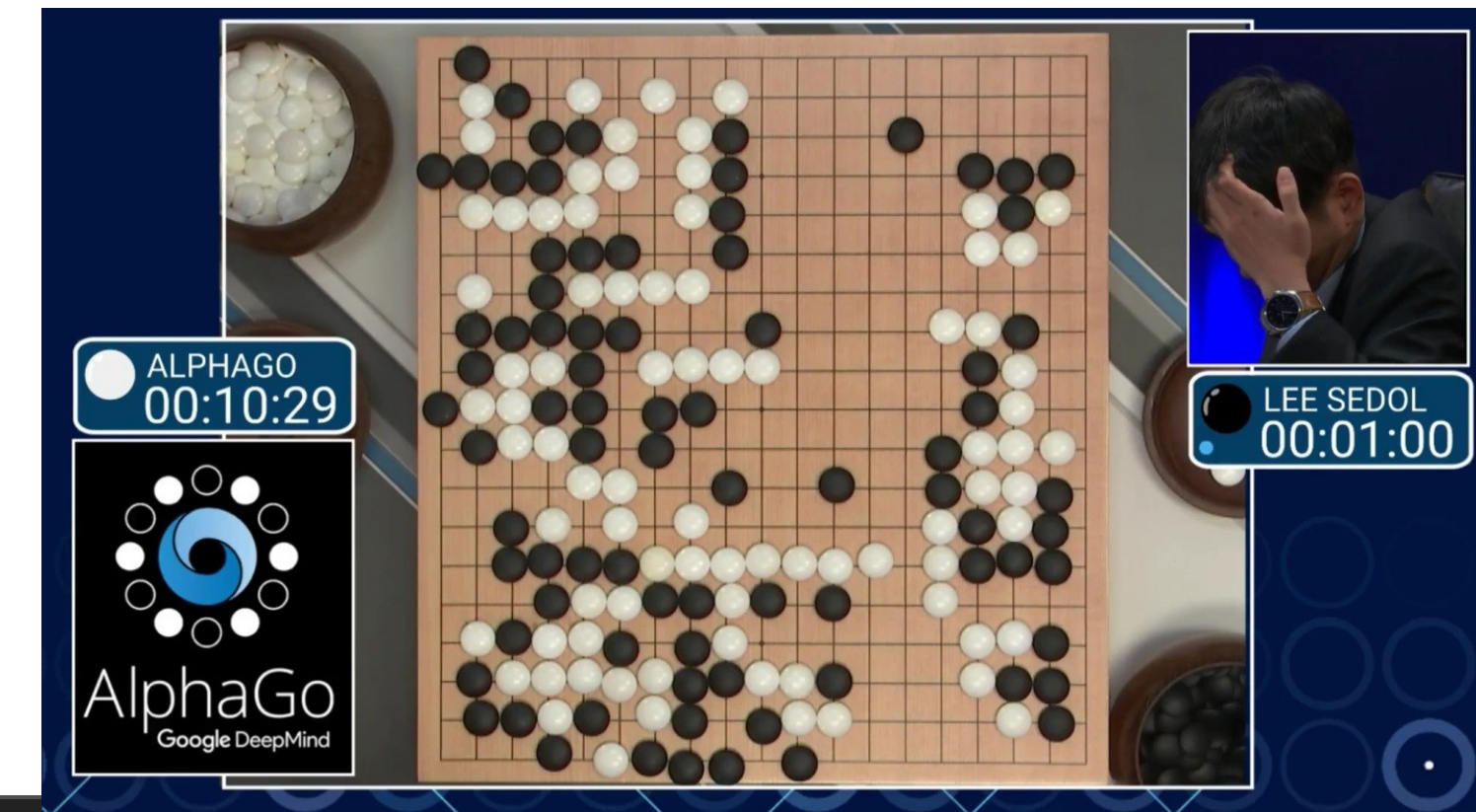
Tiphaine Viard – Associate professor
Numérique, Organisation et Société,
Operational AI Ethics
Ellis PhD summer school



tiphaine.viard@telecom-paris.fr
github.com/TiphaineV
<https://tiphaine.vi.github.io>
Mastodon: tviard@lipn.info

What is AI?

- One of the 5 **promises of technology** (WEF)
- A booming **research and development** field
- A **prism for bias, inequalities and power struggles**
- **Successes and failures** are both **flashy and silent**
- AlphaGo, spam filters, COMPAS, SyRI...
- One challenge: separating **hype** from **reality**



What is AI (in this talk)?

- “Which metric should we use, micro- or macro-F1 score?”
- “When is deep learning more suited than non-deep models?”
- “Which rules should govern ‘AI’ use in high-risk systems?”
- “Will AI exterminate humanity?”
- “We use ‘AI’ to determine recidivism risk”
- “How do scientists conciliate AI’s energy consumption with a sustainable world?”
- “Who talks about ‘AI’ and what do they say?”
- “We want fair AI, but fair to whom?”

**AI = the general concept that encompasses
systems
that learn inferences from data**

What is AI (in this talk)?

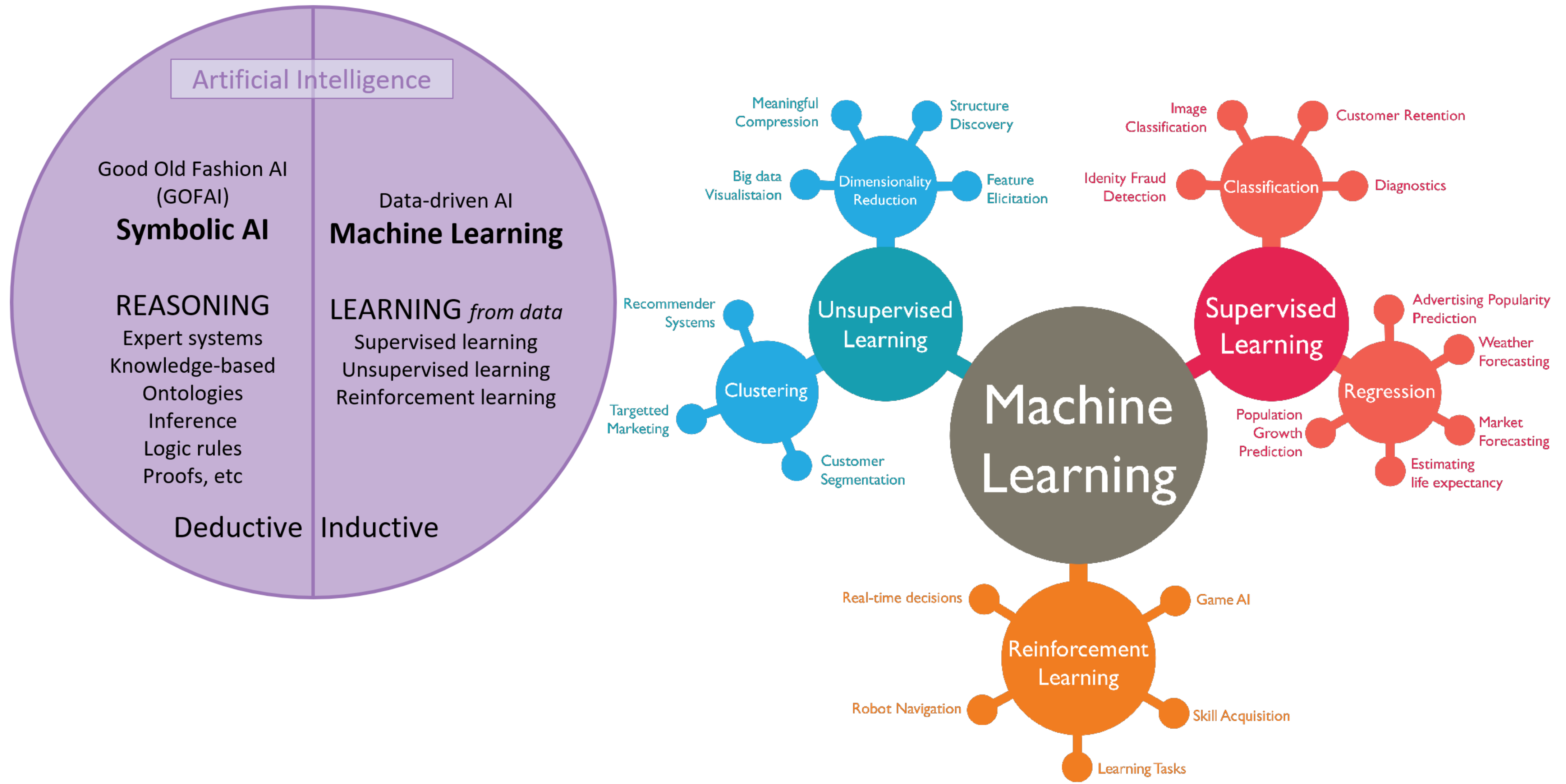
- “Which metric should we use, micro- or macro-F1 score?” (technical)
- “When is deep learning more suited than non-deep models?” (technical)
- “Which rules should govern ‘AI’ use in high-risk systems?” (regulation)
- “Will AI exterminate humanity?” (speculative)
- “We use ‘AI’ to determine recidivism risk” (regulation/usage)
- “How do scientists conciliate AI’s energy consumption with a sustainable world?” (sociology)
- “Who talks about ‘AI’ and what do they say?” (sociology)
- “We want fair AI, but fair to whom?” (technical, law, sociology...)

**AI = the general concept that encompasses
systems
that learn inferences from data**

This lecture

- Exhibit the **tension** between “artificial intelligence” and “machine learning”
- Get a broad overview of the **sociotechnical system** that is AI and its **history**
- Get into the regulation framework for AI, through **law** and **standards**
- Understand how the **landscape** of AI is shaped: which topics are discussed? Which tenets are central? How do actors **interact** and **negotiate meaning**?
- I will speak at the crossroads of **computer science** and **science and technology studies**

Some AI classifications



Sociohistory of AI

Cardon, D., Cointet, J. P., & Mazières, A. (2018). La revanche des neurones. *Réseaux*, 211(5), 173-220.

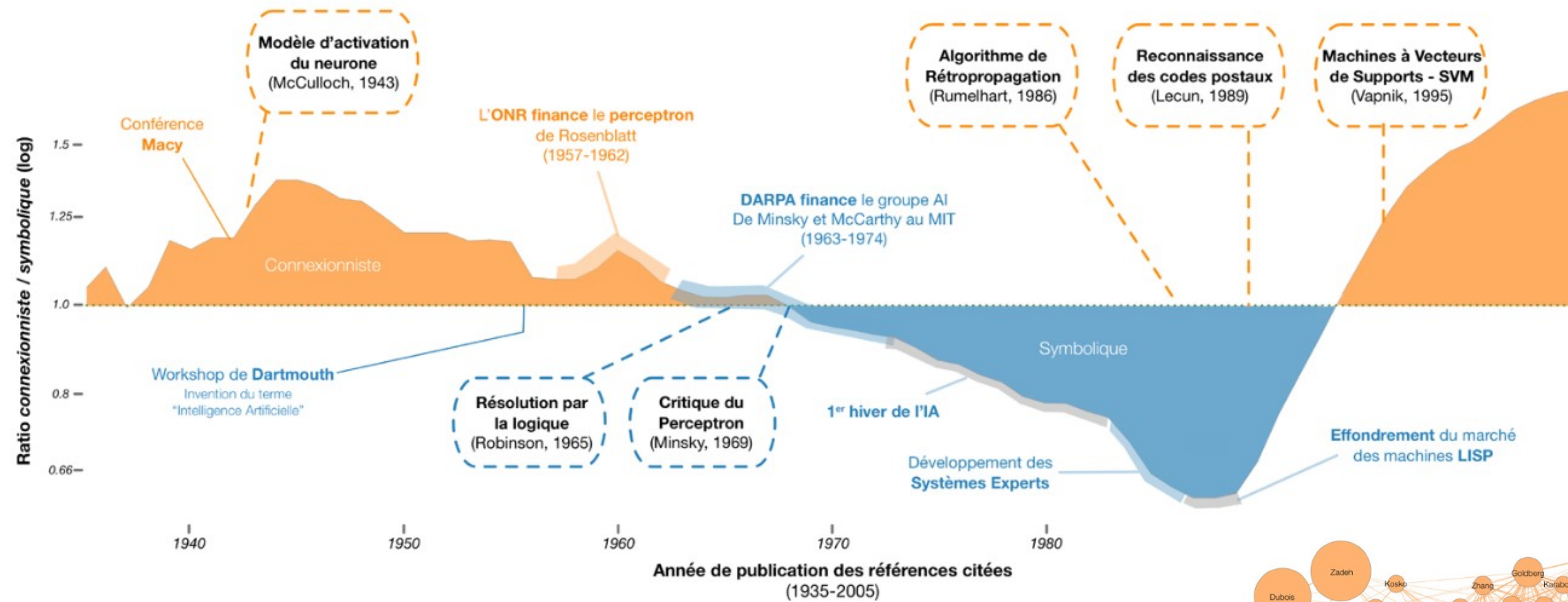
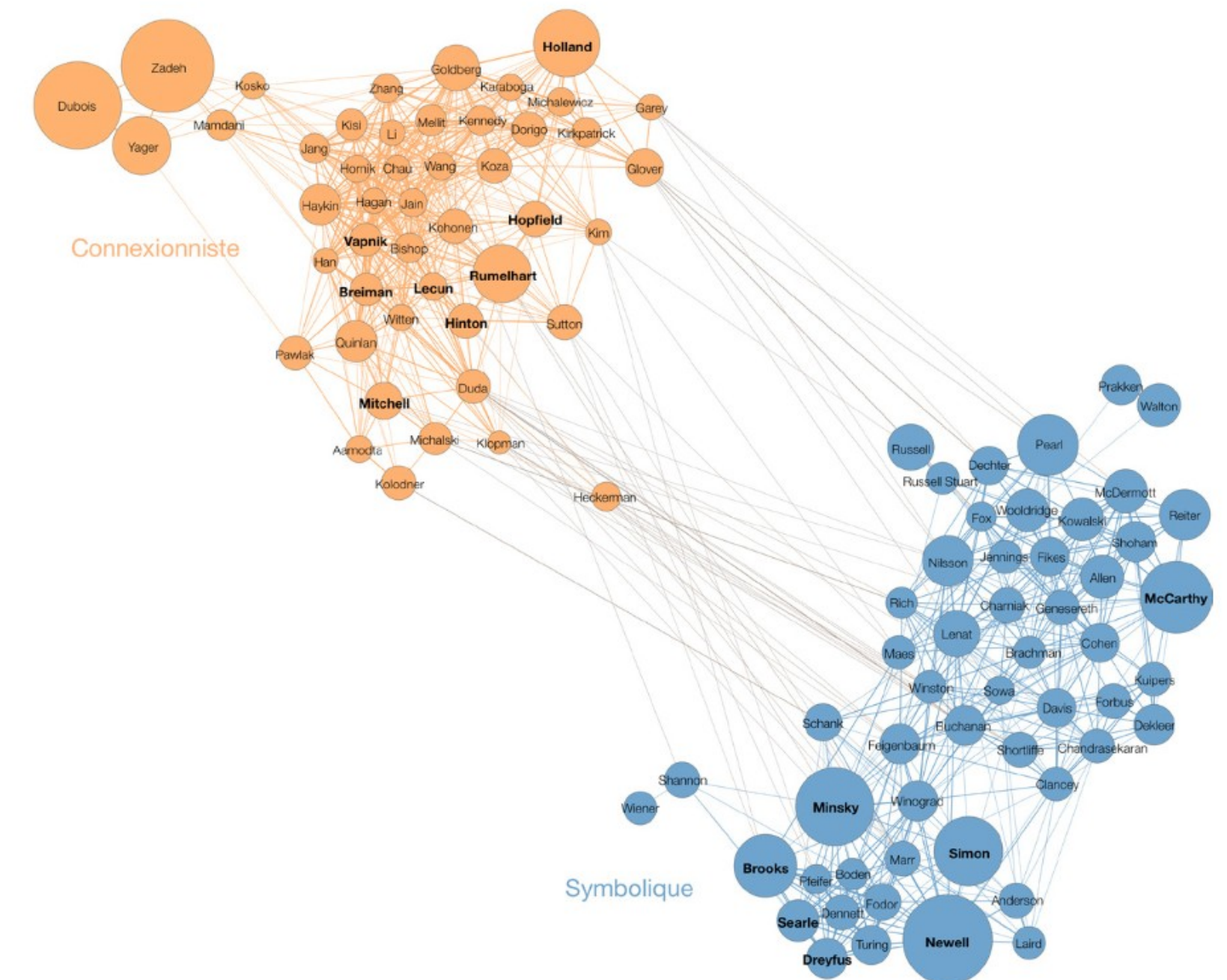


Figure 1. Machine hypothético-déductive (1) et machine inductive (2)



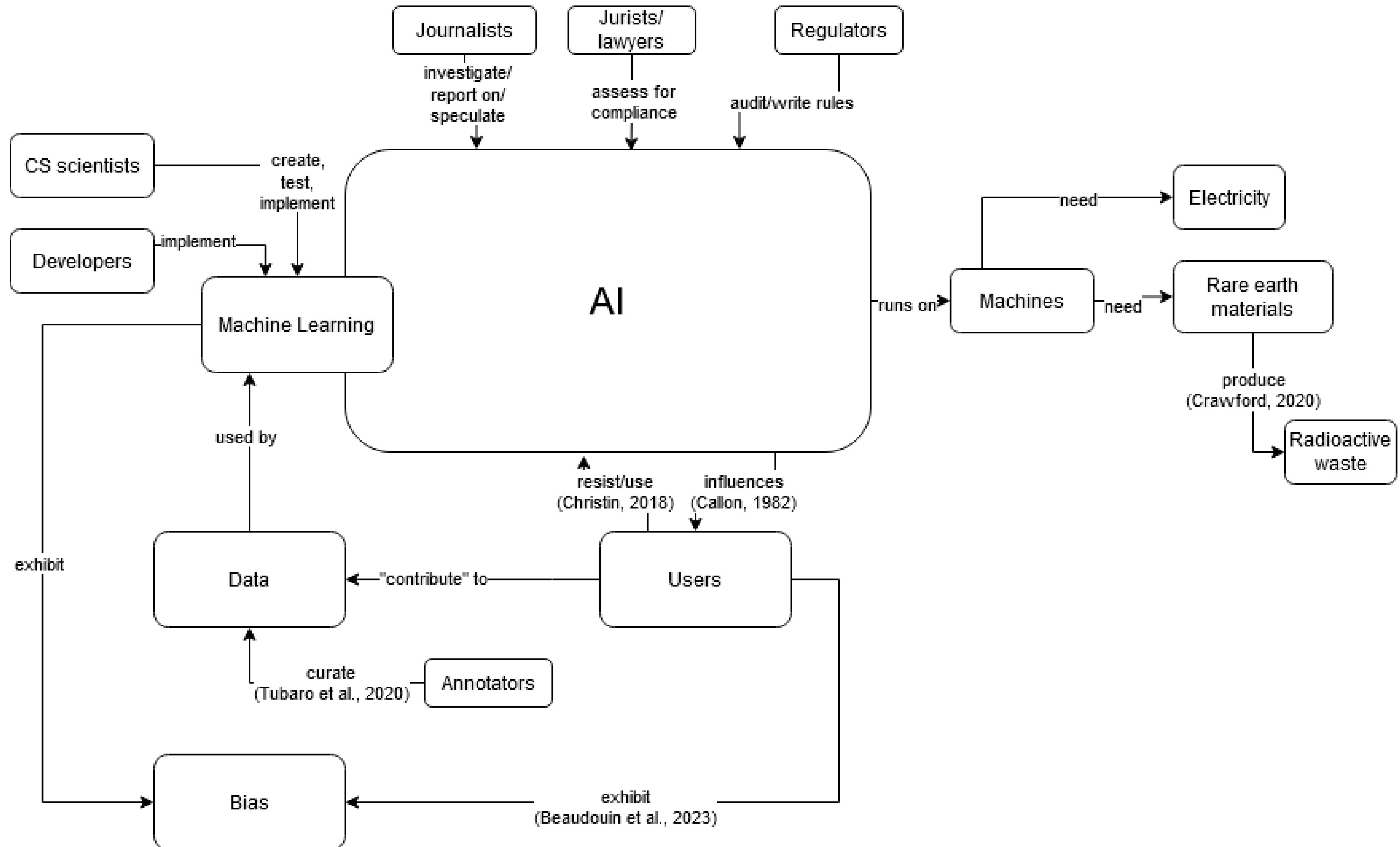
Sociotechnical systems and their analysis

- Studying technical objects in society: the focus of **Science and Technology Studies (STS)**
- **Framing** is important: where do technical objects start and end?
- The **reception** of technical objects cannot be ignored: uses emerge, evolve and disappear



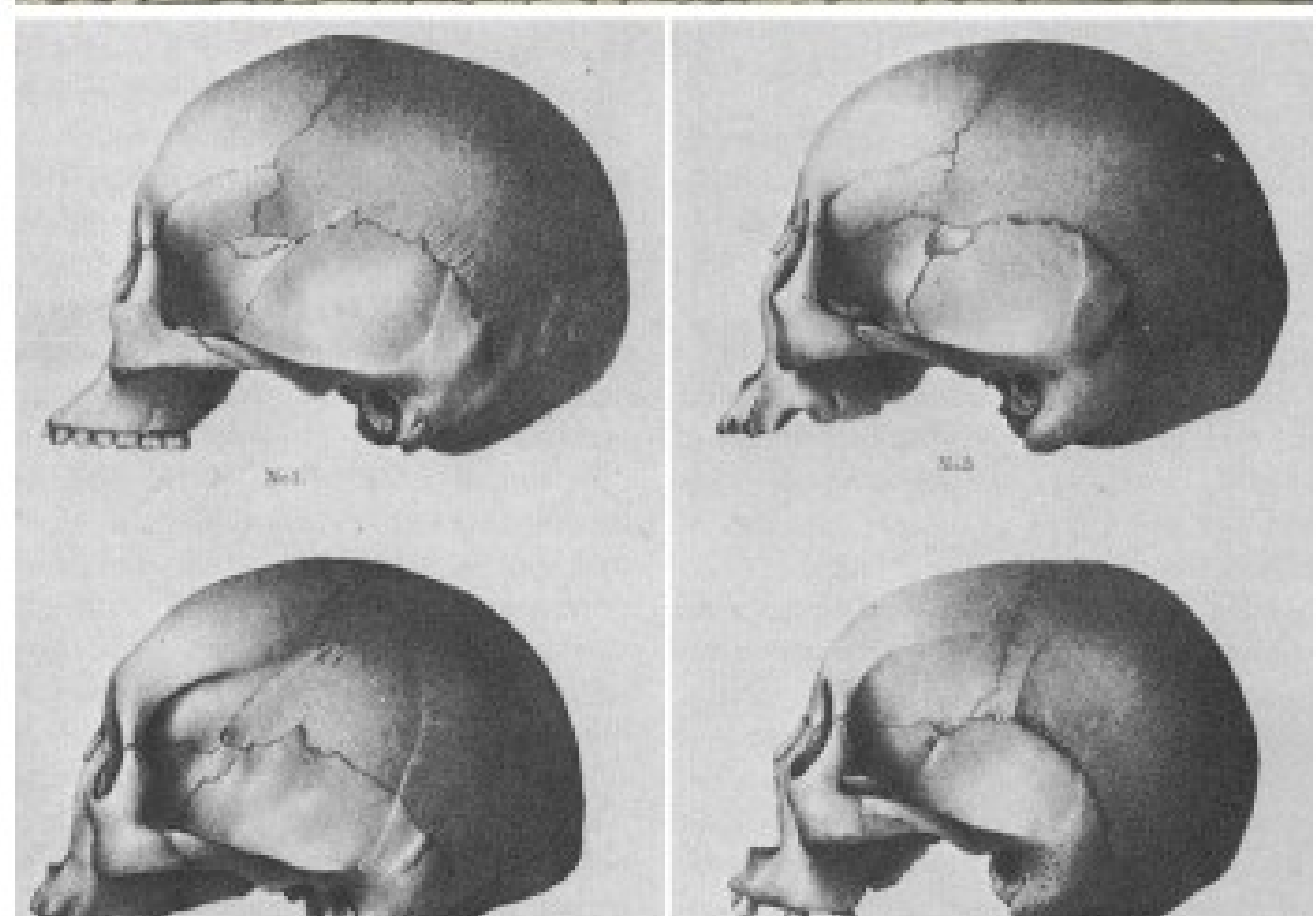
- Citizens' attention, policy makers' attention are limited resources
- Decisions are taken in **sub-political spaces** (Beck, 1986)

The (incomplete) AI sociotechnical system



The materiality of AI: human labour

- Clever Hans or the illusion of intelligence
- Intelligence was narrowly defined: Interspecies communication, public performance, a lot of patience!
- Understanding the shaping of labour:
 - Amazon's "matrix" algorithm minimises breakage but robs workers of habituation,
 - Focus on the time clock,
- Digital labour reinforces North-South dynamics, with cascades of delocalisation.
- Multiple hidden human roles: the trainer, the verifier, the imitator (Tubaro et al., 2020)



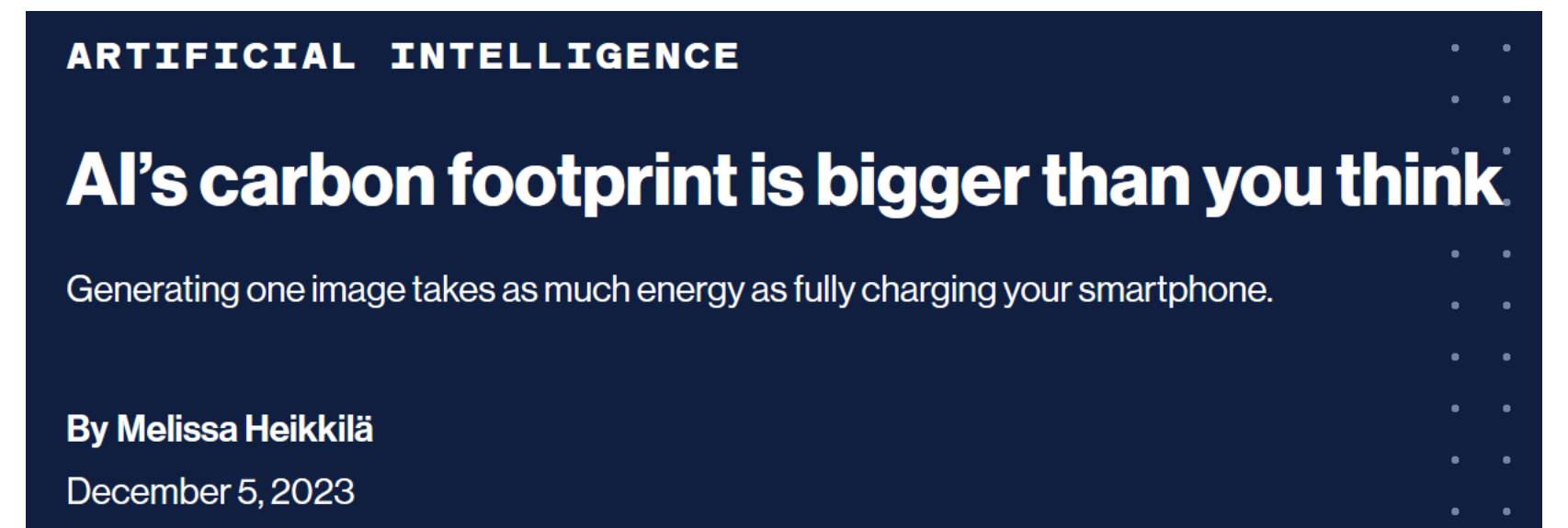
Crawford, K. (2021). *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.

Tubaro, P., Casilli, A. A., & Coville, M. (2020). The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence. *Big Data & Society*.

Le Ludec, C., Cornet, M., & Casilli, A. A. (2023). The problem with annotation. Human labour and outsourcing between France and Madagascar. *Big Data & Society*

The materiality of AI: GPUs and servers

- Rematerialising the “cloud”
- Dependent on mine sites (esp. lithium),
- 99.8% of earth is treated as waste,
- Work conditions are abhorrent,
- 1 ton of rare earth elements = 75,000L of acidic water, 1 ton radioactive residue
- Deep learning is **carbon-intensive** and **labour intensive**



Common carbon footprint benchmarks

in lbs of CO2 equivalent

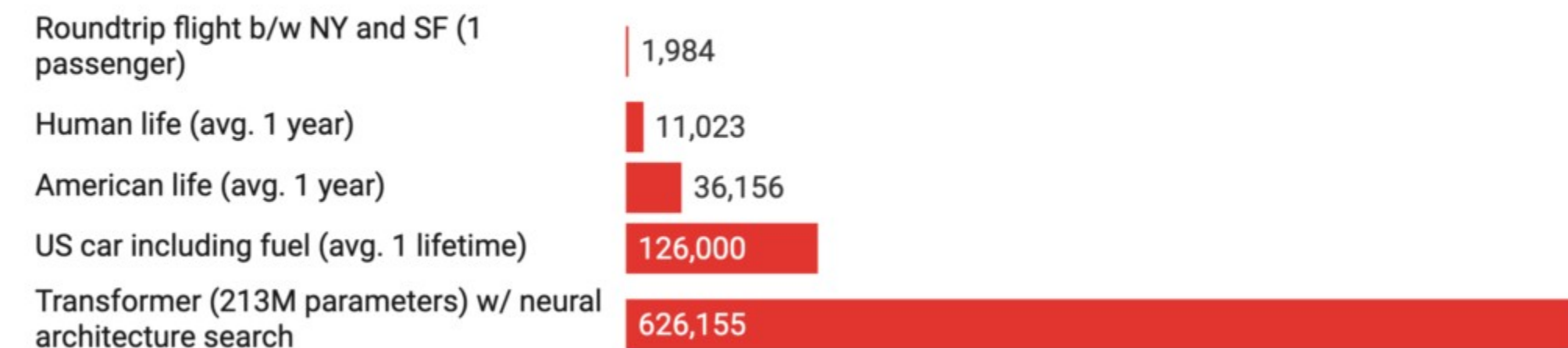


Chart: MIT Technology Review • Source: Strubell et al. • Created with Datawrapper

The regulation of AI: the EU's AI Act

- First works by the High Level Expert Group in 2021
- Final vote in May 2024
- Entry into force starting May 2025
- Main definitions: “AI systems” and “high-risk AI systems”
- Reliant on standards, national supervisory authorities
- Much left to be decided!



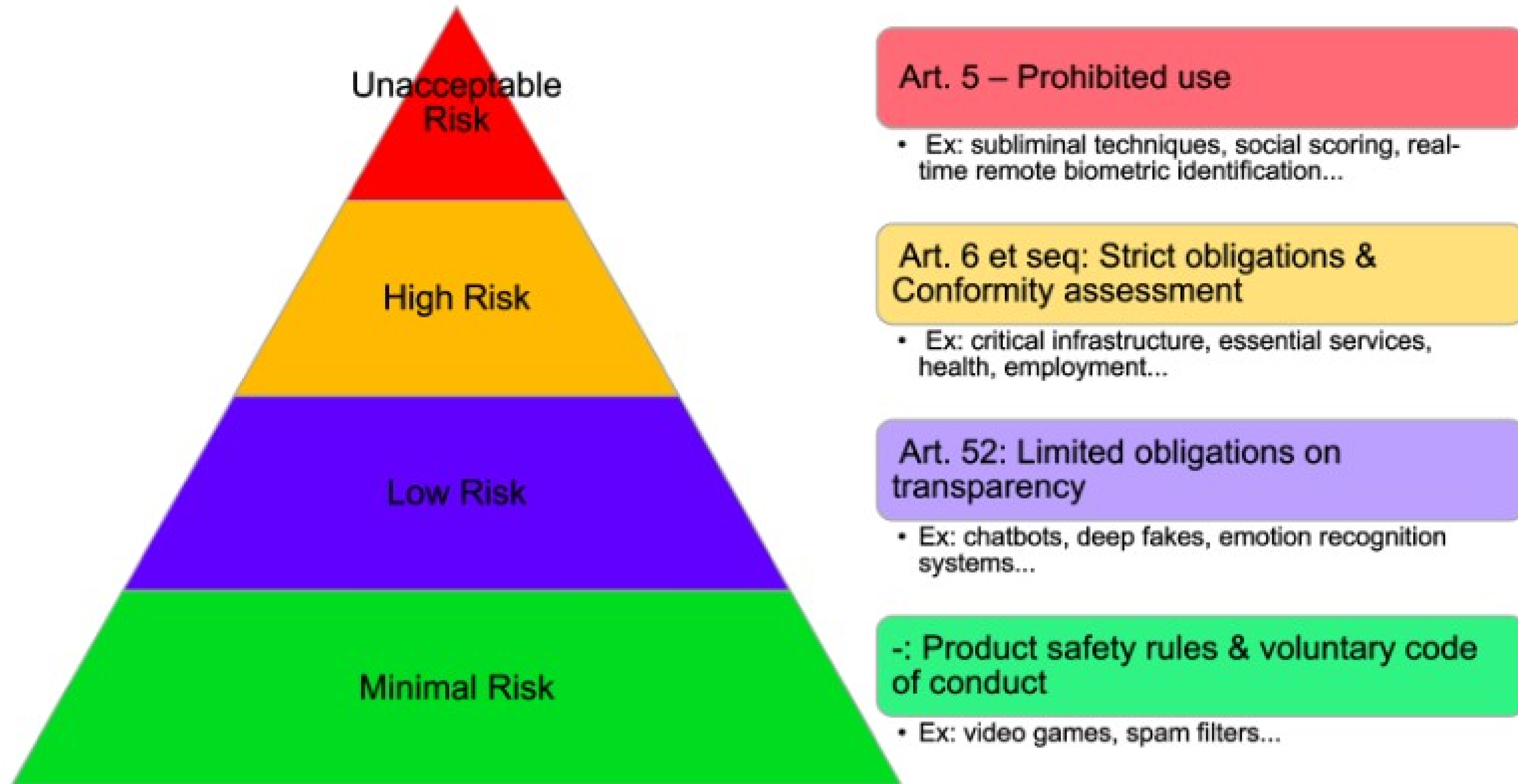
AI systems

- **‘AI system’** means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments;

Classification rules for high-risk AI systems

1. Irrespective of whether an AI system is placed on the market or put into service independently of the products referred to in points (a) and (b), that AI system shall be considered to be high-risk where both of the following conditions are fulfilled:
 - (a) the AI system is intended to be used as a safety component of a product, or the AI system is itself a product, covered by the Union harmonisation legislation listed in Annex I;
 - (b) the product whose safety component pursuant to point (a) is the AI system, or the AI system itself as a product, is required to undergo a third-party conformity assessment, with a view to the placing on the market or the putting into service of that product pursuant to the Union harmonisation legislation listed in Annex I.
2. In addition to the high-risk AI systems referred to in paragraph 1, AI systems referred to in Annex III shall be considered to be high-risk.
3. By derogation from paragraph 2, an AI system referred to in Annex III shall not be considered to be high-risk where it does not pose a significant risk of harm to the health, safety or fundamental rights of natural persons, including by not materially influencing the outcome of decision making.

Risks in the AI Act



‘risk’ means the combination of the probability of an occurrence of harm and the severity of that harm; (Art. 3, 2)

High-risk AI systems

- List in Annex 3 adds to the previous definition
 - Biometrics (incl. emotion recognition), critical infrastructure, education, employment, essential services, law enforcement, migration/asylum, justice and democratic processes
 - This list can be **regularly amended** by the EU Commission
- (Initial) Compliance + risk management system + technical documentation + record-keeping + transparency + human oversight
- Many of these concepts are **not yet specified** : this will come with **standards**

Further resources on EU's AI Act

- <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689>
- <https://artificialintelligenceact.eu/>
- Thomas Le Goff's AI Act game:
https://www.canva.com/design/DAGF2FfogqE/QpjWW1ghhCr_GMZT_gLJ3A/view



A dense regulatory ecosystem

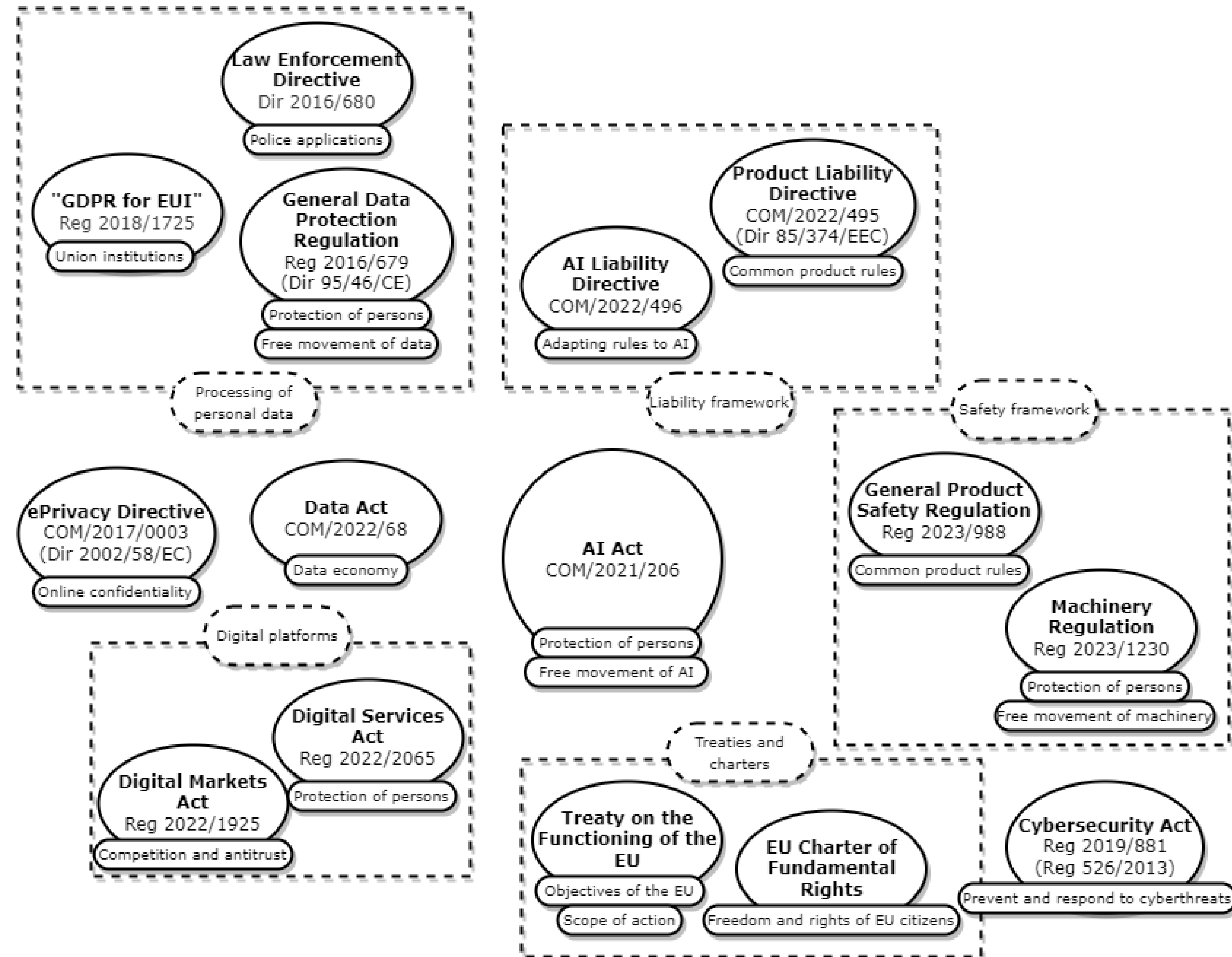
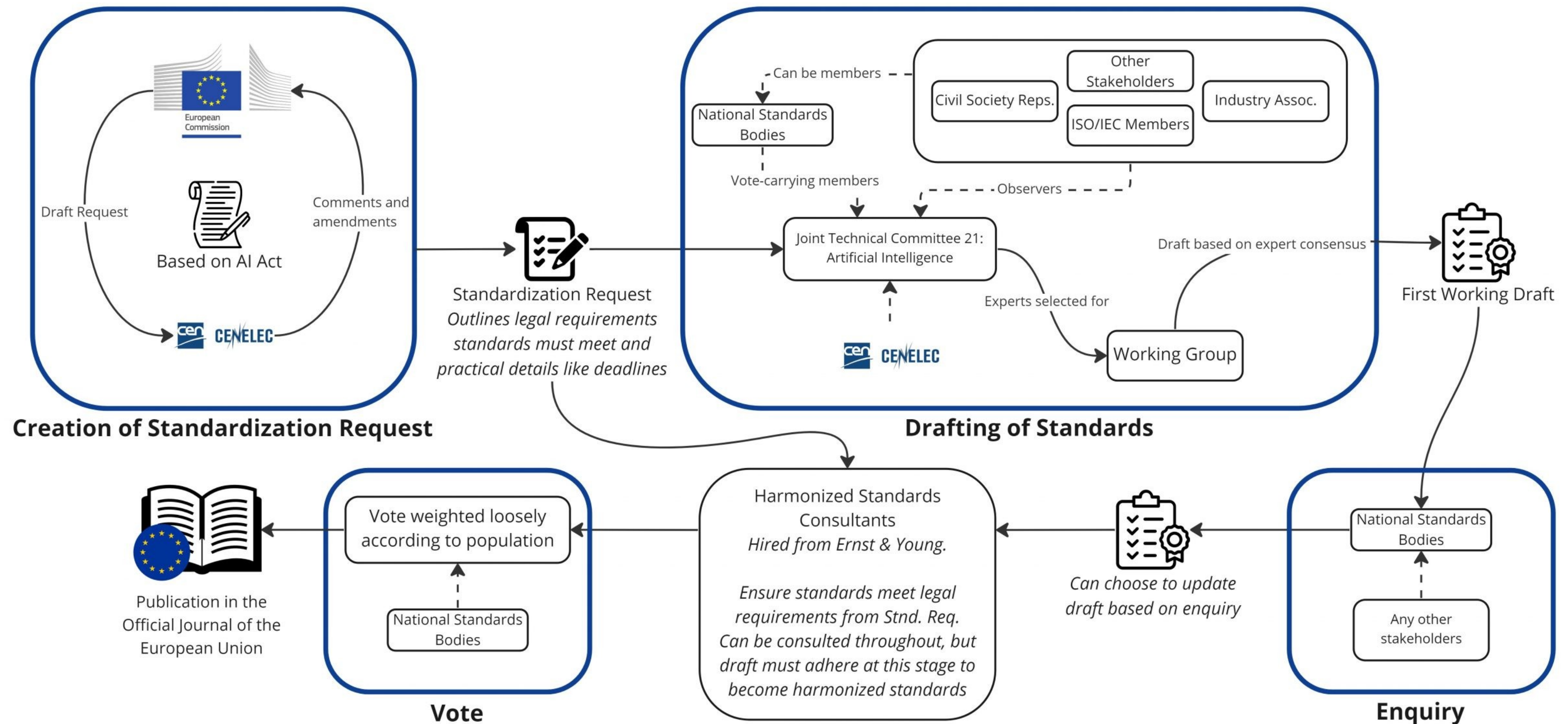


Figure by Mélanie Gornet, reproduced with authorisation

The regulation of AI: standards



The Development Process for Harmonized Standards

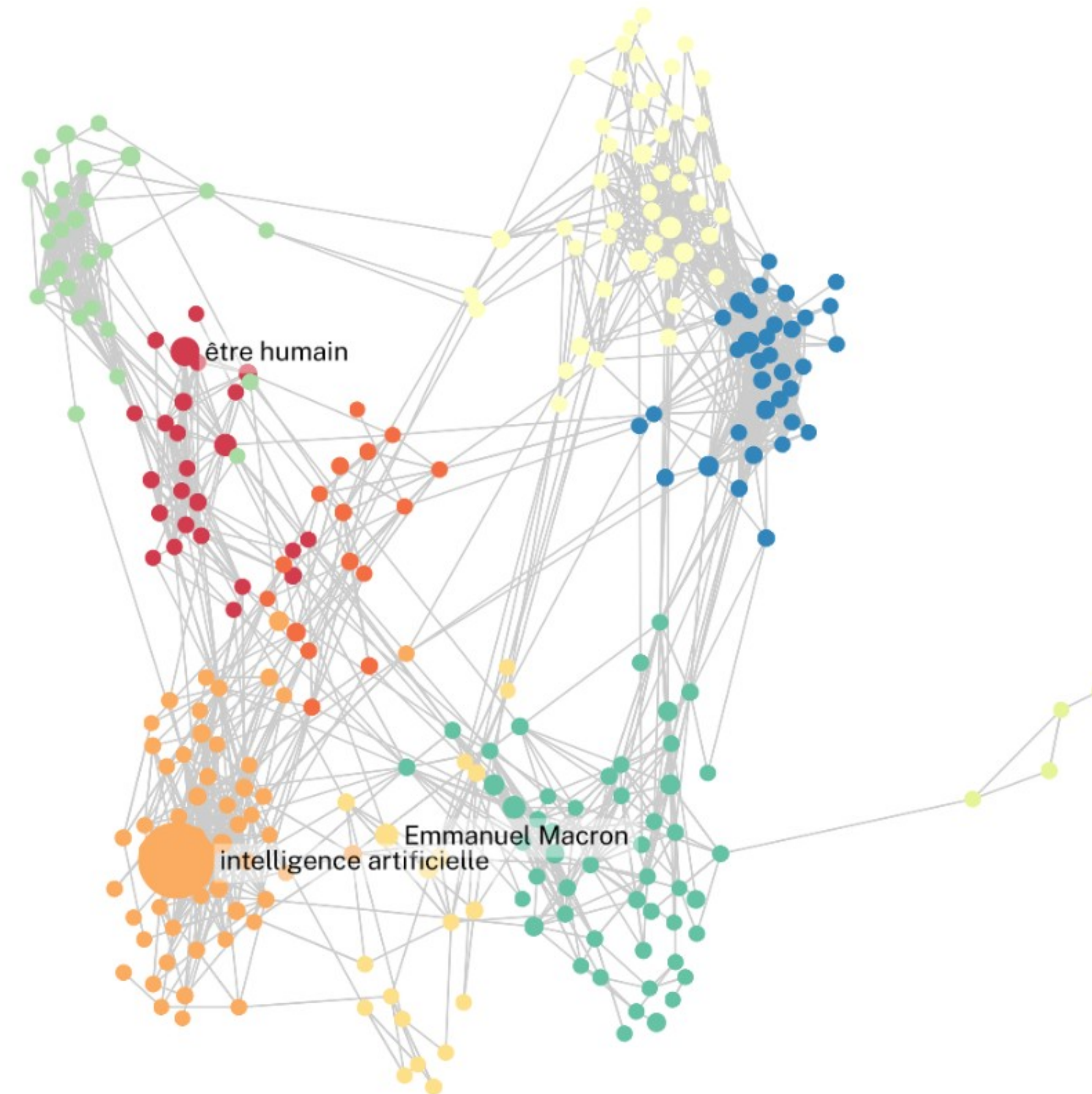
Standards and AI

- Compliance = testing against **harmonized standards** (hENs)
- A common tool for product safety, associated with the **CE mark**
- **AI Act novelty:** certification of fundamental rights
- Difficult process: fundamental rights are heavily **context-dependent** and require **judicial determination**
- **Democratic processes** in standard organisations are put into question (CJEU, 5 march 2024)



The media landscape of AI

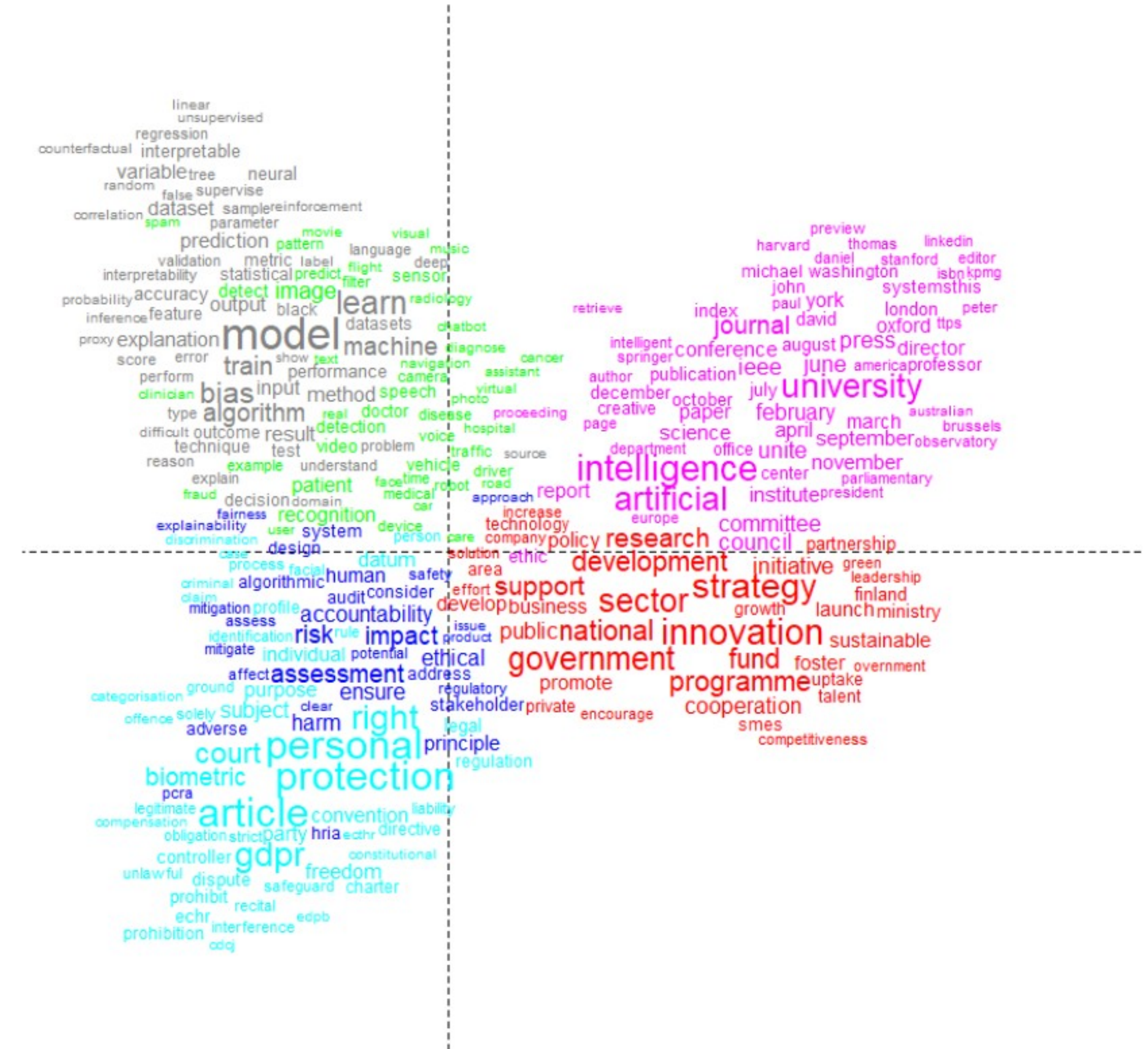
Source: T. Viard



- Pre-2023, a corpus of 1000 articles in French mentioning AI and ethics
- Let's explore it [online](#)

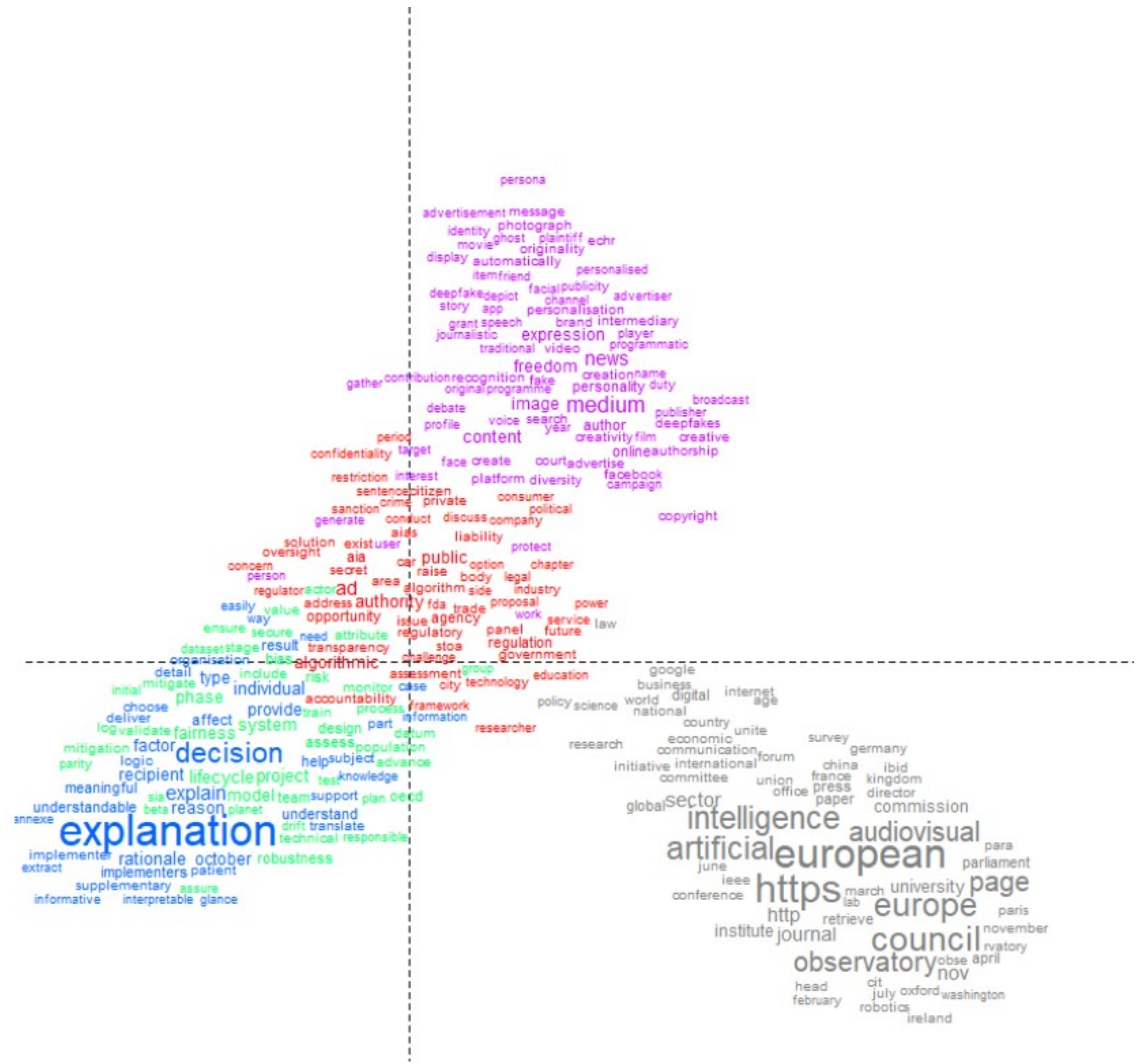
AI ethics through its manifestos

- **A tri-partition of discourses:**
technical, regulatory and
business/innovation



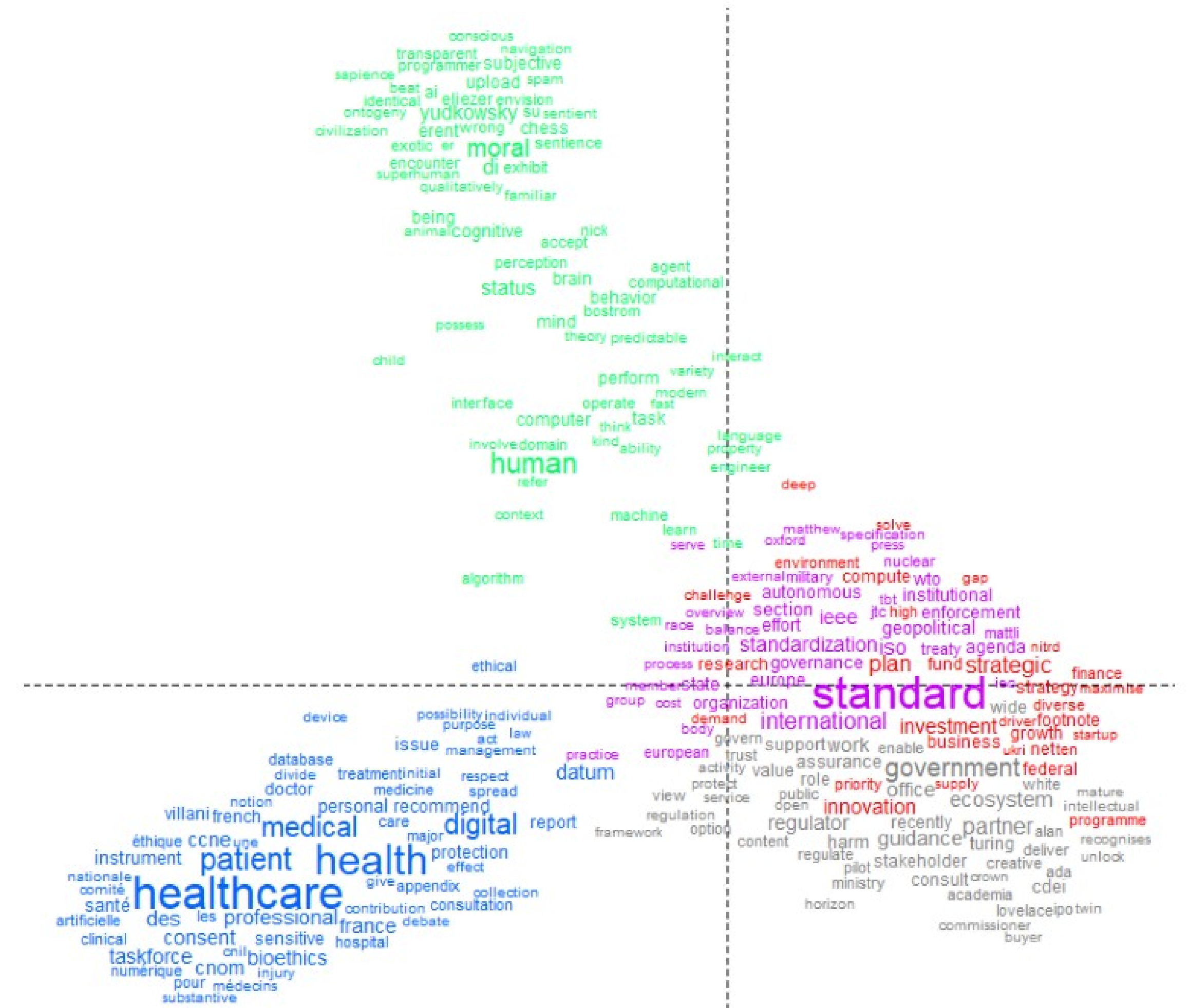
AI ethics through its manifestos

- A tri-partition of discourses: technical, regulatory and business/innovation
- On **explainable AI**: more technical and less regulation



AI ethics through its manifestos

- A **tri-partition of discourses**: technical, regulatory and business/innovation
- On **explainable AI**: more scientific/technical and less regulation
- On **fairness**: no significant change
- On **AGI**: little scientific/technical, little regulation

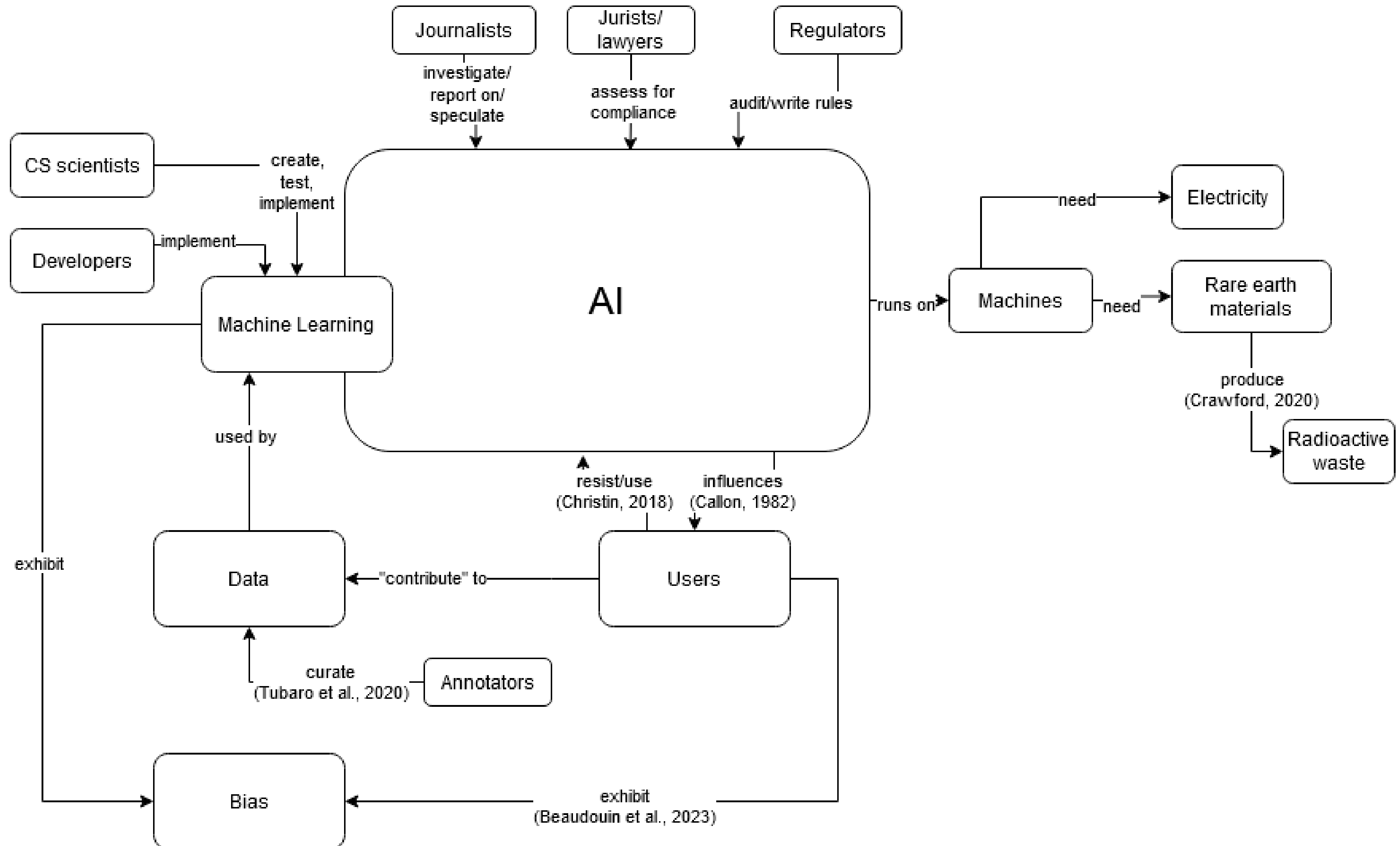


AI ethics through its manifestos

- A **tri-partition of discourses**: technical, regulatory and business/innovation
- On **explainable AI**: more scientific/technical and less regulation
- On **fairness**: no significant change
- On **AGI**: little scientific/technical, little regulation
- On **freedom/autonomy**: heavy focus on regulation and social justice



Back to the AI sociotechnical system



Thank you!