Who decides what's trustworthy?

Standards development for the EU AI Act

James White and Stefan Larsson

5 March, 2025

James White

Postdoktoral forskare

Institutionen för Teknik och Samhälle LTH
Lunds universitet

Fil.dr. sociologi

Stefan Larsson

Docent i teknik och social förändring
Institutionen för Teknik och Samhälle LTH
Lunds universitet

Jur.kand.

Fil.dr. rättssociologi

Tek.dr. fysisk planering

Source: https://www.euractiv.com/section/tech/news/
zuckerberg-calls-on-us-to-defend-tech-companies-against-european-censorship/

Bias baked in

How Big Tech sets its own AI standards

09.01.2025

TECH

Setting ‘harmonised standards’ for artificial intelligence is a crucial element of the EU’s recent AI Act. But who develops these standards? Corporate Europe Observatory reveals how Europe’s standard-setting bodies dealing with AI are heavily skewed towards the tech industry. Behind closed doors, corporations are writing rules that will govern their own AI products – including on how to

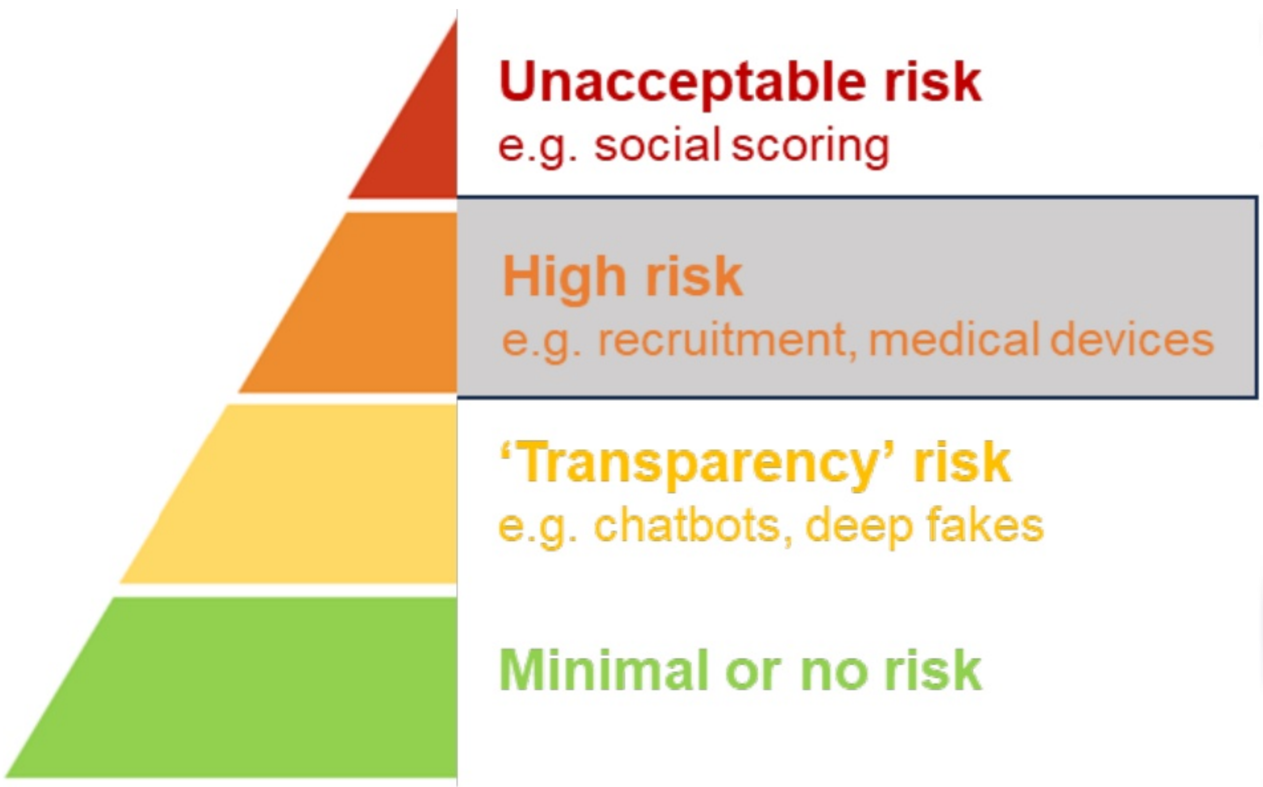
Source: https://corporateeurope.org/en/2025/01/bias-baked

"Standardization is a social process by which humans come to take things for granted. Through standardization, inventions become commonplace, novelties become mundane, and the local becomes universal. It is a historical and therefore contested process whose success depends upon the obfuscation of its founding conflicts and contingencies. Successful standards, if they are noticed at all, simply appear as authoritative, objective, uncontroversial, and natural."

Source: Andrew L. Russell, *Open Standards and the Digital Age*, 2014, p. 16

What are standards and how do they work?

- long history as a form of 'industrial self-governance'
- *de facto*, *de jure* and voluntary (consensus) standards (Russell, 2014)
- design, terminological, performance and procedural (Timmermans & Berg, 2003)
- standards developing organisations (SDOs) harmonise and co-ordinate standards development at different scales
- historical changes in voluntary standards (Yates & Murphy, 2019)
  - a traditional focus on morphological parameters and properties
  - moved towards the standardisation of organisational and business practice since the 1990s
- voluntary standards are developed through expert elicitation and participation
  - which creates obstacles for access and the potential for interference
  - these play out through working groups and national mirror committees



Source: https://publications.jrc.ec.europa.eu/repository/handle/JRC139430

Harmonised standards within the AI Act

- “Normative flexibility” risks unpredictable law and pushes normative decisions to the future and other arenas (Larsson, Hildén, Söderlund, forthcoming)
- Fundamental rights as standardised product safety is new, and raises important questions (Gornet & Maxwell, 2024)
- the role of harmonised standards, CE-marking, conformity assessment to comply with the Act
- CEN/CENELEC given the mandate and tasked to meet ten standardisation deliverables, including:
  - Risk Management
  - Transparency
  - Human Oversight
  - Conformity Assessment
- JTC 21 was established to meet this mandate
  - composed of five working groups
  - 11 published standards, 32 in progress

This dashboard shows the work items and their status in CEN-CENELEC JTC21 as of August 2024.

General note: It can be hard to make sense of some of the information in this status dashboard for anyone not directly involved in JTC21 or not familiar with standardisation at all. For more explanations of standardisation structures, processes and regulations, please refer to https://www.cencenelec.eu/ and https://www.cencenelec.eu/areas-of-work/cen-cenelec-topics/artificial-intelligence/ and https://boss.cen.eu/ as well as your national standardisation organisation with its AI mirror committee.
Comments and contributions to JTC21 have to be routed either through the AI mirror committee in your country, or through one of the Annex 3 organisations (ANEC, ETUC, SBS). Please do not address them to CEN-CENELEC nor the JTC21 management team since they cannot introduce them to the committee.

Wherever possible, work items in JTC21 are based on international work (shades of blue) to minimize divergence and speed up delivery. In some cases, it is necessary and/or more efficient to develop homegrown standards for Europe (black)

Stage code and status notes in this dashboard are only a simplified view; there are more fine-grained internal steps.
Note that there is only a loose correlation between the formal stage, the maturity of technical content that has already been produced, and the degree of consensus that has already been achieved. Hence, the formal stage is not a reliable predictor of completion dates.

Given that much of JTC21 draws on international work, mostly in ISO/IEC JTC1/SC42, the status and project editors there are of interest to JTC21.

These are only partial notes and are not meant to be in any way a comprehensive picture of the much broader work in ISO/IEC JTC1/SC42 overall.

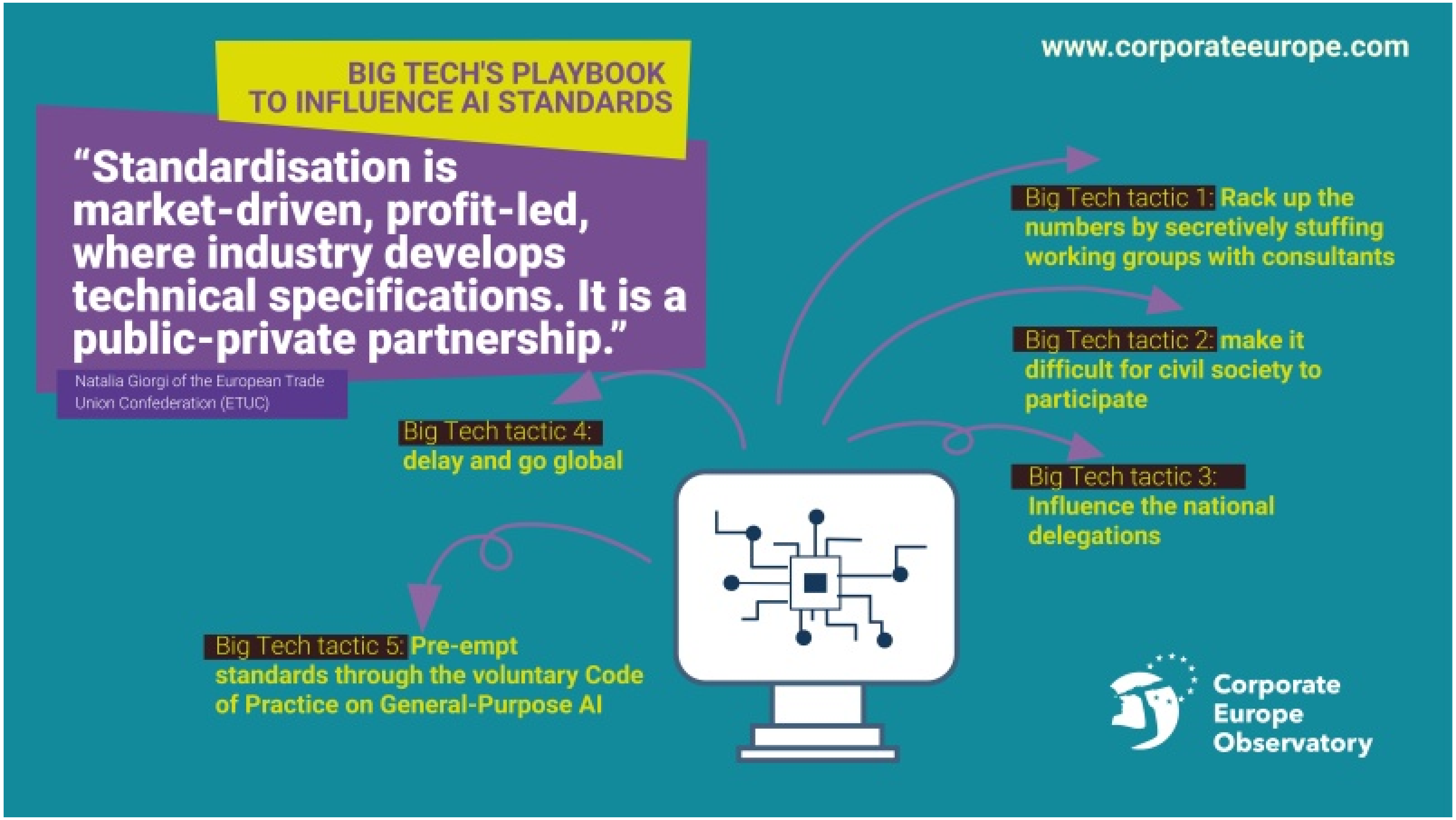
The EU has issued a formal standardisation request (SR) as a mandate for CEN-CENELEC to develop the harmonised standards for the AI Act. It consists of 10 points (with "supporting standards" added for technical reasons). Most work items contribute to fulfilling one or more points in this request since this is the top priority at present.
Note that a couple of items are checkmarked here that will not become harmonised standards but fall into the topic areas covered by the SR.

Colour coding: Shades of blue indicate how closely an item reflects work at the international level. Grey is for homegrown European standards.

Colour coding: Shades of green indicate how far along the process a work item currently is

Thematic group	WI Number	Reference	Type of deliverab	WG	Editor	Title	Relationship with work in SC42 or other applicable committee	Stage code in JTC21 (according to https://standards.cencenelec.eu/dyn/www/f?p=205:105:0:...)	Notes on status in JTC21	Stage code in ISO/IEC	Notes on status of underlying standard in ISO/IEC if applicable	Supporting standards (terminology...	Risk management for AI systems	governance and quality of datasets used to build AI systems	Record keeping through built-in logging capabilities in AI systems	Transparency and information to the users of AI systems	human oversight of AI systems	accuracy specifications for AI systems	robustness specifications for AI systems	cybersecurity specifications for AI systems	quality management system for providers of AI system, including pos market monitoring process	conformity assessment for AI system
Terminology	JT021004, JT021031 (Amendment)	EN ISO/IEC 22989:2023	EN	WG4		Information technology - Artificial intelligence - Artificial intelligence concepts and terminology (ISO/IEC 22989:2022)	Adoption of 22989	Published	EN published	60	Standard published in 2022. Editor: Wei Wei (IBM, Germany)	x										
Terminology	JT021031	EN ISO/IEC 22989/prA1	EN	WG4		Information technology - Artificial intelligence - Artificial intelligence concepts and terminology (ISO/IEC 22989 Amendment 1)	Parallel dev, ISO lead	Under Drafting		60		x										
Terminology	JT021005	EN ISO/IEC 23053:2023	EN	WG4		Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML) (ISO/IEC 23053:2022)	Adoption of 23053	Published	EN published	60	Standard published in 2022. Editor: Milan Patel (Microsoft, UK)	x										
Terminology	JT021032	EN ISO/IEC 23053/prA1	EN	WG4		Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML) (ISO/IEC 23053 Amendment 1)	Parallel dev, ISO lead	Under Drafting		60		x										
Terminology				WG3		Terminology and concepts for domain engineering of AI systems	Parallel dev, ISO lead?		informal preparatory discussions underway. Still to be discussed in WG1		Under discussion in WG3	x										
AI and risk management	JT021011	prEN ISO/IEC 42001		WG2	Marta Janczarski	Information technology - Artificial intelligence - **Management system**	Adoption of 42001	Preliminary	NWIP to activate PWI approved by JTC 21 but with No vote from BE. Tbd if adoption is still needed. Will not be harmonised by the EC.	60	Standard published										x	
AI and risk management	JT021039	prEN XXX	EN	WG2	Adam Leon Smith	AI - **quality management system** for regulatory purposes	Complements, and builds on, several ISO/IEC standards, including 42001	Under Drafting													x	
AI and risk management	JT021016	prEN ISO/IEC 23894	EN	WG2		Information technology - Artificial intelligence - Guidance on **risk management** - ISO/IEC 23894	Adoption of 23894	Published	EN published	60	Standard published		x									
AI and risk management	JT021024	prEN XXX	EN	WG2	Renaud di Francesco	AI **Risk Management**		Under Drafting	Supersedes earlier work item JT021009 (AI Risks - Check List for AI Risks Management				x									
Trustworthiness	JT021021	prEN ISO/IEC 24970	EN	WG3	Adam Leon Smith	AI system **logging**	Parallel dev, ISO lead	Under Drafting		20					x							
Terminology	JT021022	prEN ISO/IEC 12792	EN	WG4	Rania Wazir	**Transparency** taxonomy of AI systems	Parallel dev, ISO lead	Under Drafting		40	Editor: Rania Wazir (leiwand.ai, Austria)					x						
Trustworthiness	JT021015	prEN ISO/IEC 24029-2	EN	WG3	Arnault Ioualalen	Artificial intelligence (AI) - Assessment of the **robustness** of neural networks - Part 2: Methodology for the use of formal methods	Adoption of 24029-2	Preliminary		60	Same project editor (Arnault Ioualalen, Numalis, France)								x			
Trustworthiness		ISO/IEC 24029-3	EN	WG3	Arnault Ioualalen	Artificial intelligence (AI) - Assessment of the **robustness** of neural networks - Part 3: Methodology for the use of statistical methods	Parallel dev?		WG3 discussion points towards parallel development (or alternatively adoption although this takes longer overall)	20	NWIP approved and under development								x			
Trustworthiness		24029-5		WG3	Arnault Ioualalen	Artificial intelligence (AI) - Assessment of the **robustness** of neural networks - Part 5: Applicability of the methodology to other AI algorithms	Parallel dev?		To be clarified to what extent this would be needed to support the SR. Likely to apply to ML, not just neural networks.		NWIP in preparation								?			
Trustworthiness	JT021018	CEN/CLC ISO/IEC TR 24029-1	TR	WG3		Artificial Intelligence (AI) - Assessment of **robustness** of neural networks - part 1: overview	TR 24029-1 already adopted by JTC21 (N270)	Published	24029-1 has already been adopted by JTC21 but since it is a TR, another item still needs to be created in response to the SR. Not much more content is available, though, since this is the frontier of research at the moment.	60	TR published								x			
Trustworthiness	JT021012	ISO/IEC 23282	EN	WG3	Lauriane Aufrant	Evaluation methods for **accurate** natural language processing systems	Parallel dev, ISO lead	Under Drafting			Editor: Rania Wazir (leiwand.ai, Austria). Parallel dev approved, allocated to JWG5; in progress							x				
Trustworthiness	JT021002	ISO/IEC 23281	TR	WG3	Lauriane Aufrant	Artificial Intelligence - Overview of AI tasks and functionalities related to natural language processing	Parallel dev, ISO lead	Under Drafting			Parallel development approved and formally suggested to JTC21; now being done in JWG5	x						x				
Trustworthiness	no WI needed		EN	WG4		Guidelines and requirements for **accuracy** threshold definition			Conclusion in WG4: Will be covered within the Trustworthiness work item. Alignment on scope/depth with EC/JRC needed.									x				
Trustworthiness	no WI needed	ISO/IEC TS 4213	TS			**Accuracy** of classification systems	Reference to 4213	n/a	Conclusion in WG4: Adoption probably not required; TS4213 can be referenced from Trustworthiness work item.		Published in 2022							x				
Trustworthiness	no WI needed	ISO/IEC TS 4213 (update)	TS			**Accuracy** of AI systems for regression, recommendation and clustering	Reference to 4213 ed.2	n/a	Conclusion in WG4: Will be covered in imminent update to TS4213, hence no action needed in JTC21	20	Update in progress							x				
Trustworthiness	JT021008	prEN XXX	EN	WG4	Piercosma Bisconti	Artificial Intelligence **trustworthiness** framework		Under Drafting						x	x	x	x	x	x	x		
Trustworthiness	JT021025	prEN XXX	EN	WG3	Nicolas Cubaud	Evaluation methods for **accurate** computer vision systems		NWIP ballot	Ballot closes 2024-09-12									x				
Trustworthiness			EN	WG	Nicolas Cubaud	Taxonomy of AI tasks in computer vision		NWIP ballot	Ballot closes 2024-09-12			x						x				
Trustworthiness	JT021013	FprCEN/CLC ISO/IEC/TS 12791	TS	WG3	Adam Leon Smith	Information technology - Artificial intelligence - Treatment of unwanted bias in classification and regression machine learning tasks	Parallel dev, ISO lead	Under Ballot	Second DTS for voting in parallel with ISO/IEC JTC 1/SC 42 "Artificial Intelligence" in the period 2024-06-19 to 2024-09-11					x				x				
Trustworthiness	JT021036	prEN XXX	EN	WG3	Michael van Hartskamp	Artificial Intelligence – Concepts, measures and requirements for managing **bias** in AI systems		Under Drafting						x				?	?			
Data governance	JT021020	prEN ISO/IEC 8183	EN	WG3	Domenico Natale	Information technology – Artificial intelligence – Data life cycle framework	Adoption of 8183	Published		60	Published in 2023			x								

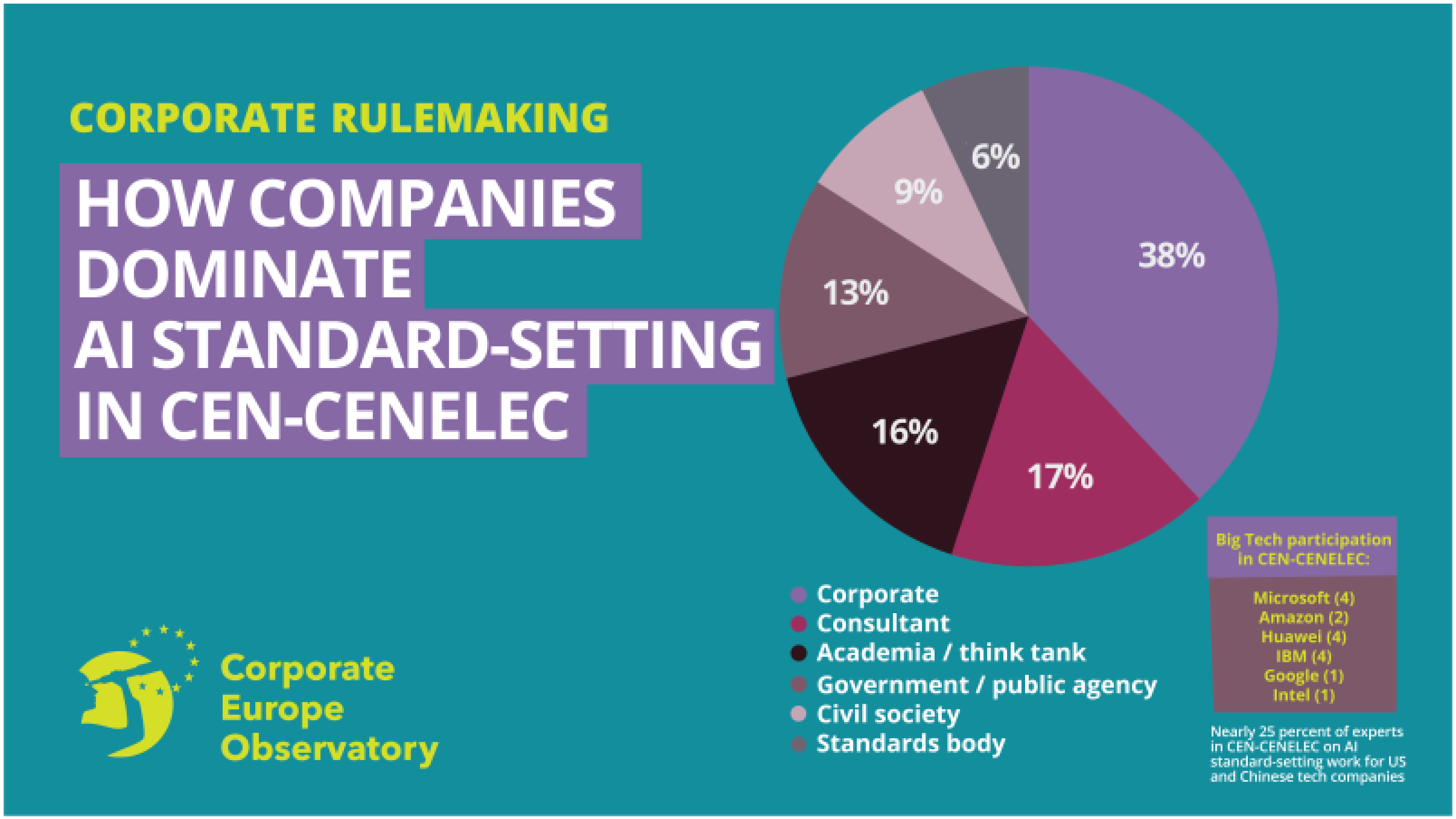
Source: https://www.linkedin.com/posts/sebastianhallensleben_status-dashboard-jtc21-european-ai-standardisation-activity-7235665875112988673-xfEH



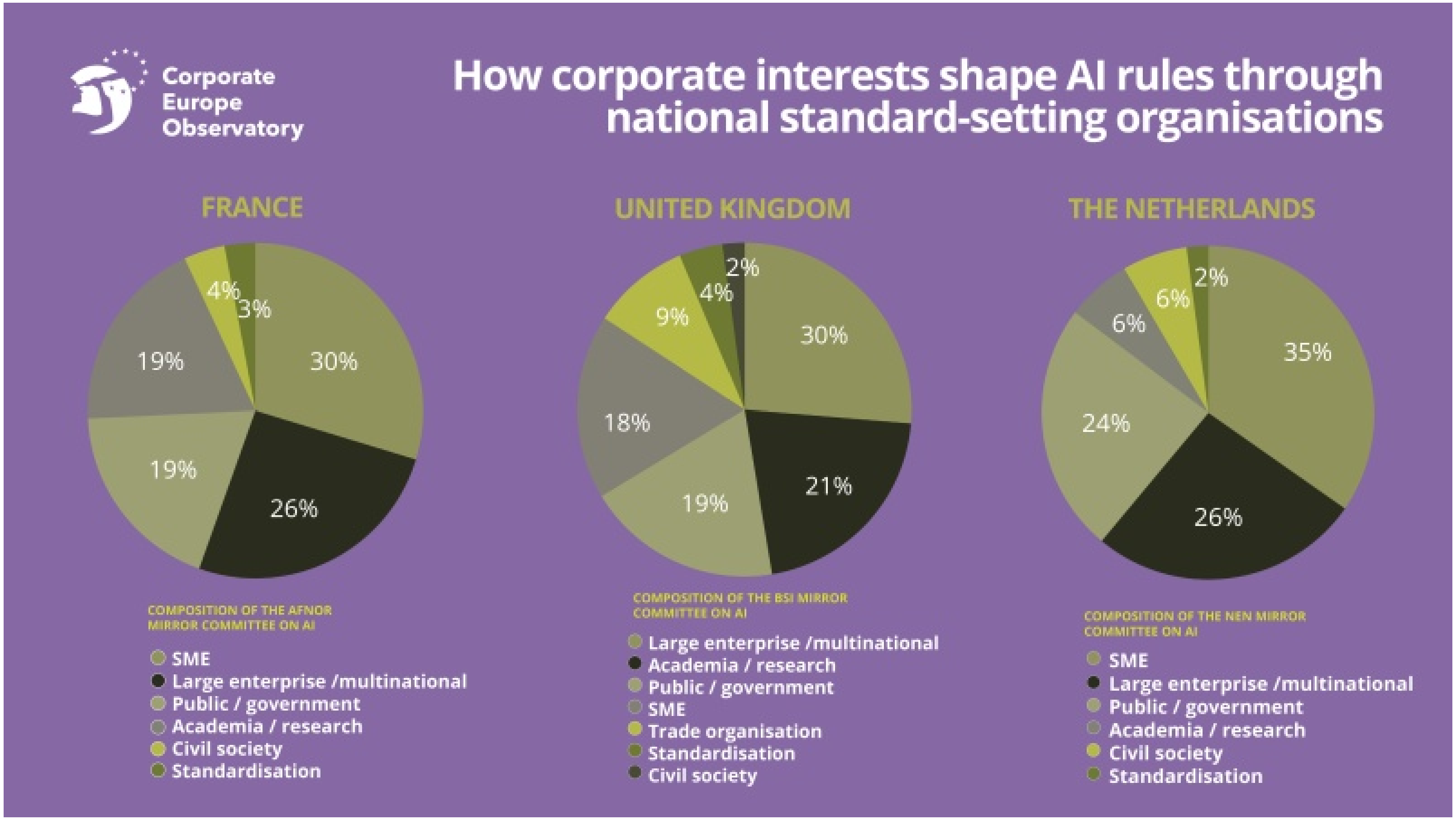
Source: https://corporateeurope.org/en/2025/01/bias-baked

The Corporate Europe Observatory report

- independent research report published January 2025
- based on expert interviews, and document and social media analysis
- claims that the high involvement of technology companies in the development of harmonised standards risks undermining the AI Act
- identify five tactics by which this is being done, most notably
  - stacking working groups with consultants
  - influencing national delegations to CEN/CENELEC
- our experience at Swedish Standards does not support (but neither does it disprove) these claims



Source: https://corporateeurope.org/en/2025/01/bias-baked



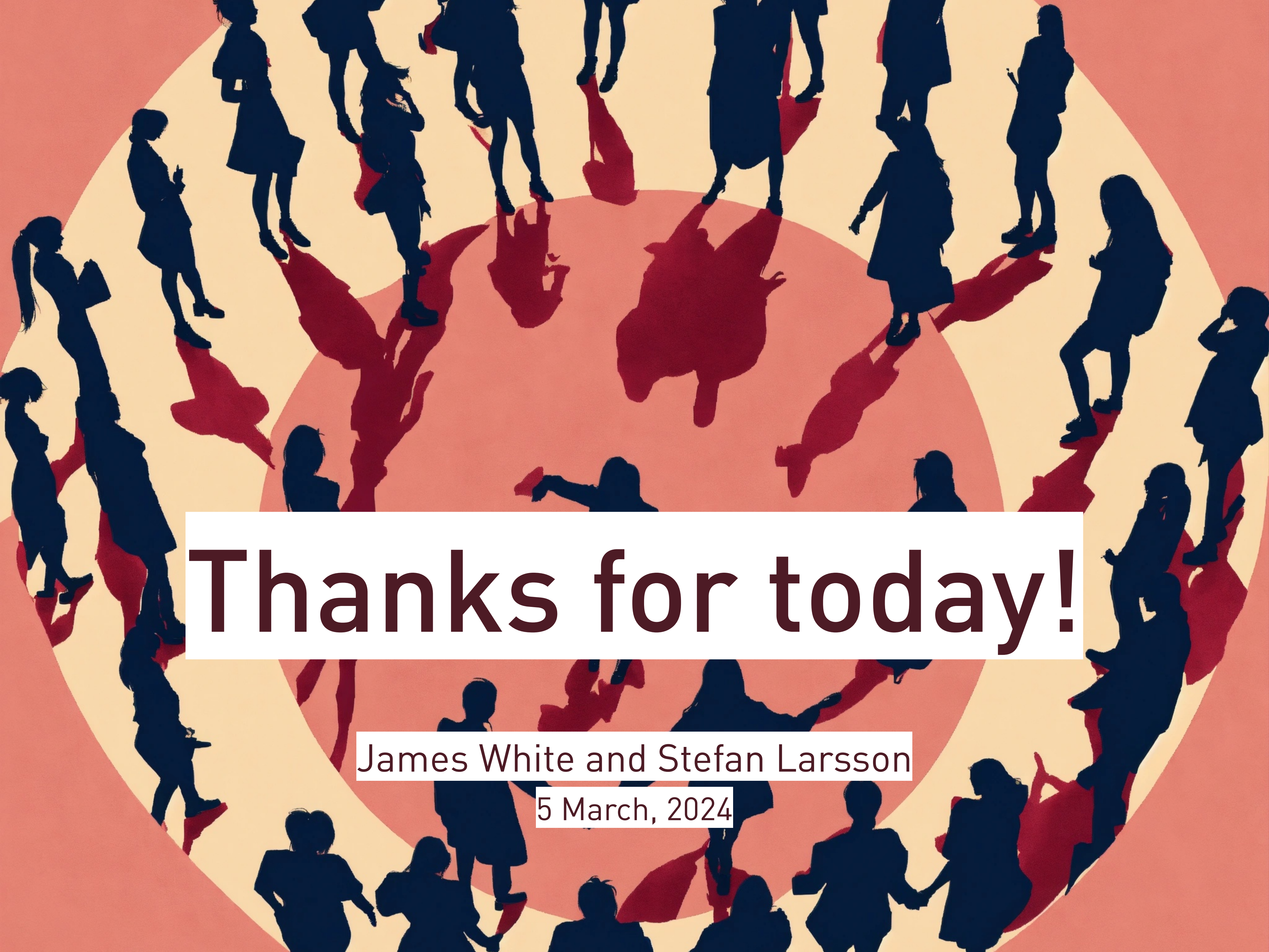
Source: https://corporateeurope.org/en/2025/01/bias-baked

So, who decides what's trustworthy?

- large technology companies are engaged in both direct opposition to and also subtle shaping of EU law
  - both appear to be true
  - and not necessarily in contradiction
- but private and corporate involvement in standards development is driven by a composite of interests
  - some will seek to undermine
  - others to create or find competitive advantage
  - still others to pursue higher ethical and social objectives
- standards emerge through conflict and compromises between of these interests
  - while certainly political, they do not inevitably champion a single, unified cause

Getting involved in standards development

- there is the need for greater pluralism in AI Act standards development
  - overcoming opportunity costs and path dependencies to the low involvement of civil society actors is difficult
  - academics ought to participate much more than they do (and there is precedent for recognition of engagement in science-policy organisations)
- any Lund University employee has the opportunity to get involved
  - AI Lund has supported membership in SIS/TK 421
  - LU AI Standardisation Research Network (LATTICE)



Thanks for today!

James White and Stefan Larsson

5 March, 2024

Resources and references

- Corporate Europe Observatory (2025). *Bias baked in: How Big Tech sets its own AI standards*. https://corporateeurope.org/en/2025/01/bias-baked.
- Edwards, Lilian (2022). *Expert explainer: The EU AI Act proposal*. Ada Lovelace Institute. https://www.adalovelaceinstitute.org/resource/eu-ai-act-explainer/.
- Russell, Andrew L (2014). *Open Standards and the Digital Age: History, Ideology, and Networks*. New York: Cambridge University Press.
- Timmermans, Stefan, & Berg, Marc (2003). *The gold standard: The Challenge of Evidence-Based Medicine and Standardization in Health Care*. Philadelphia: Temple University Press.
- Veale, Michael, & Borgesius, Frederik Zuiderveen (2021). Demystifying the Draft EU Artificial Intelligence Act: Analysing the good, the bad, and the unclear elements of the proposed approach. *Computer Law Review International*, 22(4), 97–112.
- Yates, JoAnne, & Murphy, Craig (2019). *Engineering rules: global standard setting since 1880*. Baltimore: Johns Hopkins University Press.