

From Technocracy to Democracy: Recalibrating AI Governance?

Mehmet Bilal UNVER^{a,1}

^a *University of Hertfordshire, School of Law and Education*

ORCID ID: Mehmet Bilal Unver <https://orcid.org/0000-0001-7952-2634>

Abstract. AI ethics has recently emerged as a dominant governance paradigm, increasingly implemented through institutionalised, specialised and expert-driven tools and mechanisms, thereby compelling us to revisit the question of *democratic deficit*. This ethics-oriented technocratic approach is exemplified by the EU AI Act (2024), which promotes soft-law tools as part of a layered, risk-based governance framework. This article critically examines whether that model can address the democratic deficit in AI governance, focusing on *harmonised standards*, *voluntary codes of conduct* and *regulatory sandboxes*. Drawing on democratic theory and governance scholarship, it explores the extent to which these instruments advance participation, deliberation and accountability. By assessing these mechanisms as key sites of governance, the article concludes that they prioritise flexibility, technical expertise and market integration over public contestation and participation, offering little response to the democratic deficit. It argues that AI ethics can address this deficit only when complemented by institutional reforms that embed deliberation, broaden participation and ensure meaningful public accountability, including the involvement of non-expert citizens.

Key words. AI governance, democracy, AI ethics, EU AI Act, technocracy.

1. Introduction

Artificial intelligence (AI) systems are increasingly embedded in social, economic and political life and accompanying decision-making processes. As AI systems mediate interactions between institutions and members of the society, questions concerning accountability, legitimacy and participation have moved to the forefront of legal and political debate. These developments pose a fundamental challenge to democratic AI governance, giving rise to what is commonly described as the risk of *democratic deficit*. AI can concentrate decision-making power in the hands of technical experts and private corporations, often limiting meaningful public scrutiny, participation and contestation over systems that profoundly affect citizens' lives. In this context, ethical concerns have become inseparable from wider questions about democracy: how collective decisions are made, who participates in them, and whose values shape technological futures?

¹ E-mail address: m.unver@herts.ac.uk.

In response to these challenges, AI ethics has emerged as a leading governance paradigm. Principles such as fairness, transparency, accountability and human oversight have been embedded in ethics guidelines, technical standards, codes of conduct and policy frameworks aimed at aligning AI with societal values and reducing harms. However, the rise of AI ethics has also raised concerns about democracy, as emerging governance models may privilege expert judgement, corporate self-regulation and technocratic decision-making at the expense of public participation and deliberation.

This tension is more explicit in the EU's approach to AI governance. The EU AI Act [23] establishes a risk-based framework for regulating AI while relying heavily on ethics-based governance tools in pursuit of *trustworthy AI*. *Harmonised standards*, *voluntary codes of conduct* and *regulatory sandboxes* shape compliance and the substantive values governing AI systems, yet their growing role raises questions about their ability to support democratic governance [23].

Against this backdrop, this article examines whether these expert-driven mechanisms can address the *democratic deficit* in AI governance, focusing on harmonised standards, voluntary codes of conduct and regulatory sandboxes under the EU AI Act. It evaluates their capacity to foster participation, deliberation and accountability, questioning whether they strengthen democratic legitimacy. While each mechanism offers flexibility and practical benefits, none is inherently democratic. Their effectiveness depends on the participatory safeguards, deliberative spaces and accountability structures accompanying them.

The article proceeds as follows. Section 2 examines the rise of AI ethics; Section 3 explores democracy, legitimacy and power asymmetries in AI governance; Section 4 analyses ethics and democratic governance under the EU AI Act; Section 5 evaluates key governance mechanisms; and Section 6 reassesses their implications for participatory and deliberative democracy before outlining pathways toward more democratic AI governance.

2. The Rise of AI Ethics

Ethics seeks to define principles and standards that guide moral decision-making and responsible conduct across different sectors and institutions ([3], [42]). AI ethics extends these concerns by promoting normative principles and frameworks to ensure that AI systems align with human values, fundamental rights and broader societal interests.

The rise of AI ethics reflects growing concern over the social, political and economic effects of algorithmic decision-making. As AI systems increasingly shape access to services, employment, healthcare, education and security, concerns surrounding bias, discrimination and opacity have intensified. In response, AI ethics has emerged as a governance framework addressing regulatory and legitimacy gaps at organisational, governmental and international levels.

Since the mid-2010s, AI ethics has rapidly institutionalised through ethical guidelines and declarations issued by governments, international organisations and corporations, including the Asilomar AI Principles [26], the OECD AI Principles [48], the European Commission's Ethics Guidelines for Trustworthy AI [33], and UNESCO's Recommendation on the Ethics of Artificial Intelligence [69]. Despite institutional differences, these frameworks largely converge around principles such as fairness, transparency, accountability, privacy and human oversight ([14]; [50]; [36]; [24]).

AI ethics has also evolved beyond abstract principles into concrete governance practices. Ethics boards, advisory committees, algorithmic impact assessments and auditing frameworks increasingly shape organisational compliance and risk management ([54]; [45]). Yet critics argue that AI ethics may legitimise existing power asymmetries through *ethics washing*, where organisations adopt ethical language while resisting binding regulation ([76]; [25]; [5]). Many ethics initiatives remain industry-led, weakly enforceable and limited in public accountability. Ethics boards, expert committees and corporate self-regulation assume quasi-regulatory roles which are however questionable as to their openness, public accountability and legitimacy.

The democratic implications of this development are particularly significant. AI ethics is often produced through expert-driven and technocratic processes, with limited participation from affected individuals and communities, including marginalised groups of people ([51]; [4]). This raises concerns that ethics-based governance may depoliticise fundamentally political questions about social priorities, distributive justice and acceptable risk, relocating them from democratic arenas to technical and managerial domains ([7]; [79]).

The rise of AI ethics therefore represents an ambivalent development. While it has strengthened attention to societal values in AI governance, it may also reinforce technocratic and corporate control by marginalising meaningful public participation. While the need for *democratisation* of AI intersects with ethical principles, democratic legitimacy also requires meaningful opportunities for citizens to shape, contest and influence the norms and institutions that govern them.

3. Democracy in the Age of AI

Democracy, at its core, is a system of collective decision-making that grants individuals equal opportunities to influence matters of public concern ([11]; [31]). It presupposes not only periodic elections but also deliberation, participation and association, namely the capacity of citizens to reason together about common goods ([24]). Modern democratic theory increasingly distinguishes between *participatory* and *deliberative* models: the former emphasises active involvement of citizens in governance, while the latter highlights reason-giving and collective justification as the basis of legitimacy ([15]).

AI challenges democracy in multiple ways. Algorithmic systems increasingly mediate access to information, shape political discourse and determine eligibility for public services like social benefits or credit ([70]; [7]; [79]). These systems often operate opaquely and are developed by actors beyond democratic scrutiny. The opacity of

algorithmic decision-making, along with the concentration of technological power within a handful of global corporations, risks eroding the egalitarian principles upon which democracy rests ([70]).

Governance authority increasingly resides with technology companies, SSOs, auditors and technical experts who embed value judgments into AI design and deployment. However, these actors are often insulated from democratic accountability through proprietary, contractual and transnational arrangements ([42]; [52]; [13]; [73]). This raises concerns about *democratic legitimacy*, commonly understood in terms of both *input legitimacy* (participation and representation) and *output legitimacy* (effective and acceptable outcomes) ([70]; [71]). AI governance frequently prioritises output legitimacy, such as efficiency and accuracy, while weakening input legitimacy through limited transparency and participation ([7]; [51]).

Research also shows that AI systems disproportionately affect marginalised and vulnerable groups, including welfare recipients, racial minorities, migrants and workers in precarious employment ([21]; [3]). This creates a democratic imbalance in which those most affected by automated decision-making have the least influence over its design and deployment ([63]; [71]).

While ethics emerges in response to the increasing penetration of algorithms into everyday life, current trajectories in AI governance risk narrowing the space for democratic contestation and collective value formation, particularly where a *top-down* technocratic approach is adopted. Such an approach is crystallised usually by means of compliance-heavy procedures and/or imposed value proposition. From the perspective of participatory and deliberative democratic theory, this raises concerns about exclusion and inequality. Participatory democracy emphasises the direct involvement of citizens in decision-making processes that affect their lives ([66]), while deliberative democracy stresses the importance of inclusive, reasoned public debate in legitimising collective decisions ([28]; [17]; [77]). AI governance processes, particularly those centred on ethics, standards and technical compliance, often fall short of these ideals. Participation is typically limited to consultation, if present at all, whereas deliberation is constrained by technical complexity and knowledge and power asymmetries ([49]; [67]).

4. Ethics and Democracy under the EU AI Act

The EU AI Act represents the most ambitious attempt to date to establish a comprehensive legal framework for the governance of AI systems. Grounded in the EU's product safety tradition, it adopts a risk-based approach aimed at balancing innovation with the protection of fundamental rights, democratic values and the rule of law. Ethics is central to this framework, with the Act translating principles from the EU's *trustworthy AI* agenda, rooted in the European Commission's Ethics Guidelines for Trustworthy AI ([33]), into binding legal and procedural requirements.

From a democratic perspective, the Act seeks to enhance legitimacy primarily through compliance mechanisms. Its risk-based classification system distinguishes between unacceptable-, high-, limited- and minimal-risk AI, imposing extensive

obligations on high-risk systems, including risk management, data governance, technical documentation, transparency, human oversight, robustness and post-market monitoring. These requirements are designed to operationalise ethical principles in an auditable and harmonised way across the internal market.

However, while the Act strengthens *output legitimacy* by improving safety and reliability, its contribution to *input legitimacy* is more limited. Democratic participation is largely indirect and mediated through expert bodies, regulators and standardisation organisations such as CEN-CENELEC, ETSI and ISO/IEC. Decisions about acceptable risk levels, fairness trade-offs or the scope of human oversight are largely resolved within expert-driven processes, rather than democratic forums ([64]; [71]; [65]). This *delegation* raises concerns about democratic governance and legitimacy, given the limited participation of civil society and affected communities in these processes ([40]).

This technocratic orientation risks reinforcing knowledge and power asymmetries while depoliticising ethical choices by reframing them as technical compliance issues.² Although the Act includes self- and co-regulatory instruments, these may also strengthen industry influence in areas where democratic control remains weak ([66]; [55]).

At the same time, the Act introduces certain ethics-oriented mechanisms, including transparency obligations (Articles 13 and 52), post-market monitoring (Article 72), individual rights (Articles 85-86) and fundamental rights impact assessments (Article 27). These tools offer some scope for scrutiny, but they remain embedded within administrative and compliance frameworks and provide limited opportunities for meaningful public deliberation or collective value formation ([44]; [77]).

In sum, the EU AI Act embodies an insecure model of hybridisation of ethics and democracy. While it represents an important step towards embedding ethical principles into binding law, its compliance-driven and expert-led structure raises ongoing questions about whether it can meaningfully enhance democratic legitimacy in AI governance.

5. Can Ethical Instruments Fill the Democratic Vacuum: A Closer Look at the EU AI Act

As highlighted above, the EU AI Act's technocratic governance approach raises concerns about whether democratic values are genuinely embedded or instead subordinated to ethics-based compliance. Addressing this requires re-examining whether key regulatory tools enhance democratic participation and deliberation. This section therefore evaluates *harmonised standards*, *voluntary codes of conduct* and *regulatory sandboxes*, selected for their relevance to democratic legitimacy under the Act. As self- and co-regulatory mechanisms, these governance tools are designed to support regulatory learning and

² From a philosophical standpoint, this tension mirrors the long-standing debate between Kantian ethics and discursive ethics. Kantian ethics emphasises moral autonomy and duty grounded in rational universality ([37]), whereas discursive ethics, articulated by Habermas ([29]), centres moral validity on processes of communication and mutual justification. In the context of AI governance, these traditions collide: Kantian approaches underpin much of the rule-based, principle-oriented design of ethical frameworks, while discursive approaches would demand open, participatory dialogue about the moral implications of AI. Bridging these paradigms requires reimagining AI ethics as a public practice rather than a set of compliance checklists.

experimentation, potentially opening space for the *democratisation* of AI governance. Drawing on democratic theory and governance scholarship, the following section critically examines their capacity to support meaningful democratic legitimacy.

5.1. Standardisation

Standardisation is a central pillar of contemporary AI governance, translating abstract legal and ethical principles into concrete technical and organisational requirements. Standards developed by SSOs specify processes, metrics and design practices that AI providers must follow to access markets and demonstrate compliance, thereby functioning as a quasi-regulatory mechanism shaping the normative contours of socio-technical systems.

Within the EU AI Act, standardisation plays a decisive role. Under Articles 40 and 41, compliance with harmonised standards developed by bodies such as CEN, CENELEC and ETSI creates a presumption of conformity with legal obligations for high-risk AI systems.³ In practice, standards operationalise requirements relating to transparency, robustness, human oversight and risk management, effectively becoming the main vehicle through which legal norms are applied.⁴ This can be described as *governance by standardisation*, where technical norms substitute for political decision-making ([58]).

While this approach enhances legal certainty, scalability and cross-border harmonisation, it also raises concerns from the perspective of democratic legitimacy. Standard-setting processes are typically dominated by industry actors and technical experts. Participation by civil society organisations, affected communities and lay citizens remains limited, uneven and often resource-dependent ([40]). Deliberation tends to focus on technical feasibility and market needs, while broader questions of justice, distribution and value trade-offs are often marginalised ([38]).

This structure risks shifting norm-setting into expert-driven fora. Although standards are not formally binding, their integration into conformity assessment procedures gives them *de facto* regulatory force. Decisions that shape how AI systems affect society are largely made outside traditional channels of democratic *accountability*.

Moreover, the ethical dimension of AI standardisation remains contested. Standardisation processes operationalise values through checklists, risk metrics and documentation requirements, rather than engaging with substantive ethical disagreement

³ This delegation of standardisation to SSOs reflects a long-standing EU regulatory strategy in product safety and technical regulations. In accordance with Article 40 of the EU AI Act, the European Commission issued a formal standardisation request to CEN and CENELEC in May 2023 (Commission Implementing Decision C(2023) 3215), mandating the development of harmonised standards to operationalise the Act's requirements for high-risk AI systems. However, progress has been slower than anticipated, with several core standards still under development and timelines extending into 2026, raising concerns about regulatory uncertainty during the Act's initial application phase ([18]; [8]; [17]).

⁴ For example, ISO/IEC 23894 on AI risk management and ISO/IEC 42001 on AI management systems set out organisational processes for identifying, assessing and mitigating AI-related risks, thereby shaping how responsibility and accountability are distributed within organisations. Similarly, the IEEE P7000 series – particularly IEEE P7001 on transparency and IEEE P7003 on algorithmic bias considerations – seek to embed ethical reflection into system design lifecycles ([39]).

([47]; [60]). Their capability to translate plural moral values and ethical theories – whether rooted in deontology, utilitarianism, virtue ethics or non-Western traditions – into technical requirements is inherently fraught.

The EU AI Act exemplifies such technocratic approach by prioritising harmonised standards over deeper democratic deliberation. As critics have observed, ethics functions as a seal of approval for AI-based products, “serv[ing] the purpose of appeasing public criticism and pressure as well as helping re-brand European AI, without necessarily guaranteeing substantial ethical reforms within the industry” ([62]; [30]; [76]).

Notably, democratic potential of standard-setting processes depends on their openness, inclusiveness and transparency as well as being subject to democratic oversight. Although initiatives such as IEEE’s *Ethically Aligned Design* framework and proposals ([38]; [22]) to increase civil society participation suggest possible reforms, their impact remains uneven ([32]). Overall, standardisation improves regulatory compliance but only partially contributes to democratisation. Without stronger mechanisms for participation, deliberation and accountability, it risks reinforcing a technocratic model in which key normative choices are made by experts rather than through democratic engagement.

5.2. *Voluntary codes of conduct*

Voluntary codes of conduct represent a second, increasingly institutionalised form of ethics-based AI governance. Unlike technical standards, they are typically principle- and process-oriented instruments adopted by companies, industry bodies or public–private initiatives. The EU AI Act explicitly recognises them as a governance tool, encouraging providers of non-high-risk AI systems to voluntarily apply requirements aligned with those imposed on high-risk systems (Art. 69).

This reflects a broader regulatory trend in digital governance where soft law instruments operate alongside binding rules, delegating a degree of normative discretion to private actors ([72]; [2]). In practice, codes of conduct sit between ethics and law, allowing organisations to determine how far regulatory expectations are internalised beyond mandatory thresholds.

Existing corporate and international frameworks already illustrate this landscape. Companies such as Microsoft and Google have developed internal AI governance principles covering risk assessment, oversight and impact evaluation ([46]; [28]), while instruments like the OECD AI Principles ([48]) and UNESCO’s Recommendation on AI Ethics ([68]) promote voluntary adoption of safeguards including impact assessment, transparency and accountability.

From a democratic perspective, this model is ambivalent. Codes of conduct may enhance democratic responsiveness and support the articulation of shared values and context-sensitive governance. However, their democratic potential is constrained by the dominance of corporate actors in agenda-setting and interpretation, with limited external oversight or enforceability. This weakens accountability and restricts opportunities for affected individuals to contest decisions or seek remedies.

Structural limitations posing further democratic constraints embody a range of powers from agenda-setting to interpretation of norms ([78]). This raises questions about whose values are prioritised in this context. Although codes may formally incorporate ethical principles, their translation into operational norms often lacks external oversight, independent monitoring or enforceable sanctions – features central to democratic accountability ([5]; [67]). This shifts key value judgments, such as acceptable risk or fairness trade-offs, into internal corporate processes rather than public deliberation, raising concerns about the privatisation of norm-setting. Under Article 69 of the EU AI Act, providers retain discretion over which governance elements to adopt, potentially leading to uneven and opaque practices in effect. This discretion can obscure the collective negotiation of value trade-offs that democratic governance ideally entails.

Finally, voluntary codes may inadvertently contribute to *ethics washing*, a phenomenon whereby organisations signal ethical commitment without undertaking substantive governance changes ([30]; [76]; [59]). In fact, analyses of corporate AI ethics initiatives reveal frequent reliance on high-level principles with limited operational specificity or measurable outcomes ([36]; [24]).

Nonetheless, voluntary codes can contribute to democratic governance if they are integrated into hybrid frameworks that combine voluntary commitment with external accountability. Linking codes to public reporting requirements, independent audits, stakeholder review processes or supervisory incentives could strengthen their legitimacy and enforceability. For instance, the conformity assessment and post-market monitoring procedures could be extended to codes through requirements for transparency reporting or third-party evaluation of voluntarily adopted governance practices. Similarly, multi-stakeholder coalitions that include civil society in code development and review may help align codes with broader social expectations rather than internal corporate priorities. Without such elements, codes risk reinforcing a technocratic, corporate-centric model of ethical governance, where ethical commitments serve as badges of legitimacy rather than vehicles for empowered democratic engagement.

5.3. *Regulatory sandboxes*

Regulatory sandboxes constitute a third ethics-led governance mechanism designed to reconcile innovation with risk mitigation in AI governance. They are controlled environments in which regulators and innovators jointly test and refine regulatory approaches based on real-world evidence ([27]). The EU AI Act recognises sandboxes as formal instruments (Arts 53–54), intended to support responsible innovation while safeguarding fundamental rights, democratic values and the rule of law.

Member States are required to establish sandboxes enabling providers to test AI systems – particularly novel or high-risk applications – before market deployment. Importantly, participation does not exempt actors from core legal obligations, including fundamental rights and safety requirements (EU AI Act, Article 57(6)(12)). Instead, sandboxes operationalise governance-by-experimentation, allowing regulators and developers to assess risks and mitigation strategies in practice.

Sandboxes have already been used across Europe and the UK. The UK Information Commissioner's Office (ICO) sandbox has supported AI applications in areas such as

healthcare and credit scoring ([10]; [34]), while coordinated initiatives under the Digital Regulation Cooperation Forum have promoted cross-sector experimentation ([35]). In the EU, pilot sandboxes have been used in fields such as healthcare, mobility and public administration, often focusing on explainability, bias and human oversight ([75]).⁵

From a democratic perspective, sandboxes are often framed as participatory tools that enable iterative dialogue between stakeholders. By exposing regulators to socio-technical contexts of AI deployment, sandboxes can support not only innovation but also reflexive regulation and adaptive learning, addressing the limitations of static, *ex ante* legislative approaches ([19]; [56]; [27]).

However, their democratic potential is limited in practice. Participation is typically restricted to selected firms, regulators and technical experts, while affected individuals and communities are largely excluded. Key decisions about design, risk thresholds and evaluation remain concentrated in expert and corporate hands, raising concerns that sandboxes prioritise innovation and market access over public deliberation, particularly where public participation is limited to *post hoc* consultation or transparency reporting (see [55]). Similar concerns emerged in the deployment of autonomous vehicle regulatory sandboxes in jurisdictions such as the UK and the US, where industry and regulators largely shaped testing parameters, while local communities and road users had little meaningful input into the governance of public road trials ([15], [68]).

While sandboxes may incorporate ethical safeguards such as algorithmic impact assessments, human-in-the-loop requirements or bias testing, these mechanisms are typically evaluated internally or by regulators, rather than through participatory or adversarial procedures accessible to citizens (see [19]; [61]). As a result, citizens have limited opportunities to shape or contest outcomes during experimentation phases and newly introduced individual rights (Art 85-86) offer only limited corrective capacity.

The democratic value of sandboxes therefore depends heavily on institutional design. Where they include civil society participation, user representation and transparent evaluation processes, they may function as deliberative spaces for negotiating societal values. Similarly, embedding sandboxes within public accountability frameworks through mandatory transparency reporting, independent evaluation and parliamentary or judicial oversight could mitigate risks of regulatory capture and technocratic drift ([56]; [1]).

Overall, while regulatory sandboxes enhance learning and regulatory flexibility, their contribution to democratising AI governance remains constrained. Without stronger participatory structures and accountability mechanisms, they risk reinforcing an experimental governance model that privileges expertise and innovation over democratic contestation and collective reasoning.

⁵ For a comparative analysis of regulatory sandboxes across different jurisdictions, examining their institutional design, governance models and practical implementation in sectors such as fintech and AI, see [61].

6. What is Lost in Translation? Re-evaluating the Prospect of Democratic AI Governance

The analysis of *harmonised standards*, *voluntary codes of conduct* and *regulatory sandboxes* shows that while these mechanisms help operationalise ethical principles, they also have structural limits that contribute to a *democratic deficit* in AI governance. This section synthesises their regulatory implications, examines weaknesses in participation, deliberation and accountability, and outlines pathways for democratisation from the perspective of participatory and deliberative theories.

6.1. Participatory Deficits

Although each mechanism nominally allows stakeholder input, participation remains uneven in practice. In standardisation, engagement is often dependent on technical expertise and organisational resources, which privileges industry actors and specialist communities over civil society and affected publics ([22]; [40]). Voluntary codes of conduct, even when informed by consultation processes, tend to reflect broad ethical consensus rather than the situated concerns of marginalised groups ([30]; [76]). Sandboxes, while enabling real-world experimentation, are typically limited to selected firms and regulators, with restricted involvement of wider societal actors ([55]).

6.2. Deliberative Deficits

The translation of ethical principles into standards, codes and sandbox practices often prioritises procedural compliance over substantive normative debate. Although values such as fairness and transparency are widely endorsed, their meaning in context is frequently resolved through technocratic processes rather than open deliberation over competing conceptions of distributive justice and equity ([30]; [76]; [41]; [62]). SSOs, corporate ethics frameworks and sandbox environments tend to prioritise feasibility and efficiency, limiting space for contestation over alternative social and ethical visions. Yet, deliberative democratic theory underscores that public reasoning and collective justification over value trade-offs is central to legitimate governance ([29]; [15]).

6.3. Accountability Gaps

A further concern across all three instruments is the weakness of accountability structures. Standardisation lacks direct democratic oversight; voluntary codes of conduct are typically non-binding and rely on reputational enforcement; and regulatory sandboxes provide experimental flexibility without robust mechanisms for remedy when harms occur. These characteristics have led to critiques of *ethics washing*, where ethical commitments function more as symbolic signals than enforceable obligations ([76]; [59]; [30]). Accountability requires not only transparency but also mechanisms for contestation, redress and answerability to affected communities ([74]; [12]). By contrast, the current governance ecosystem tends to confine ethical oversight within expert committees, regulatory bodies or corporate governance structures, limiting the ability of individuals and groups affected by AI systems to challenge decisions or seek remedies.

Overall, selected governance tools collectively illustrate a broader shift of normative authority away from democratically accountable institutions toward expert, regulatory and private actors. Standards under Article 40 institutionalise technical and ethical norms through SSO processes shaped largely by industry and experts. Codes of conduct under Article 69 enable providers to voluntarily extend compliance beyond mandatory requirements, reinforcing self-regulatory dynamics. Sandboxes facilitate regulatory experimentation and innovators but remain largely insulated from broader societal participation. These mechanisms reveal structural participatory, deliberative and accountability deficits in the EU AI governance framework, as summarised in Table 1.

Table 1. Governance tools and their democratic constraints

Governance Tool	Participation Status	Deliberation Status	Accountability Status
Harmonised standards	Expert-privileged: Participation is contingent on organisational capacity and technical expertise, favouring corporate actors and experts over civil society and the lay public.	Technocratic Moralism: Value trade-offs are resolved through technocratic considerations and feasibility rather than democratic debate over competing values.	Direct deficit: Standardisation processes currently lack direct political accountability.
Voluntary codes of conduct	Top-down consensus: Often reflects high-level ethical consensus rather than the specific concerns of marginalised groups affected by AI.	Power Asymmetry: While deliberation is central, it is frequently shaped by power imbalances that risk overshadowing public goods and concerns.	Symbolic compliance: Mechanisms generally lack binding enforcement or sanctions beyond reputational effects, making them susceptible to ethics washing.
Regulatory sandboxes	Market-oriented: Involvement is typically limited to a small set of participants assessed primarily on regulatory and commercial criteria.	Optimisation-focused: These environments prioritise technical feasibility and efficiency over open normative contestation and divergent social visions.	Temporary flexibility: Grant temporary regulatory flexibilities without providing clear avenues for redress or substantive behavioural change when harms occur.

The regulatory landscape shaped by these instruments shifts ethical governance toward corporate discretion rather than public deliberation and participation. This delegation of normative authority has been criticised for lacking a sustainable democratic foundation. As democratic theory highlights, expertise cannot replace public reasoning and accountability; democratic legitimacy depends on opportunities for those subject to regulation to influence, contest and justify norms through inclusive and publicly accessible processes ([29]; [15]; [20]).

7. Conclusion: Towards Democratic AI Governance

The EU's shift toward ethics-led governance reflects an ambivalent regulatory landscape in which normative authority has been partially transferred from democratically accountable institutions to technical experts and private actors. Although EU governance tools aim to embed ethical principles across the AI lifecycle, their design raises concerns about meaningful public engagement in the formation of those norms.

Under the current framework, *harmonised standards*, *voluntary codes of conduct* and *regulatory sandboxes* play a central role in shaping both the ethical and democratic dimensions of AI governance. If ethics is to address the democratic vacuum in AI governance, it must itself be democratised. This requires moving beyond expert consensus and corporate self-regulation towards participatory and deliberative structures that enable affected publics to influence, contest and hold accountable AI systems and their governance. To this end, the key mechanisms could be reconfigured as follows:

- Democratising standardisation through participatory oversight: Standard-setting must be reclaimed as an inherently political process. This requires extending democratic oversight to technical committees by mandating participation quotas for public interest representatives, maintaining transparent records of deliberation and conducting rigorous impact assessments of the value choices embedded in technical requirements.
- Strengthening accountability in voluntary codes: Voluntary codes must be made externally verifiable through institutional oversight and accountability mechanisms, which should include linking these codes to external monitoring, independent audits and credible sanctions through regulatory certification schemes, ensuring alignment with civic values rather than corporate self-regulation.
- Transforming sandboxes into deliberative Arenas: Regulatory sandboxes should evolve beyond testing environments into participatory spaces, incorporating periodic reviews of societal impact, citizens' panels, participatory technology assessments and structured stakeholder deliberation on the societal implications of AI systems.

In conclusion, ethics-led governance tools under the EU AI Act cannot, on their own, remedy the democratic deficit they seek to address. While *harmonised standards*, *voluntary codes of conduct* and *regulatory sandboxes* enhance flexibility and responsiveness, their current design privileges technocratic and expert-driven forms of governance that limit meaningful public participation. Reconfiguring these mechanisms along democratic lines is therefore a normative necessity. Embedding *participation*, *accountability* and *deliberation* within them would enable ethics to function not merely as a governance technique, but as a vehicle for democratic AI governance capable of addressing the underlying democratic vacuum.

References

- [1] Anjos LC, Leurquin P. Regulatory experimentation as a panacea in emerging markets: The Brazilian Central Bank's approach to financial innovation through a regulatory sandbox. *Cambridge Forum on AI: Law and Governance*. 2025; 1:e41. doi: 10.1017/cfl.2025.10016.
- [2] Bąkowski T. *Regulating artificial intelligence: Binary ethics and the law*. London: Routledge; 2021.
- [3] Beauchamp TL, Childress JF. *Principles of biomedical ethics* (5th ed.). Oxford: Oxford University Press; 2001.
- [4] Benjamin R. *Race after technology: Abolitionist tools for the new Jim Code*. Cambridge: Polity Press; 2019.
- [5] Bietti E. From ethics washing to ethics bashing: A view on tech ethics from within moral philosophy. In: *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAT*)*; 2020 Jan; New York, NY, USA. ACM; 2020. p. 210–19. doi: 10.1145/3351095.3372860
- [6] Brennen S, Kreiss D. Digitalization. In: Jensen KB, Rothenbuhler EW, Pooley JD, Craig RT, editors. *The International Encyclopedia of Communication Theory and Philosophy*. Wiley-Blackwell; 2016. p. 1–11. doi: 10.1002/9781118766804.wbiect111.
- [7] Cavaliere P, Romero G. From poisons to antidotes: Algorithms as democracy boosters. *Eur J Risk Regul*. 2022; 13(2):421–42. doi: 10.1017/err.2021.57.
- [8] CEN-CENELEC. Update on CEN and CENELEC's Decision to Accelerate the Development of Standards for Artificial Intelligence. 2025. <https://www.cencenelec.eu/news-events/news/2025/brief-news/2025-10-23-ai-standardization/>.
- [9] Centre for Data Ethics and Innovation. *Centre for Data Ethics and Innovation: Two year review*. Government of the United Kingdom; 2021. https://assets.publishing.service.gov.uk/media/60e2d02d3bf7f7c30482802/Centre_for_Data_Ethics_and_Innovation_-_Two_year_review.pdf.
- [10] Choi SJ. ICO opens beta phase of privacy “regulatory sandbox”. Covington, 2 April 2019. <https://www.insideprivacy.com/international/united-kingdom/ico-opens-beta-phase-of-privacy-regulatory-sandbox/>.
- [11] Christiano T, Bajaj S. Democracy. In: Zalta EN, Nodelman U, editors. *The Stanford Encyclopedia of Philosophy* (Summer 2024 Edition). 2024. <https://plato.stanford.edu/archives/sum2024/entries/democracy/>.
- [12] Cobbe J, Lee MS, Singh J. Reviewable automated decision-making: A governance framework. *Legal Stud*. 2023; 43(4):640–60.
- [13] Cohen, JE. *Between truth and power: The legal constructions of informational capitalism*. Oxford: Oxford University Press; 2019.
- [14] Corrêa NK, Galvão C, Santos JW, Del Pino C, Pontes Pinto E, Barbosa C, et al. Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance. *Patterns*. 2023; 4(10):100857. doi: 10.1016/j.patter.2023.100857.
- [15] Department for Transport. *Future of transport regulatory review: Regulatory sandboxes - Government response*. London: GOV.UK; 2021. <https://www.gov.uk/government/consultations/future-of-transport-regulatory-review-regulatory-sandboxes>.
- [16] Dryzek JS. *Deliberative democracy and beyond: Liberals, critics, contestations*. Oxford: Oxford University Press; 2002.
- [17] Euronews. EU standards bodies flag delays to work on AI Act. Euronews; 16 Apr 2025.
- [18] European Commission. Commission implementing decision C(2023) 3215 of 22 May 2023 on a standardisation request to the European Committee for Standardisation and the European Committee for Electrotechnical Standardisation in support of the Artificial Intelligence Act. European Union. 2023. https://single-market-economy.ec.europa.eu/publications/standardisation-request-artificial-intelligence-act_en.
- [19] Español AG, Koenig PD. Regulatory sandboxes for AI in the majority world: A learning-centric approach to legal adaptation. *Cambridge Forum on AI: Law and Governance*. 2025; 1:e42. doi: 10.1017/cfl.2025.10015.
- [20] Elstub S. Deliberative and participatory democracy. In: Bächtiger A, Dryzek JS, Mansbridge JJ, editors. *The Oxford Handbook of Deliberative Democracy*. Oxford: Oxford University Press; 2018. p. 491–505.
- [21] Eubanks V. *Automating inequality: How high-tech tools profile, police, and punish the poor*. New York, NY: St. Martin's Press; 2018.
- [22] European Parliament. Resolution of 9 May 2023 on a standardisation strategy for the single market (2022/2058(INI)). European Union. 2023.
- [23] European Parliament and Council. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending multiple regulations and directives (Artificial Intelligence Act). *Official Journal of the European Union*. 2024.

- [24] Fjeld J, Achten N, Hilligoss H, Nagy AC, Srikumar M. Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI. Berkman Klein Center for Internet & Society Research Publication No. 2020-1, Harvard University; 2020.
- [25] Floridi L. Translating principles into practices of digital ethics: Five risks of being unethical. *Philos Technol.* 2019; 32(2):185–93.
- [26] Future of Life Institute. Asilomar AI principles. 2017. <https://futureoflife.org/open-letter/ai-principles/>.
- [27] Genicot N, Guimaraes Moraes T. Exploring the boundaries of AI regulatory sandboxes under the AI Act: Flexibility and real-world testing. *Cambridge Forum on AI: Law and Governance.* 2025; 1:e36. doi: 10.1017/cfl.2025.10013.
- [28] Google. AI at Google: Our principles. 2018. <https://ai.google/static/documents/ai-principles-2023-progress-update.pdf>.
- [29] Habermas J. Between facts and norms: Contributions to a discourse theory of law and democracy. Translated by Rehg W. Cambridge, MA: MIT Press; 1996.
- [30] Helm P, Gerlek S. Empirical AI ethics: Reconfiguring ethics towards a situated, plural, and transformative approach. *arXiv.* 2025. <https://arxiv.org/abs/2509.17727>.
- [31] Himmelreich J. Should we automate democracy? In: Véliz C, editors. *The Oxford Handbook of Digital Ethics.* Oxford: Oxford University Press; 2024. p. 631–684.
- [32] IEEE. Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems. IEEE Standards Association. 2019. https://standards.ieee.org/ieee/White_Paper/10594/.
- [33] Independent High-Level Expert Group on Artificial Intelligence. Ethics guidelines for trustworthy artificial intelligence. European Commission. 2019. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.
- [34] Information Commissioner's Office (ICO). Regulatory sandbox. 2025. <https://ico.org.uk/for-organisations/advice-and-services/regulatory-sandbox/>.
- [35] Information Commissioner's Office (ICO). Digital Regulation Cooperation Forum. 2020. <https://ico.org.uk/about-the-ico/what-we-do/digital-regulation-cooperation-forum/>.
- [36] Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. *Nat Mach Intell.* 2019; 1(9):389–99.
- [37] Kant I. Groundwork of the metaphysics of morals. Translated by Gregor M. Cambridge: Cambridge University Press; [1785] 2012.
- [38] Kilian R, Jäck L, Ebel D. European AI standards - Technical standardisation and implementation challenges under the EU AI Act. *Eur J Risk Regul.* 2025; 16(3):1038–62. doi: 10.1017/err.2025.10032.
- [39] Koene A, et al. A governance framework for algorithmic accountability and transparency. European Parliamentary Research Service; 2019 Apr. [https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624262/EPRS_STU\(2019\)624262_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624262/EPRS_STU(2019)624262_EN.pdf).
- [40] Laux J, Wachter S, Mittelstadt B. Three pathways for standardisation and ethical disclosure by default under the European Union Artificial Intelligence Act. *Comput Law Secur Rev.* 2024; 53:105957. doi: 10.1016/j.clsr.2024.105957.
- [41] Laux J, Wachter S, Mittelstadt B. Trustworthy artificial intelligence and the European Union AI Act: On the conflation of trustworthiness and acceptability of risk. *Reg Gov.* 2023. doi: 10.1111/rego.12512.
- [42] MacIntyre A. *After virtue: A study in moral theory* (2nd ed.). Notre Dame: University of Notre Dame Press; 1984.
- [43] Mansell R, Durach F, Kettemann M, Lenoir T, Procter R, Tripathi GP. Information ecosystems and troubled democracy: The state of knowledge on news media, AI, and data governance. Gothenburg: Nordicom, University of Gothenburg; 2025. doi: 10.48335/9789189864238.
- [44] Mantelero A. AI and big data: A blueprint for a human rights, social and ethical impact assessment. *Comput Law Secur Rev.* 2018; 34(4):754–72. doi: 10.1016/j.clsr.2018.05.017.
- [45] Metcalfe J, Moss E, Watkins EA, Singh R, Elish MC. Algorithmic impact assessments and accountability: The co-construction of impacts. In: *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAccT'21)*; 2021 Mar; Virtual Event. New York (NY): ACM; 2021. doi: 10.1145/3442188.3445935.
- [46] Microsoft. Microsoft responsible AI standard (Version 2). 2022.
- [47] Nathan C, Poder I. Towards a comprehensive framework for ethical and responsible standardisation. *Ethics Inf Technol.* 2025; 27:27. doi: 10.1007/s10676-025-09839-0.
- [48] OECD. Recommendation of the Council on Artificial Intelligence. 2024. <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>.
- [49] O'Neil C. *Weapons of math destruction: How big data increases inequality and threatens democracy.* London: Penguin Books; 2016.

- [50] Palladino N. A 'biased' emerging governance regime for artificial intelligence? How AI ethics get skewed moving from principles to practices. *Telecomm Policy*. 2023; 47(5):102479. doi: 10.1016/j.telpol.2022.102479.
- [51] Parthasarathy A, Phalnikar A, Jauhar A, Somayajula D, Krishnan GS, Ravindran B. Participatory approaches in AI development and governance: A principled approach. *arXiv*. 2024. <https://arxiv.org/abs/2407.13100>.
- [52] Pasquale F. *The black box society: The secret algorithms that control money and information*. Cambridge, MA: Harvard University Press; 2015.
- [53] Pateman C. *Participation and Democratic Theory*. Cambridge: Cambridge University Press; 1970.
- [54] Raji ID, Smart A, White RN, Mitchell M, Gebru T, Hutchinson B, et al. Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In: *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT'20)*; 2020 Jan; New York, NY, USA. ACM; 2020. p. 33–44. doi: 10.1145/3351095.3372873.
- [55] Ranchordás S. Regulatory sandboxes and innovation-friendly regulation: Between collaboration and capture. *Ital J Public Law*. 2024; 1. doi: 10.2139/ssrn.4696442.
- [56] Ruschmeier H. Experimental regulation for AI governance with regulatory sandboxes. *Cambridge Forum on AI: Law and Governance*. 2025; 1:e38. doi: 10.1017/cfl.2025.10035.
- [57] Sankaran A. Enhancing trust through standards: A comparative risk-impact framework for aligning ISO AI standards with global ethical and regulatory contexts. *arXiv*. 2025. <https://arxiv.org/abs/2504.16139>.
- [58] Schepel H. *The constitution of private governance: Product standards in the regulation of integrating markets*. Oxford: Hart Publishing; 2005.
- [59] Schultz MD, Conti L, Seele P. Digital ethicswashing: A systematic review and a process perception outcome framework. *AI Ethics*. 2025; 5(2):805–18. doi.org/10.1007/s10551-022-05054-9.
- [60] Sciacca F, Floridi L. From principles to practice: Proceduralisation and governance of AI ethics. *AI Ethics*. 2024. doi: 10.48550/arXiv.2412.04683.
- [61] Seferi FA. comparative analysis of regulatory sandboxes from selected use cases: Insights from recurring operational practices. In: *Regulatory sandboxes for AI and cybersecurity – Questions and answers for stakeholders*. Cybersecurity National Lab; 2025. <https://cybersecnatlab.it/wp-content/uploads/2025/02/CybersecNatLab-White-Paper-Regulatory-Sandboxes.pdf>.
- [62] Siapka A. Towards a feminist metaethics of AI. In: *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society (AIES '22)*. ACM; 2022. p. 665–674. doi: 10.1145/3514094.3534197.
- [63] Sieber R, Brandusescu A, Adu-Daako A, Sangiambut S. Who are the publics engaging in AI? *Public Understanding of Science*. 2024; 33(5):634–653.
- [64] Smuha N. Algorithmic rule by law: How algorithmic regulation in the public sector erodes the rule of law. Cambridge: Cambridge University Press; 2025.
- [65] Smuha N, Ahmed-Rengers E, Harkens A, Li W, MacLaren J, Piselli R, Yeung K. How the EU can achieve legally trustworthy AI. *LEADS Lab, University of Birmingham*; 2021. <https://leads.blogs.bham.ac.uk/2021/06/15/trustworthy-ai/>.
- [66] Smuha N, Radu R. Delegating the law of artificial intelligence: Governance, expertise and democratic accountability. *EUI Working Paper, European University Institute*; 2022.
- [67] Taylor RR, Murphy JW, Hoston WT, et al. Democratizing AI in public administration: Improving equity through maximum feasible participation. *AI Soc*. 2025. doi: 10.1007/s00146-024-02120-w.
- [68] Tennant C, Stares S, Vucevic S, et al. *Driverless Futures? A survey of US public attitudes (Report)*. London: Department of Science & Technology Studies, University College London; 2022. <https://openaccess.city.ac.uk/id/eprint/29210/>.
- [69] UNESCO. *Recommendation on the Ethics of Artificial Intelligence*, adopted 23 November 2021. Paris: United Nations Educational, Scientific and Cultural Organization; 2021. <https://unesdoc.unesco.org/ark:/48223/pf0000380455>.
- [70] Unver MB. Exploring Democratic AI Governance: A Cross-Jurisdictional Inquiry. *J AI Law Regul*. 2025; 2(4):337–59.
- [71] Unver MB. Shaping Democratic AI Governance: A critical look at EU's AI Act and beyond. In: Arvidsson M, Tranter K, editors. *The Routledge Handbook on AI, Law and Society*. Routledge; 2026 (In press).
- [72] Varošanec I. On the path to the future: Mapping the notion of transparency in the EU regulatory framework for AI. *Int Rev Law Comput Technol*. 2022; 36(2):95–117.
- [73] Veale M, Matus K, Gorwa R. AI and Global Governance: Modalities, Rationales, Tensions. *Annu Rev Law Soc Sci*. 2023; 19:255–75. doi: 10.1146/annurev-lawsocsci-020223-040749.
- [74] Veale M, Van Kleek M, Binns R. Fairness and accountability design needs political philosophy. *Philos Technol*. 2024; 37(1):1–27.

- [75] Veale M, Zuiderveen Borgesius F. Demystifying the Draft EU Artificial Intelligence Act – Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach. *Comput Law Rev Int.* 2021; 22(4):97–112. doi: 10.9785/cri-2021-220402.
- [76] Wagner B. Ethics as an Escape from Regulation: From Ethics Washing to Ethics Shopping? In: Hildebrandt M, editor. *Being Profiling. Cogitas ergo sum.* Amsterdam: Amsterdam University Press; 2018. p. 86–90.
- [77] Wieringa M. What to account for when accounting for algorithms: A systematic literature review on algorithmic accountability. In: Hildebrandt M, Castillo C, editors. *FAT*20: Proceedings of the Conference on Fairness, Accountability, and Transparency.* ACM; 2020;1–18. doi: 10.1145/3351095.3372833.
- [78] Wei K, Ezell C, Gabrieli N, Deshpande C. How do AI companies “fine tune” policy? Examining regulatory capture in AI governance. *arXiv.* 2024. <https://arxiv.org/abs/2410.13042>.
- [79] Yeung K. Algorithmic Regulation: A Critical Interrogation. *Reg Gov.* 2018; 12(4):505–23. doi: 10.1111/rego.12158.