TAKING STANDARDS SERIOUSLY: THE CASE FOR A PRIVATE STANDARDS-BASED APPROACH TO AI GOVERNANCE

Alexander R. Mueller[†] *and Christopher S. Yoo*[††]

ABSTRACT

The rapid proliferation of artificial intelligence (AI) across industry verticals and other domains of social and economic life is forcing policymakers to confront an urgent challenge: how to govern AI systems effectively without stifling innovation or beneficial deployment. This Article argues that private standards-based governance—the adoption of voluntary consensus standards developed through open, multistakeholder processes—offers the most promising "second-best" approach in a world where practical constraints make ideal governance unattainable. Drawing from the successes of standards in governing past digital technologies, the Article highlights how private standards can serve as a regulatory modality for AI by embedding expectations and constraints directly into technological design and organizational practice. Through comparative institutional analysis, it demonstrates how standards-based governance outperforms traditional regulation across four key dimensions: governance architecture, technical expertise and inclusive participation, adaptability to rapid change, and global scalability. At the same time, the Article confronts a number of tradeoffs and challenges such as the nonbinding nature of standards and their susceptibility to industry capture. These challenges and tradeoffs, however, can be managed more efficiently through thoughtful institutional design and strategic government support. The Article ultimately contends that, although standards are not a cure-all, they represent a vital opportunity to build a sustainable, responsive, and highly effective AI governance ecosystem—provided that stakeholders approach their development with the urgency and intentionality the moment demands.

TABLE OF CONTENTS

I. INTRODUCTION 1274
II. AI STANDARDS: A BRIEF OVERVIEW 1278

DOI: https://doi.org/10.15779/Z38WS8HP0D

† Research Fellow, Center for Technology, Innovation & Competition, University of Pennsylvania Carey Law School.

†† Imasogie Professor in Law & Technology, Communications, and Computer & Information Science and Founding Director of the Center for Technology, Innovation & Competition, University of Pennsylvania. This work benefited from discussions following its initial presentation at the 28th Annual BTLJ-BCLT Spring Symposium on "AI Governance at the Crossroads" and was supported by the National Research Foundation of Korea (2022R1A5A7083908).

III. PRIVATE STANDARDS-BASED GOVERNANCE: PROMISES AND SURMOUNTABLE CHALLENGES 1284
A. GOVERNANCE ARCHITECTURE 1285
1. The Benefits of Bottom-Up, Decentralized Governance 1285
2. Second-Best Considerations: Voluntariness and Possible Races to the Bottom 1291
B. STAKEHOLDER INVOLVEMENT 1295
1. Leveraging Field-Level Expertise and Inclusive Participation 1295
2. Second-Best Considerations: Industry Capture and Participation Barriers 1297
C. AGILITY 1302
1. Greater Speed and Adaptability 1302
2. Second-Best Considerations: Prematurity and Tension with Legitimacy 1305
D. SCALE 1307
1. Better Positioning to Scale Across Borders 1307
2. Second-Best Considerations: Geopolitical Competition 1309
IV. CONCLUSION: THE PATH FORWARD 1311

I. INTRODUCTION

As artificial intelligence (AI) systems grow increasingly capable and widespread, the question of how to govern them effectively has become one of the most vexing contemporary policy challenges. The stakes could hardly be higher: AI promises transformative benefits across healthcare, scientific discovery, and many other domains of human endeavor,[1] yet it also introduces

1. *See, e.g.*, W. Nicholson Price II, *Artificial Intelligence in the Medical System: Four Roles for Potential Transformation*, 21 YALE J.L. & TECH. (SPECIAL ISSUE) 122 (2019) (identifying four ways AI may transform medical care); Hanchen Wang, Tianfan Fu, Yuanqi Du, Wenhao Gao, Kexin Huang, Ziming Liu, Payal Chandak, Shengchao Liu, Peter Van Katwyk, Andreea Deac, Anima Anandkumar, Karianne Bergen, Carla P. Gomes, Shirley Ho, Pushmeet Kohli, Joan Lasenby, Jure Leskovec, Tie-Yan Liu, Arjun Manrai, Debora Marks, Bharath Ramsundar, Le Song, Jimeng Sun, Jian Tang, Petar Veličković, Max Welling, Linfeng Zhang, Connor W. Coley, Yoshua Bengio & Marinka Zitnik, *Scientific Discovery in the Age of Artificial Intelligence*, 620 NATURE 47 (2023) (examining the various ways AI is increasingly being used to augment and accelerate scientific research); Francesco Filippucci, Peter Gal, Cecilia Jona-Lasinio, Alvaro Leandro & Giuseppe Nicoletti, *The Impact of Artificial Intelligence on Productivity, Distribution and Growth: Key Mechanisms, Initial Evidence and Policy Challenges* 15–37 (OECD Artificial Intelligence Papers No. 015, 2024), https://www.oecd.org/en/publications/the-impact-of-artificial-intelligence-on-productivity-distribution-and-growth_8d900037-en.html (highlighting early evidence of AI's potential for short-term, firm-level productivity gains while acknowledging that long-term macro-level gains will depend on numerous conditions being realized); NAT'L SCI. & TECH. COUNCIL & U.S. DEP'T OF TRANSP., ENSURING AMERICAN LEADERSHIP IN AUTOMATED VEHICLE TECHNOLOGIES: AUTOMATED VEHICLES 4.0, at 2–3 (2020)

a host of new individual and societal risks.[2] In this high-stakes environment, policymakers face difficult institutional choices about how to structure AI governance so as to enable the benefits of continued AI innovation and deployment while also mitigating the risks of harm.

Standards offer a potential answer to this governance question. From the internet to mobile wireless networks to cybersecurity and encryption, standards have emerged over the past few decades as the dominant modality for governing digital technologies.[3] These mutually agreed-upon specifications defining how a technology should function, perform, and be designed have become key instruments for building order within their respective technological domains. Consistent with a body of literature underscoring how design choices can regulate digital technologies, standards can embed constraints directly into a technology's architecture, shaping how it is used and operates in the world.[4] Not all aspects of standards are technical, however. They can also be directed at organizational practices surrounding a technology's development and use, shaping how firms interact with and oversee the systems they deploy.

Building upon their established role in governing prior digital technologies, standards are increasingly being explored as mechanisms for governing AI.[5] Together, AI standards of both the technical and nontechnical variety have the capacity to shape and constrain how AI systems are developed, tested, deployed, and managed over their lifecycle, steering them towards socially desirable ends. Perhaps more importantly, they can do so in a manner that

(highlighting a variety of safety, economic, and social benefits presented by AI-powered autonomous vehicles); Carlos Mureithi, *High Tech, High Yields? The Kenyan Farmers Deploying AI to Increase Productivity*, GUARDIAN (Sep. 30, 2024), https://www.theguardian.com/world/2024/sep/30/high-tech-high-yields-the-kenyan-farmers-deploying-ai-to-increase-productivity (discussing how small-scale farmers in parts of sub-Saharan Africa are utilizing AI tools to enhance crop yields).

2. *See generally* MIA HOFFMANN & HEATHER FRASE, ADDING STRUCTURE TO AI HARM: AN INTRODUCTION TO CSET'S AI HARM FRAMEWORK 12–13 (2023), https://cset.georgetown.edu/publication/adding-structure-to-ai-harm/ (establishing a taxonomy of potential tangible and intangible AI-related harms).

3. For a concise yet helpful overview of standards and their role in governing digital technologies, see *Digital Standards*, GENEVA INTERNET PLATFORM DIGITAL WATCH OBSERVATORY, https://wp.dig.watch/topics/digital-standards (last visited July 8, 2025).

4. *See generally, e.g.*, Lawrence Lessig, CODE AND OTHER LAWS OF CYBERSPACE 6 (1999); Joel R. Reidenberg, *Lex Informatica: The Formulation of Information Policy Rules Through Technology*, 76 TEX. L. REV. 553 (1998); Daniel Benoliel, *Technological Standards, Inc.: Rethinking Cyberspace Regulatory Epistemology*, 92 CALIF. L. REV. 1069 (2004); Deirdre K. Mulligan & Kenneth A. Bamberger, *Saving Governance-by-Design*, 106 CALIF. L. REV. 697 (2018).

5. *See infra* Part II (identifying several ongoing efforts to develop AI standards across various fora).

offers several key advantages.[6] As this Article explores in greater detail below, the fact that standards are typically developed through collaborative, multistakeholder processes makes them well-positioned to draw directly on the expertise of those closest to the technology.[7] This proximity to the developments on the ground also allows for rapid responses to changes in the technological landscape.[8] Moreover, the voluntary, market-driven nature of standards adoption can foster parallel experimentation and competition among prospective standards, enabling the ecosystem to learn from and choose among the most effective practices.[9] Standards can also easily scale across borders, serving as a form of transnational coordination in an increasingly global AI landscape.[10] For these reasons and others, this Article holds that a private standards-based regime currently represents a highly promising approach to AI governance.

To be clear, the argument here is not that standards represent the *perfect* or *optimal* solution to the AI governance question. In the case of AI, a hypothetical "first-best" solution would be democratically accountable, expertly informed, and highly effective, striking a perfect balance between upholding public values and mitigating AI risks on one hand, and fostering innovation and permitting beneficial AI uses on the other. It would articulate expectations for responsible AI development and deployment in terms specific enough to provide clear guidance to AI actors and appropriately tailored to the unique contextual demands of different types of AI systems or use cases. It would be capable of responding swiftly to changing conditions as AI technology continues to evolve, doing so without the need to resort to advance speculation. Finally, it would be able to meaningfully monitor and enforce compliance at relatively low costs and without demanding significant expansions in capacity.

Although standards largely perform well in these dimensions, they do not do so perfectly. Then again, no institutional arrangement could. The reality is that many of the desirable governance attributes above are in tension with one another. For example, the mechanisms for ensuring democratic accountability often come at the expense of speed and responsiveness. Similarly, technically

6. Many of these advantageous features have been discussed at length by scholars examining the "New Governance" model, an alternative regulatory paradigm under which private standards-setting can be located. *See generally, e.g.*, Orly Lobel, *The Renew Deal: The Fall of Regulation and the Rise of Governance in Contemporary Legal Thought*, 89 MINN. L. REV. 342 (2004); Kenneth W. Abbott & Duncan Snidal, *Strengthening International Regulation Through Transnational New Governance: Overcoming the Orchestration Deficit*, 42 VAND. J. TRANSNAT'L L. 501 (2009).

7. *See infra* Section III.B.

8. *See infra* Section III.C.

9. *See infra* Section III.A.

10. *See infra* Section III.D.

detailed and context-sensitive specifications are valuable insofar as they provide clear, actionable guidance to AI developers and users, yet they are also more prone to obsolescence in the face of rapid technological advancements and are more difficult to centrally maintain. Those with the deepest technical expertise frequently have interests or incentives that diverge from the broader public good, while those best equipped to represent societal values tend to lack the technical fluency needed to develop concrete, effective governance measures. Economist Harold Demsetz's pathbreaking work on the "nirvana" fallacy cautions against dismissing a particular governance arrangement because it falls short of some unattainable ideal.[11] Standards should not be rejected as a governance modality simply because they fail to measure up against a benchmark that, while normatively attractive, is not feasible in practice.

Instead, the more appropriate inquiry adopts a comparative second-best approach that compares private standards-based governance to institutional alternatives as they exist in the real world. A complete comparative second-best analysis would consider the full combinatorial suite of possible institutional arrangements and configurations. However, given the abbreviated format of this Article, it mostly narrows its focus to comparing private standards-based governance with the alternative that seems to be at the forefront of the AI governance debate: traditional government regulation. Such regulation is exemplified by several newly adopted state-level AI statutes[12] and by the European Union's Artificial Intelligence Act, the latter of which is being held up as a model for other countries to follow.[13] Yet, even within this limited frame, the relative advantages of private standards-based governance come into clear view. This does not mean, of course, that private standards-based governance is without challenges or tradeoffs; those certainly exist, whether it be the nonbinding nature of standards, potential democratic deficit in their development, or susceptibility to industry capture. However, there are concrete actions that can be taken—drawing lessons from the governance of other digital technologies on how to design legitimate, effective

11. Harold Demsetz, *Information and Efficiency: Another Viewpoint*, 12 J.L. & ECON. 1, 1–4 (1969).

12. *See generally, e.g.*, Act Concerning Consumer Protections in Interactions with Artificial Intelligence Systems, ch. 198, 2024 Colo. Sess. Laws 1199 (codified at COLO. REV. STAT. § 6-1-1701); Artificial Intelligence Policy Act, ch. 186, 2024 Utah Laws (codified at UTAH CODE §§ 13-2-12, 13-70-101 to -305, 76-2-107 and as amended at §§ 13-11-4, 13-61-101, 63I-2-213).

13. *See generally* 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 Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), 2024 O.J. (L 1689) 1.

multistakeholder standards development processes—to make these tradeoffs more favorable such that the benefits of private standards-based governance far outweigh the costs.[14]

The remainder of this Article proceeds as follows. Part II provides a brief overview of AI standards, their functions, and current standardization efforts underway. Part III discusses the comparative strengths of private standards-based governance across four critical dimensions—governance architecture, stakeholder involvement, agility, and scale—while also considering the tradeoffs involved and how they might be better managed. Part IV concludes by offering some final thoughts about the best way forward for AI governance.

II. AI STANDARDS: A BRIEF OVERVIEW

Standards, in the most basic terms, are commonly agreed-upon ways of doing something, whether that be designing a product, organizing information, or performing an activity. Standards serve the critical functions of facilitating interoperability, compatibility, consistency, and quality across different products, systems, and organizations. By defining criteria that enable the formation of shared expectations, standards help reduce uncertainty and transaction costs between parties as well as facilitate consumer trust. They often take the form of *technical standards*: documents specifying common interface designs enabling components to interconnect,[15] formats for structuring information to enable shared understandings,[16] rules for how a technology should behave or interact under different scenarios,[17] or uniform benchmarks and procedures for testing reliability and safety.[18] They can also

14. The term "multistakeholder" lacks a universally accepted definition, but it generally entails two or more classes of actors with shared authority in a common governance enterprise addressing an issue of traditional public concern. *See* Mark Raymond & Laura DeNardis, *Multistakeholderism: Anatomy of an Inchoate Global Institution*, 7 INT'L THEORY 572, 573–82 (2015).

15. *See, e.g.*, *Universal Serial Bus Type-C Cable and Connector Specification Release 2.2*, USB IMPLEMENTERS F., INC. (Oct. 18, 2022), https://usb.org/document-library/usb-type-cr-cable-and-connector-specification-release-22 [https://web.archive.org/web/20221120045921/https://usb.org/document-library/usb-type-cr-cable-and-connector-specification-release-22].

16. *See, e.g.*, *HTML Living Standard*, WEB HYPERTEXT APPLICATION TECH. WORKING GRP., https://html.spec.whatwg.org/multipage/ (last visited Apr. 16, 2025); *Extensible Markup Language (XML) 1.0 (5th ed.)*, WORLD WIDE WEB CONSORTIUM (Nov. 26, 2008), https://www.w3.org/TR/xml/.

17. *See generally, e.g.*, Mike Belshe, Roberto Peon & Martin Thomson, *Hypertext Transfer Protocol Version 2 (HTTP/2)*, INTERNET ENG'G TASK FORCE (May 2015), https://datatracker.ietf.org/doc/html/rfc7540.

18. *See generally, e.g.*, INST. OF ELEC. & ELEC. ENG'RS, *IEEE 3130-2024: Standard for Security Requirements and Testing Methods of Operating Systems in Connected Vehicles* (2024), https://standards.ieee.org/ieee/3130/h10757/.

take the form of *governance* or *management standards*, which are standards that specify organizational structures, processes, and practices for managing internal assets, capabilities, and risk as well as for meeting legal or ethical obligations.[19] Standards, whether they are of the technical or management variety, also vary in their intended scope of applicability: some are general and broadly applicable across domains, whereas others are more domain-specific.

Standards typically emerge through deliberative, consensus-based processes at institutions known as Standards Development Organizations (SDOs). To be sure, a particular technology or product design created unilaterally by a single firm can achieve the status of *de facto* standard simply by becoming widely accepted in the marketplace. An example relevant to the AI context would be NVIDIA's CUDA, a GPU-accelerated computing platform that, at one point, was virtually indispensable (i.e., "the industry standard") for advanced model training and inference.[20] Still, most standards are developed cooperatively in an institutional setting, though the institutions themselves can differ in many important respects. SDOs vary in terms of their composition, with some taking the form of industry consortia made up of mostly corporate stakeholders, while many others are multistakeholder bodies that engage civil society, academia, and technical community in addition to industry. SDOs also vary in terms of geographic scope, ranging from nationally accredited bodies with geographically limited activities to more globally focused bodies.[21] Furthermore, each has its own unique rules and processes governing participation, agenda-setting, voting, and intellectual property rights, all of which can significantly influence the content and adoption of resulting

19. *See generally, e.g.*, INT'L ORG. FOR STANDARDIZATION, *ISO/IEC 27001: Information Security, Cybersecurity and Privacy Protection—Information Security Management Systems—Requirements* (2022), https://www.iso.org/standard/82875.html.

20. *See* Hasan Chowdhury, *CUDA Is Nvidia's Secret Sauce—and Now It's in the Sights of European Regulators*, BUS. INSIDER (July 2, 2024), https://www.businessinsider.com/nvidia-secret-sauce-regulators-gpu-chips-jensen-huang-2024-7. *But see* Anton Shilov, *DeepSeek's AI Breakthrough Bypasses Industry-Standard CUDA for Some Functions, Uses Nvidia's Assembly-Like PTX Programming Instead*, TOM'S HARDWARE (Jan. 28, 2025), https://www.tomshardware.com/tech-industry/artificial-intelligence/deepseeks-ai-breakthrough-bypasses-industry-standard-cuda-uses-assembly-like-ptx-programming-instead (reporting a breakthrough from Chinese firm DeepSeek that has cast doubt on CUDA's continued indispensability).

21. It is not uncommon for an SDO to have either a loosely defined geographic scope or none at all. For example, the Internet Engineering Task Force (IETF), the organization that has historically been responsible for standardizing core and application-layer internet protocols, has neither a physical headquarters nor a formal membership requirement; standards work occurs mainly through electronic mailing list discussions or tri-annual in-person meetings that rotate between host cities, both of which are open to anyone who wishes to participate. Its "geographic scope" is thus a function of where its participants happen to be located, something that has become increasingly diverse over time.

standards.[22] However, what all of these SDOs have in common is that, unlike traditional laws and regulations, the standards they produce are voluntary, meaning adoption is primarily driven by market forces rather than state mandates.

Before proceeding any further, there are a few points worth clarifying. First, standards are distinct from the numerous AI ethical frameworks and high-level principles that have sprung up over the past decade. Much of the supra- and non-national cooperation around AI governance to date has indeed consisted of articulating broad principles—accountability, transparency, fairness, robustness, etc.—that should guide the development and use of AI.[23] While these efforts have been well-intentioned and constitute a good first step towards laying a shared normative foundation, abstract principles alone say little about how AI systems should be developed or used in practice, limiting their effectiveness as AI governance mechanisms.[24] Other commentators have struck a more critical tone, accusing these principles of enabling a form of ethics washing: a rhetorical tool for convincing regulators and the public that concerns are being addressed, yet vague and unenforceable enough so as to require industry to make few meaningful commitments or changes.[25] Regardless of one's perspective on these previous efforts, it is important to recognize that a standards-based regime is not just a continuation of the

22. For a comprehensive comparison of the various rules governing major technical SDOs, see generally Justus Baron, Jorge Contreras, Martin Husovec, Pierre Larouche & Nikolaus Thumm, *Making the Rules: The Governance of Standard Development Organizations and Their Policies on Intellectual Property Rights*, JOINT RSCH. CTR. SCI., EUR. COMM'N EUR 29655 EN (Nikolaus Thumm ed., 2019).

23. *See, e.g.*, Organisation for Economic Co-operation and Development [OECD], *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449 (adopted May 22, 2019; amended May 3, 2024), https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449; United Nations Educational, Scientific and Cultural Organization [UNESCO], *Recommendation on the Ethics of Artificial Intelligence* (adopted Nov. 23, 2021), https://unesdoc.unesco.org/ark:/ 48223/pf0000381137; G20, *G20 Ministerial Statement on Trade and Digital Economy*, annex (June 9, 2019), https://www.mofa.go.jp/policy/economy/g20_summit/osaka19/pdf/documents/en/annex_08.pdf.

24. *See* Brent Mittelstadt, *Principles Alone Cannot Guarantee Ethical AI*, 1 NATURE MACH. INTEL. 501, 503–04 (2019); Jess Whittlestone, Rune Nyrup, Anna Alexandrova & Stephen Cave, *The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions*, 2019 AAAI/ACM CONF. ON AI, ETHICS, & SOC'Y 195, 196–97 (2019), https://doi.org/10.1145/3306618.3314289.

25. *See, e.g.*, Ryan Calo, *Artificial Intelligence and the Carousel of Soft Law*, 2 IEEE TRANSACTIONS ON TECH. & SOC'Y 171, 172 (2021), https://doi.org/10.1109/TTS.2021.3136025; Luke Munn, *The Uselessness of AI Ethics*, 3 AI & ETHICS 869, 872 (2023), https://doi.org/10.1007/s43681-022-00209-w; Karen Hao, *In 2020, Let's Stop AI Ethics-Washing and Actually Do Something*, MIT TECH. REV. (Dec. 27, 2019), https://www.technologyreview.com/2019/12/27/57/ai-ethics-washing-time-to-act/.

"abstract principle" approach to governance. Rather, the point of AI standards is to operationalize these high-level principles, translating them into concrete, technically implementable measures that can be objectively assessed and verified.[26]

This understanding of standards as concrete and well-specified may be counterintuitive to many lawyers, as it represents the inverse of the way the term "standard" has long been employed within the legal tradition.[27] Here, standards have historically been contrasted with rules—the former being open-ended (e.g., reasonableness in tort law, materiality in securities law, or good faith in contract law) and applied ex post and the latter being highly precise and specific and determined ex ante, leaving much less discretion to the party applying the rule.[28] Despite this terminological confusion, the type of technical and governance standards discussed in this Article are, in terms of substance, much more akin to legal rules: they aim to provide specific, actionable ex ante guidance rather than broad discretionary principles.

With that understanding in mind, there are several places where AI standardization activities could beneficially focus.[29] One is the development of different procedures and benchmarks for evaluating model robustness, security, and/or bias—both before deployment and throughout the model's lifecycle as conditions change.[30] These types of standards will almost surely be

26. Some may view this project with skepticism, arguing that framing complex ethical challenges as matters of technical design and implementation reinforces a kind of "technological solutionism" that obscures and sidesteps deeper normative questions. *See, e.g.*, Mittelstadt, *supra* note 24, at 505. Though these concerns are not without merit, nothing about translating principles into implementable practices requires ignoring or oversimplifying the normative questions involved, treating them as purely technical matters that should be addressed exclusively by those with relevant technical expertise. To the contrary, well-designed standards processes can provide structured and inclusive forums for stakeholders to negotiate how values like fairness, accountability, or transparency should be understood and expressed in different technical and organizational contexts.

27. Cass R. Sunstein, *Problems with Rules*, 83 CALIF. L. REV. 953, 959 (1995) ("Lawyers have customarily compared standards ('do not drive unreasonably fast') to rules ('do not go over 60 miles per hour'), with rules seeming hard and fast, and standards seeming open-ended.").

28. *See id.*; Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 DUKE L.J. 557, 559–60 (1992).

29. Many of the areas highlighted here overlap with those outlined in a recent NIST report on the future of U.S. standards engagement. *See generally* JESSE DUNIETZ, ELHAM TABASSI, MARK LATONERO & KAMIE ROBERTS, NAT'L INST. OF STANDARDS & TECH., NIST TRUSTWORTHY AND RESPONSIBLE AI 100-5: A PLAN FOR GLOBAL ENGAGEMENT ON AI STANDARDS (2024), https://www.nist.gov/publications/plan-global-engagement-ai-standards [https://doi.org/10.6028/NIST.AI.100-5] [hereinafter NIST AI Standards Plan].

30. *See id.* at 10–11; *see also* NAT'L TELECOMMS. & INFO. ADMIN., U.S. DEP'T OF COM., NTIA ARTIFICIAL INTELLIGENCE ACCOUNTABILITY POLICY REPORT 26–27 (2024), https://

highly application- or domain-specific. Automated driving systems, for instance, demand much different testing and validation protocols than AI-based medical diagnostic systems.

Another place where standardization efforts are likely to be directed is the definition of standardized formats for disclosing important information about a model, including that about the underlying architecture, training data used, benchmarking/testing results, and any known limitations.[31] Such disclosures help enable downstream actors to accurately assess the model's suitability for specific applications and implement complementary safeguards to enhance overall system safety and reliability.

Similarly, a standardized protocol for reporting discovered model flaws back to the original developer and any affected downstream actors would help facilitate the remediation of potential ecosystem-wide risks in a responsible and streamlined manner.[32] There are also places where AI standards may not be imminent but may emerge in the future as the technology matures. For example, as advancements continue to be made in areas such as model interpretability and explainability, related techniques and technical mechanisms could become promising candidates for standardization activities.[33]

Just as technical standards can help ensure AI systems function reliably and safely, management and governance standards can help address the human and organizational aspects of AI development, use, and oversight. These standards might include—among many other things—implementation best

---

www.ntia.gov/sites/default/files/ntia-ai-report-final.pdf (identifying appropriate information flows to downstream actors as "a critical input to AI accountability").

31. Model cards, including those published by organizations like OpenAI, offer a promising starting point for model disclosures. *See, e.g.*, OpenAI, *GPT-4o System Card* (Aug. 8, 2024), https://openai.com/index/gpt-4o-system-card/. However, such disclosures are not standardized across organizations, vary widely in scope and granularity, and often lack sufficient detail to enable meaningful third-party evaluation or implementation of downstream safeguards.

32. *See* Shayne Longpre, Kevin Klyman, Ruth E. Appel, Sayash Kapoor, Rishi Bommasani, Michelle Sahar, Sean McGregor, Avijit Ghosh, Borhane Blili-Hamelin, Nathan Butters, Alondra Nelson, Amit Elazari, Andrew Sellars, Casey John Ellis, Dane Sherrets, Dawn Song, Harley Geiger, Ilona Cohen, Lauren McIlvenny, Madhulika Srikumar, Mark M. Jaycox, Markus Anderljung, Nadine Farid Johnson, Nicholas Carlini, Nicolas Miailhe, Nik Marda, Peter Henderson, Rebecca S. Portnoff, Rebecca Weiss, Victoria Westerhoff, Yacine Jernite, Rumman Chowdhury, Percy Liang & Arvind Narayanan, *In-House Evaluation Is Not Enough: Towards Robust Third-Party Flaw Disclosure for General-Purpose AI* (Mar. 25, 2025), https://arxiv.org/abs/2503.16861 (identifying the need for a standardized AI flaw report template and coordinated disclosure process).

33. *See* NIST AI Standards Plan, *supra* note 29, at 13 (identifying explainability and interpretability as areas where standards are needed, but that still require "significant foundational work").

practices, internal accountability and governance structures, frameworks in areas such as data quality management or secure development, guidelines for responsible AI procurement, methodologies for conducting impact assessments, and incident response protocols for AI system failures. Sector-specific governance standards may further tailor these approaches to ensure AI deployment aligns with the unique legal, ethical, and operational expectations of different industry verticals.

Many stakeholders have already recognized the need for AI standards, as evidenced by the numerous standardization efforts underway across various organizations and governance levels.[34] A joint technical committee (JTC) of the International Organization for Standardization (ISO) and the International Electrotechnical Committee (IEC), bodies comprised of 165 member countries each represented by a designated national standards organization, launched the earliest coordinated effort to develop global AI standards.[35] Since 2017, ISO/IEC JTC1 has developed and issued dozens of AI standards, though most of its focus thus far has been on the governance and management side as well as on defining foundational concepts and terminology.[36] The Institute of Electrical and Electronics Engineers Standards Association (IEEE-SA), a global multistakeholder body, has also initiated work on a mix of different terminological, technical, and governance standards.[37] At the regional and national levels, U.S. National Institute of Standards and Technology (NIST) has continued to build around its AI Risk Management Framework (RMF) in addition to launching various technical evaluation and testing initiatives.[38] Similarly, the European Committee for Standardization

34. The UK's Alan Turing Institute hosts an extensive and continuously updated online database that compiles both published and in-progress AI standards from various organizations worldwide. *See* ALAN TURING INST., AI STANDARDS HUB, https://aistandardshub.org/ (last visited Apr. 16, 2025).

35. *See* Wael William Diab & Mike Mullane, *How the ISO and IEC Are Developing International Standards for the Responsible Adoption of AI*, UNESCO (Aug. 2, 2024), https://www.unesco.org/en/articles/how-iso-and-iec-are-developing-international-standards-responsible-adoption-ai.

36. *See* NIST AI Standards Plan, *supra* note 29, at 23–27.

37. *See, e.g.*, INST. OF ELEC. & ELEC. ENG'RS, *IEEE 7000-2021: Standard Model Process for Addressing Ethical Concerns During System Design* (2021), https://standards.ieee.org/ieee/7000/6781/; INST. OF ELEC. & ELEC. ENG'RS, *IEEE 3119-2025: Approved Draft Standard for the Procurement of Artificial Intelligence and Automated Decision Systems* (approved Mar. 27, 2025), https://standards.ieee.org/ieee/3119/10729/; INST. OF ELEC. & ELEC. ENG'RS, *IEEE 3129-2023: Standard for Robustness Testing and Evaluation of Artificial Intelligence (AI)-Based Image Recognition Service* (2023), https://standards.ieee.org/ieee/3129/10747/.

38. *See* NAT'L INST. OF STANDARDS & TECH., NIST AI 100-1: ARTIFICIAL INTELLIGENCE RISK MANAGEMENT FRAMEWORK (2024), https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf; *see also, e.g.*, NAT'L INST. OF STANDARDS & TECH., AI RISK

and the European Committee for Electrotechnical Standardization (CEN/CENELEC) have been charged with developing "harmonized standards" to support that implementation of the European Union's landmark AI Act.[39]

Beyond these legacy SDOs, several new fora have emerged to address specific aspects of AI governance. The Partnership on AI is a coalition of technology companies and civil society groups focusing on actionable guidance around the ethical development and use of AI.[40] MLCommons is another multistakeholder consortium that develops performance, quality, and safety benchmarks for AI systems.[41] The Frontier Model Forum is a small alliance of leading frontier AI labs dedicated to identifying best practices for the safe and secure development of large advanced AI models.[42] The Coalition for Content Provenance and Authenticity (C2PA) is an industry-led consortium that creates technical standards for certifying the source of media content amid the rise of deepfakes.[43] All of these initiatives reflect a growing recognition that effective AI governance requires more than just high-level principles, but also concrete steps and measures for putting those principles in practice. As AI capabilities continue to advance and AI systems become more deeply embedded into economic and social life, the role of standards will only grow in importance.

III. PRIVATE STANDARDS-BASED GOVERNANCE: PROMISES AND SURMOUNTABLE CHALLENGES

As AI systems become increasingly pervasive, more granular articulations of expected or desired behavior surrounding AI will become a virtual necessity. There are several forces that may contribute to this need. It could be the demand for common baselines against which ex ante audits or tests can be conducted, a way to foster trust throughout the ecosystem by signaling

MANAGEMENT FRAMEWORK PLAYBOOK (2023), https://airc.nist.gov/docs/AI_RMF_Playbook.pdf.

39. *See* Hadrien Pouget & Ranj Zuhdi, *AI and Product Safety Standards Under the EU AI Act*, CARNEGIE ENDOWMENT FOR INT'L PEACE (Mar. 5, 2024), https://carnegieendowment.org/research/2024/03/ai-and-product-safety-standards-under-the-eu-ai-act?lang=en.

40. *About Us*, P'SHIP ON AI, https://partnershiponai.org/about (last visited Apr. 16, 2025).

41. *About Us: An Open AI Engineering Consortium*, MLCOMMONS, https://mlcommons.org/about-us/ (last visited Apr. 16, 2025).

42. *About Us*, FRONTIER MODEL F., https://www.frontiermodelforum.org/about-us/#mission (last visited Apr. 16, 2025).

43. *Coalition for Content Provenance and Authenticity (C2PA) Charter*, COAL. FOR CONTENT PROVENANCE & AUTHENTICITY, https://c2pa.org/about/charter/ (last visited Apr. 16, 2025).

information about a model's quality or the adequacy of an organization's compliance efforts. Similarly, it could be the demand for clear frameworks that help guide ex post determinations of responsibility when AI-enabled harms occur. Perhaps detailed specifications are sought to facilitate vertical coordination and enable firms at different levels of the supply chain to jointly provide AI-based products or services in a safe and reliable fashion. It could even be as simple as the need to codify knowledge about best practices for AI risk mitigation so that firms have actionable guidance as they navigate an uncertain technological frontier.

Regardless of what the main driver is, standards-like specifications for AI—whether promulgated by a legislative or regulatory body or inferred ex post from a body of decisions assessing liability—are an inevitability. One of the few questions that remains is who should determine their content. The question of "who decides," as legal scholar Neil Komesar explains in his book *Imperfect Alternatives*, is fundamentally a question of institutional choice.[44] The answer to this question may ultimately dictate the degree to which social and policy goals around AI are realized.[45]

Among the institutional alternatives available for governing AI, the most promising—particularly in its ability to balance AI harm prevention with continued innovation and deployment—is private standards-based governance. To reiterate an important point raised in the introduction, this Article does not argue that private standards-based governance represents the *optimal* or *ideal* solution to the AI governance puzzle. It is extremely unlikely that such a "first-best" solution exists. In what follows, this Article examines the relative strengths of standards across the four critical dimensions mentioned briefly above—governance architecture, stakeholder involvement, agility, and scale—while also considering the tradeoffs that exist and how they might be better managed. When compared to other institutional arrangements, each viewed in light of the real-world limitations and the tradeoffs they entail, private standards-based governance emerges as the most desirable of the second-best options.

A. GOVERNANCE ARCHITECTURE

1. The Benefits of Bottom-Up, Decentralized Governance

Classic command-and-control regulation seeks to achieve outcomes by issuing legal commands—typically in the form of prescribed conduct or technical configurations—that leave regulated entities with little discretion

44. NEIL K. KOMESAR, IMPERFECT ALTERNATIVES: CHOOSING INSTITUTIONS IN LAW, ECONOMICS, AND PUBLIC POLICY 3 (1994).

45. *See id.* ch. 2.

over implementation.[46] While this can provide regulated entities with greater certainty from a compliance standpoint, the requirements issued by regulators may be suboptimal or even ineffective.[47] This is particularly cause for concern given governments' less-than-stellar track record in setting or selecting appropriate technical standards in high-technology domains.[48] Flaws in government-mandated AI specifications can be costly, undermining both AI innovation and the effectiveness of governance efforts, as the binding nature of these specifications compels adoption regardless of any shortcomings.[49]

By contrast, the voluntary market-driven nature of private standards adoption allows for more meritocratic selection of the best approaches. When multiple competing standards emerge—and they frequently do—decisions about which standards prevail are made through the decentralized choices of developers, users, and other ecosystem participants.[50] To gain traction in the marketplace in the form of adoption, private standards must prove their value

46. *See* Kenneth A. Bamberger, *Regulation as Delegation: Private Firms, Decision Making, and Accountability in the Administrative State*, 56 DUKE L.J. 377, 386 (2006) [hereinafter Bamberger, *Regulation as Delegation*]; Lobel, *supra* note 6, at 376 ("Their agency is limited to choosing whether to comply with the regulations to which they are subjected.").

47. *See* Daniel Gervais, *The Regulation of Inchoate Technologies*, 47 HOU. L. REV. 665, 704 (2010).

48. For surveys of past efforts showing the shortcomings of government-set standards, see STANLEY M. BESEN & LELAND L. JOHNSON, COMPATIBILITY STANDARDS, COMPETITION, AND INNOVATION IN THE BROADCASTING INDUSTRY 135 (1986); JEFFREY H. ROHLFS, BANDWAGON EFFECTS IN HIGH-TECHNOLOGY INDUSTRIES 201 (2001). One of the most notorious examples was the U.S. government's initiative to standardize a hardware-based key escrow system—better known as the "Clipper Chip"—that would permit law enforcement to obtain "exceptional access" to encrypted communication. The system saw little adoption and was abandoned just a few years later, due in no small part to a widely cited paper revealing significant technical vulnerabilities. *See generally* Matt Blaze, *Protocol Failure in the Escrowed Encryption Standard*, 2 PROC. ASS'N FOR COMPUTING CONF. ON COMPUT. & COMMC'NS SEC. 59 (1994) (finding major security flaws in the government's Escrowed Encryption Standard); *see also* Stacy Baird, *The Government at the Standards Bazaar*, 18 STAN. L. & POL'Y REV. 35, 67–70 (2007) (examining several other instances of "government failure" in technical standards setting); STEPHEN BREYER, REGULATION AND ITS REFORM 115 (1982) (identifying several challenges that traditional regulators face in attempting to set standards).

49. BREYER, *supra* note 48, at 102.

50. The emphasis placed on parallel experimentation and competition between prospective standards has often been described as a uniquely American approach to standardization. *See* DIETER ERNST, EAST-WEST CTR., POLICY STUDIES NO. 66, AMERICA'S VOLUNTARY STANDARDS SYSTEM: A 'BEST PRACTICE' MODEL FOR ASIAN INNOVATION POLICIES? 32 (Edward Aspinall & Dieter Ernst eds., 2013); TIM BÜTHE & WALTER MATTLI, THE NEW GLOBAL RULERS: THE PRIVATIZATION OF REGULATION IN THE WORLD ECONOMY 162 (2011). However, it is also the model that the digital technology standards-setting landscape has come to mostly closely resemble—perhaps a reflection of the United States' outsized influence in this domain.

to prospective implementers.[51] The choices the market ultimately converges around can be difficult to predict in advance. For example, there was a point in time when many expected Bluetooth would become the dominant wireless technology for connecting user devices to local area networks (among its many other uses). Yet, it was the superior range and transfer speeds provided by the IEEE 802.11 standard—known commercially as Wi-Fi—that ultimately caused it to triumph over the low-cost, low-power consumption Bluetooth.[52]

The same competitive dynamic that exists between rival standards can also extend to the SDO level, as standard-setting venues themselves must vie to attract and retain participants in order to stay relevant.[53] This process resembles Charles Tiebout's famous model under which interjurisdictional competition can lead to the efficient provision of local public goods.[54] Just as residents in the Tiebout model are free to move to communities that match their preferred level of local public good provision,[55] stakeholders in the standardization process have the opportunity to "vote with their feet" in choosing where to participate. The resulting competitive pressures can help ensure that standard-setting remains responsive to needs rather than stagnating under some combination of complacency and institutional inertia.[56]

An example of this phenomenon can be found in the efforts to standardize the World Wide Web during the mid-1990s. After growing frustrated with the slow pace and "endless philosophical rat holes" they had encountered at the Internet Engineering Task Force (IETF), web pioneer Tim Berners-Lee and several other early browser developers decided to defect and form the World Wide Web Consortium (W3C).[57] While certainly not without faults of its own, the W3C has proven instrumental in the Web's rapid and successful

51. *See* NIST AI Standards Plan, *supra* note 29, at 5.

52. Joseph Farrell & Timothy Simcoe, *Four Paths to Compatibility*, *in* THE OXFORD HANDBOOK OF DIGITAL ECONOMY 34, 40 (Martin Peitz & Joel Waldfogel eds., 2012) ("For example, Bluetooth (IEEE 802.15) was conceived as a home networking standard, but ceded that market to Wi-Fi (IEEE 802.11) and is now widely used in short-range low-power devices, such as wireless headsets, keyboards, and remote controls.").

53. *See* Baron et al., *supra* note 22, at 64–65; *see also* Errol Meidinger, *Competitive Supragovernmental Regulation: How Could It Be Democratic?*, 8 CHI. J. INT'L L. 513, 531 (2008) (making this same observation about the transnational private regulatory landscape more generally).

54. *See generally* Charles M. Tiebout, *A Pure Theory of Local Expenditures*, 64 J. POL. ECON. 416 (1956).

55. *Id.* at 418.

56. *Cf.* Baron et al., *supra* note 22, at 67 (suggesting that inter-venue competition can serve as a check on a given SDO's ability to impose unfavorable new policies in areas such as IP rights).

57. TIM BERNERS-LEE, WEAVING THE WEB: THE ORIGINAL DESIGN AND ULTIMATE DESTINY OF THE WORLD WIDE WEB BY ITS INVENTOR 62 (1999).

standardization, demonstrating the benefits of responsive governance when faced with competitive pressure.

A private standards-based regime would also allow room for early-stage and continuous experimentation, permitting several different approaches to develop in parallel so the ecosystem can learn from different models before either codifying them into standards or deciding to adopt them.[58] Such experimentation is especially valuable given the heterogeneity among AI systems and use cases, which almost certainly renders a one-size-fits-all approach inappropriate.[59] Instead of a regulator attempting to determine *a priori* if and where sector-specific approaches are needed, this differentiation can be dictated by the needs and challenges experienced by those operating within each sector.[60] Groups of related actors such as medical AI developers or autonomous vehicle manufacturers can determine whether a horizontal approach is appropriate for their intended use-case and, if not, split off and develop their own specialized standards that better address domain-specific concerns.

Top-down specifications imposed by regulators, on the other hand, stand to foreclose the possibility of private experimentation around potential improvements or alternatives.[61] In order to avoid noncompliance, regulated entities would need to continue adhering to the mandated specifications until the regulator updates them, even if the specifications grew outdated or were poorly suited for a new use case.[62] Such concerns are particularly acute in the

58. *See* Daniel E. Walters & Hannah J. Wiseman, *Self-Regulation in Emerging and Innovative Industries*, 62 HOU. L. REV. 543, 564–65 (2025); *see also* Lobel, *supra* note 6, at 380, 382 (explaining that the diversity and experimentation enabled under decentralized new governance approaches is more than just a temporary process for identifying and ultimately converging on the best solution, but it also serves as a "means for continuous change and improvement").

59. Cary Coglianese, *Regulating Machine Learning: The Challenge of Heterogeneity*, COMPETITION POL'Y INT'L: TECHREG CHRON. 8 (Feb. 2023) [hereinafter Coglianese, *Machine Learning*]; *see also* Lobel, *supra* note 6, at 379–80 (arguing that governance approaches emphasizing diversification and pluralized solutions are better suited for complex, dynamic regulatory domains than uniform, top-down models).

60. Walters & Wiseman, *supra* note 58, at 571.

61. *See* BREYER, *supra* note 48, at 105 (explaining that government-mandated design standards diminish firms' "incentive to look for better methods"); Lobel, *supra* note 6, at 393 (explaining that one reason we might prefer softer alternatives to traditional hard law is that the former, unlike the later, allows for deviance and trial and error without the fear of sanctions).

62. *See* BREYER, *supra* note 48, at 115–16 (highlighting the tendency of government mandated design standards to "freeze technology"). *But see* Cary Coglianese, *The Limits of Performance-Based Regulation*, 50 U. MICH. J. L. REFORM 525, 542 (2017) (recognizing there are actions that regulators can take to reduce technology "lock-in" and accommodate new

context of AI, a technology most would agree is still in its infancy.[63] It is extremely likely that any initial regulatory specifications would need significant refinement over time as both understanding of AI and the technology itself continue to evolve, yet the limited feedback channels within a command-and-control regime make adaptation significantly more difficult.[64]

Concededly, command-and-control regulation is not the only item traditional regulators have in their toolkit. Performance-based regulation, which focuses solely on outcomes rather than the means of achieving those outcomes, may appear to offer a more flexible yet nonetheless state-centric alternative.[65] Here, regulators define a specific performance target or goal that must be met, but leave it up to each regulated entity to determine how to achieve it.[66] The main weakness of the performance-based approach, however, is that not every area of regulation is conducive to measurable outcomes and AI is arguably one of them.[67] Unlike domains such as air pollution where regulators can set specific, quantifiable emissions thresholds, AI systems present complex, multidimensional risks that often resist simple performance metrics.[68]

Less prescriptive forms of state-centric regulation can also create significant uncertainty by threatening to penalize firms without simultaneously offering them concrete guidance or a means of easily verifying compliance. With a cutting-edge technology like AI, where knowledge pertaining to the most advanced systems has an almost esoteric quality, many firms—small and medium-sized enterprises (SMEs) in particular—are poorly equipped to navigate this uncertainty on their own.[69] Even though the technical expertise

innovations, such as establishing waiver mechanisms that allow firms to depart from otherwise binding specifications).

63. *See* Gervais, *supra* note 47, at 671–73, 702 (arguing that traditional regulatory interventions aimed at inchoate technologies—those that "are far from completely developed" and for whom the "future is unpredictable"—are more likely to miss their targets).

64. *See id.* at 676 ("[T]here are rarely feedback loops to adjust the regulatory framework if the target is missed.").

65. Coglianese, *The Limits of Performance-Based Regulation*, *supra* note 62, at 526.

66. *Id.*

67. *See* Bamberger, *Regulation as Delegation*, *supra* note 46, at 389 ("Certain public problems . . . lend themselves to neither specific behavioral commands nor measurable outcomes."); *see also, e.g.*, David Thaw, *The Efficacy of Cybersecurity Regulation*, 30 GA. ST. U. L. REV. 287, 301–02 (2014) (discussing how the lack of well-defined outcome metrics in information security complicates the application of performance-based regulatory approaches).

68. *See* Coglianese, *Regulating Machine Learning*, *supra* note 59, at 8 ("[I]n many cases it will be unlikely that regulators can develop sufficiently clear, monitorable performance tests for algorithms themselves.").

69. A similar phenomenon of disproportionate compliance burden on smaller entities due to regulatory complexity and uneven distribution of expertise has been observed with the

possessed by private actors is superior to that of regulators *in the aggregate*, this expertise is unevenly distributed: larger firms can afford to employ in-house AI specialists or engage outside consultants, but SMEs are much less likely to possess the resources to determine best practices on their own. As a result, regulatory uncertainty may disproportionately burden SMEs, chilling AI adoption and deployment due to the unpredictable legal and compliance risks they would face.[70]

To briefly summarize thus far: highly concrete, implementable specifications are desirable for effective AI governance, but state-centric modes of regulation are poorly positioned to provide this concreteness and even risk stunting the technology's development by mandating suboptimal approaches. A private standards-based regime, meanwhile, would allow approaches to emerge more organically from the bottom-up. By enabling decentralized experimentation and choice, such a regime fosters more effective, adaptive, and innovation-permitting governance.

Finally, under a private standards-based governance regime, the role of government would undergo a shift towards areas where it retains comparative institutional advantages: enforcing commitments, addressing market failures, and policing anticompetitive conduct within the standardization process. This might also include holding firms accountable ex post when they make representations to the public or other firms about their adherence to particular standards yet fail to do so.[71] The shift in function here reflects Judge Frank Easterbrook's influential insight from the early days of the internet: in the face of rapid technological change and uncertainty, the most productive role for public law is *not* to anticipate the future and attempt to engineer bespoke solutions but to instead support private ordering.[72] In Easterbrook's words, "If you don't know what is best, let people make their own arrangements."[73]

EU's General Data Protection Regulation (GDPR). *See, e.g.*, Sean Sirur, Jason R.C. Nurse & Helena Webb, *Are We There Yet? Understanding the Challenges Faced in Complying with the General Data Protection Regulation (GDPR)*, arXiv:1808.07338v1, at 5–6, 8 (Aug. 22, 2018), https://arxiv.org/pdf/1808.07338v1.

70. *See id.*

71. The Federal Trade Commission (FTC) has broad authority to police unfair and deceptive acts under Section 5 of the FTC Act and has already used it to bring enforcement actions against firms accused of misrepresenting the capabilities of their AI systems. *See* 15 U.S.C. § 45(a)(1); *see also* Press Release, Fed. Trade Comm'n, FTC Announces Crackdown on Deceptive AI Claims and Schemes (Sep. 25, 2024), https://www.ftc.gov/news-events/news/press-releases/2024/09/ftc-announces-crackdown-deceptive-ai-claims-schemes.

72. *See* Frank H. Easterbrook, *Cyberspace and the Law of the Horse*, 1996 U. CHI. LEGAL F. 207, 207–13.

73. *Id.* at 210.

Notably, the re-definition of the government's role would dramatically reduce administrative costs.[74] This is particularly salient given the large number of actors across different industry verticals who are already beginning to develop and use AI systems for a variety of different purposes, as any form of centralized oversight and enforcement will be both challenging and demand significant state resources.[75] While administrative costs are neither the only nor an overriding consideration, they assume heightened significance in the current U.S. political climate, which has seen a sharp rise in hostility towards the administrative state under the second Trump administration.[76] Repositioning government more as a facilitator and backstop instead of the primary standard-setter offers a more resource-efficient pathway to effective AI governance.

2. Second-Best Considerations: Voluntariness and Possible Races to the Bottom

The room for experimentation and competition that exists under a private standards-based regime does not come without tradeoffs. Unlike traditional state-centric regulation, where standards can be given the force of law when widespread adoption is desirable, private standards are still entirely voluntary. The manner in which certain types of AI technical standards enable interoperability, enhance product compatibility, and facilitate multi-firm provision provides strong incentives for adoption. However, other types of AI standards will impose compliance costs on adopting firms without offering an immediate, tangible benefit. For example, standards aimed at minimizing algorithmic bias and discrimination might not directly contribute to a company's bottom line, at least in the short term. Even when standards of the latter variety are developed through multistakeholder, consensus-based

74. *See* Abbott & Snidal, *supra* note 6, at 525 ("Decentralization thus reduces demands on the state, a significant advantage in an era when many states and agencies face both shrinking resources and growing demands for action.").

75. *See* Coglianese, *Regulating Machine Learning*, *supra* note 59, at 12 ("[I]t is machine learning's heterogeneity that poses regulators' greatest challenge of all. These algorithms' varied forms, multiple uses, and dynamic proper ties make most conventional regulatory strategies obsolete."); *see also* Lobel, *supra* note 6, at 396 (explaining that centralized command-and-control oversight is less preferable under conditions of rapid advancement, heterogeneity, and complexity); Richard B. Stewart, *Administrative Law in the Twenty-First Century*, 78 N.Y.U. L. REV. 437, 446 (2003) (arguing that centralized regulation "suffers from the inherent problems involved in attempting to dictate the conduct of millions of actors in a quickly changing and very complex economy and society throughout a large and diverse nation"); Cass Sunstein, *Administrative Substance*, 40 DUKE L.J. 607, 627 (1991) (identifying the use of "highly bureaucratized 'command-and-control' regulation" to regulate a large number of subjects in an diverse country as a "large source of regulatory failure in the United States").

76. *See* Jody Freeman & Sharon Jacobs, *President Trump's Campaign of 'Structural Deregulation,'* LAWFARE (Feb. 12, 2025), https://www.lawfaremedia.org/article/president-trump-s-campaign-of--structural-deregulation.

processes, their nonbinding nature raises a fundamental question: will firms implement them on their own volition? Some self-regulation skeptics are likely to maintain that, absent an enforcement mechanism, a system of voluntary standards will only lead to more of the same "ethics washing" that existing, high-level AI frameworks have been charged with perpetuating.[77]

A closely related challenge is that competitive pressures are not guaranteed to push standardization outcomes in directions that align with public interests. Given that industry actors—to a much greater extent than other stakeholders—will be the ones facing standards adoption decisions and are thus primary arbiters of a standard's value, their criteria for evaluating the attractiveness of a standard will not necessarily be consistent with broader governance objectives. Similarly, in the context of inter-SDO competition, it is possible that standard-setting venues are selected not because they foster high-quality, effective standards but because they cater to specific stakeholder groups seeking to advance their own priorities. Critics may even contend that this invites a race to the bottom, where firms inevitably gravitate toward venues that allow them to shape less stringent standards with weaker protections for public values and interests.[78]

Though the tradeoffs identified above are undoubtedly real, they are not as unfavorable as they first appear, and there are even steps that can be taken to manage them more efficiently. One of the most powerful forces counteracting the risk of weak or inconsistently implemented standards is the overhanging threat of public regulatory intervention. If private governance efforts are perceived as inadequate or as failing to meaningfully address AI risks, regulators may respond by adopting a more top-down, coercive regulatory posture.[79] This latent threat may provide firms with strong incentives to voluntarily adopt more rigorous standards in order to stave off heavier-handed (and potentially disruptive) government action.[80] In fact, harnessing this dynamic and giving the threat some credibility—continuing to

77. *See supra* note 25 and accompanying text.

78. *See* Walters & Wiseman, *supra* note 58, at 562–64 (acknowledging the possibility of a race-to-the-bottom standards and SDO-based competition, though exploring several contextual factors that may mitigate it).

79. *See* IAN AYRES & JOHN BRAITHWAITE, RESPONSIVE REGULATION: TRANSCENDING THE DEREGULATION DEBATE 38–39 (1992) (maintaining that by signaling its willingness to escalate towards more coercive, interventionist regulatory strategies, the government gives industry actors incentives to "make regulation work at lower levels of interventionism"); Abbott & Snidal, *supra* note 6, at 523 (noting that "the threat of [state] intervention reinforces softer New Governance measures.").

80. John W. Maxwell, Thomas P. Lyon & Steven C. Hackett, *Self-Regulation and Social Welfare: The Political Economy of Corporate Environmentalism*, 43 J.L. & ECON. 583, 612–13 (2000) (finding empirical support for this proposition).

monitor the general landscape and periodically assessing the need for intervention—can be one of the most effective tools the government has for inducing firms to act without resorting to mandates.[81]

It is also important not to lose sight of the broader legal backdrop against which debates over AI regulation are taking place. The threat of new binding regulations is not the only reason firms have for taking AI safety and trustworthiness seriously on their own. The development and use of AI systems—at least in the United States—is already subject to several generally applicable, technology-neutral legal and regulatory frameworks.[82] For example, those who delegate hiring, housing, or loan application decisions to AI-based systems must still comply with employment, housing, and lending discrimination laws, respectively.[83] Likewise, when an AI system that is developed or used in a careless manner goes on to cause someone injury, the responsible firm can still be held liable in tort under negligence and/or products liability theories.[84] While existing legal frameworks may not offer a

81. The government acting in this capacity straddles the line between a pure self-regulatory approach and what has been described as "meta-regulation" (i.e., active state oversight and influence over self-regulatory efforts). *See* Cary Coglianese & Evan Mendelson, *Meta-Regulation and Self-Regulation*, *in* THE OXFORD HANDBOOK OF REGULATION 147, 161–62 (Robert Baldwin, Martin Cave & Martin Lodge eds., 2010). This logic also parallels what Tim Wu has identified as the strategic use of "agency threats" (e.g., warning letters, public speeches hinting at possible action, etc.) to influence private behavior absent formal rulemaking, which he argues is most justified when industries are experiencing rapid change and thus a great amount of uncertainty. *See generally* Tim Wu, *Agency Threats*, 60 DUKE L.J. 1841 (2011).

82. *See* Mariano-Florentino Cuéllar, *A Common Law for the Age of Artificial Intelligence: Incremental Adjudication, Institutions, and Relational Non-Arbitrariness*, 119 COLUM. L. REV. 1773, 1781 (2019) ("[S]ociety already 'regulates' AI . . . even in the absence of statutes and regulatory rules governing AI . . . . [T]he ultimate regulatory backstop here is the common law.").

83. *See, e.g.*, Exec. Order No. 14,110, § 7.3, 88 Fed. Reg. 75191, 72213 (Nov. 1, 2023), https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence (instructing the Consumer Financial Protection Bureau and the Department of Housing and Urban Development to provide guidance to the private sector on how to ensure AI-based decision making tools did not violate existing credit and housing discrimination laws); CONSUMER FIN. PROT. BUREAU, U.S. DEP'T OF JUST., EQUAL EMP. OPPORTUNITY COMM'N & FED. TRADE COMM'N, *Joint Statement on Enforcement Efforts Against Discrimination and Bias in Automated Systems* (Apr. 25, 2023), https://www.ftc.gov/system/files/ftc_gov/pdf/EEOC-CRT-FTC-CFPB-AI-Joint-Statement%28final%29.pdf ("[E]xisting legal authorities apply to the use of automated systems and innovative new technologies just as they apply to other practices.").

84. *See* Bryan H. Choi, *Negligence Liability for AI Developers*, LAWFARE (Sep. 26, 2024), https://www.lawfaremedia.org/article/negligence-liability-for-ai-developers (exploring how negligence liability might apply to AI developers).

perfect fit for AI and the process of adaptation is both slow and ongoing,[85] they do present some degree of legal risk for firms. AI standards not only help firms navigate this risk by defining implementable measures for addressing the underlying source (e.g., model flaws, bias), but they can even come to shape the law itself by serving as benchmarks that courts or regulators reference when assessing legal responsibilities. For instance, a firm's adherence to widely accepted, relevant technical standards may serve as evidence (albeit non-dispositive evidence) of reasonableness when facing negligence or defective design claims.[86]

Finally, though true that it is mostly commercial actors that will be developing and deploying AI systems and thus ultimately deciding which standards get adopted—giving them more leverage than other stakeholder groups when "voting with their feet"—this does not mean a race to the bottom is either the inevitable or likely result of inter-SDO competition. Unlike the Tiebout model, where it is assumed that community residents have perfect mobility,[87] there are still some constraints on the ability of industry actors to move freely to new venues that match their individual preferences. Consider a scenario in which industry stakeholders shifted to a more lenient multistakeholder venue that allowed them to develop more relaxed AI standards. Meanwhile, stakeholders from civil society, academia, and the technical community stayed behind and, through a consensus-based deliberative process, developed far more stringent standards designed to better uphold public interests.[88] Those adopting the industry-developed standard would risk facing immense public backlash for consciously spurning a more democratically legitimate alternative in favor of weaker public interest protections—backlash that, in addition to hurting a company's reputation, could culminate in the type of public regulatory interventions firms seek to avoid.[89] This gives commercial actors yet another incentive to cooperate with

85. *See generally, e.g.*, Andrew D. Selbst, *Negligence and AI's Human Users*, 100 B.U. L. REV. 1315 (2020) (identifying several challenges that negligence law poses for plaintiffs seeking redress for AI-related harms).

86. For an overview of how private technical standards interact with the tort system, see generally GARY E. MARCHANT, SWORDS AND SHIELDS: IMPACT OF PRIVATE STANDARDS IN TECHNOLOGY-BASED LIABILITY (2022), https://ssrn.com/abstract=4178750.

87. *See* Tiebout, *supra* note 54, at 419.

88. *Cf.* Gregory C. Shaffer & Mark A. Pollack, *Hard vs. Soft Law: Alternatives, Complements, and Antagonists in International Governance*, 94 MINN. L. REV. 706, 795–98 (2010) (explaining that non-state actors will often attempt to counter transnational soft law instruments they find unfavorable by either developing their own rival soft-law instruments or pushing international organizations, such as the UN, to do so).

89. *See* Julia Black, *Constructing and Contesting Legitimacy and Accountability in Polycentric Regulatory Regimes*, 2 REGUL. & GOVERNANCE 137, 146 (2008) (describing the concept of pragmatic legitimacy, wherein industry actors are motivated solely by economic interests to

other affected stakeholders at more inclusive SDOs: not just to shape the content of emerging standards, but to ensure those standards carry the legitimacy needed to secure public trust and forestall heavier-handed government oversight.

B. STAKEHOLDER INVOLVEMENT

1. Leveraging Field-Level Expertise and Inclusive Participation

Effective AI governance requires a deep understanding of the technology itself. Yet, because AI represents a highly complex and cutting-edge field, regulatory agencies—even specialized ones—inevitably face limitations in their technical capacity and knowledge.[90] That is not to say that governments are incapable of expanding this capacity. Indeed, both the first Trump and Biden administrations prioritized efforts to bring in top AI talent to federal agencies, commencing various initiatives to streamline the hiring process for those with AI-related technical experts.[91] However, competing with the private sector over a highly in-demand talent pool may be unrealistic, as private sector salaries for new AI PhDs can reach as high as $800,000.[92] Further compounding the difficulties here is that AI expertise is often highly domain-specific: different deployment contexts and use cases present unique considerations, meaning that hiring a handful of generalist experts may not be sufficient to address the full range of governance challenges AI presents.[93]

---

comply with private regulatory standards that consumers perceive as legitimate, even when they themselves find the standards normatively undesirable); Meidinger, *supra* note 53, at 525 ("Easy exit, however, is often seriously constrained by practical power. Many firms choose to subject themselves to supragovernmental regulatory standards not so much because they wish to live under them as because they feel that they must in order to avoid significant economic losses.").

90. *See* Gary E. Marchant & Carlos Ignacio Gutierrez, *Soft Law 2.0: An Agile and Effective Governance Approach for Artificial Intelligence*, 24 MINN. J.L. SCI. & TECH. 375, 384 (2023); *see also* Coglianese, *Regulating Machine Learning*, *supra* note 59, at 6 (emphasizing the challenges posed by the domain-specific nature of AI-related technical knowledge, which makes it unlikely that generalist AI agency could ever possess the expertise needed to regulate all of AI's heterogenous uses).

91. *See* Exec. Order No. 13,960, Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government, § 7, 85 Fed. Reg. 78939, 78942 (Dec. 3, 2020); Exec. Order No. 14,110, *supra* note 83, § 10.2.

92. *See* Sam Shedden, *AI Researcher Salary: Eye-Watering Pay of Top Tech Job Revealed*, READWRITE (Jan. 2, 2024), https://readwrite.com/ai-researcher-salary-eye-watering-pay-of-top-tech-job-revealed/; Chris Stokel-Walker, *Regulators Need AI Expertise. They Can't Afford It*, WIRED (Mar. 14, 2024), https://www.wired.com/story/regulators-need-ai-expertise-cant-afford-it/.

93. *See* Coglianese, *Regulating Machine Learning*, *supra* note 59, at 6.

It would not be a stretch to suggest that the technical knowledge possessed by government agencies is unlikely to ever surpass or even match that which is distributed across the private sector.[94] Hence, one of the key advantages of private standards-based governance is its ability to *directly* leverage the collective technical expertise and experience of those working at the forefront of AI development.[95] In addition to the advantages the private sector has in attracting human capital, those actively developing and deploying AI systems possess firsthand experience identifying different risks and failure modes as well as designing mitigation strategies.[96] These actors are also well-positioned to conduct ongoing real-world testing and provide continuous feedback on what works in practice versus what merely sounds good in theory.[97] This continuous cycle of iteration and learning gives private actors a unique advantage in shaping standards that are both technically sound and practically implementable.

When standards development is conducted in an open and cooperative setting, inclusive participation in the process has the potential to yield additional benefits. Allowing regulated entities to meaningfully contribute to the creation of standards can increase industry buy-in and enhance the likelihood that implementation will go beyond mere box-checking.[98] Equally important is the inclusion of non-industry voices in the process—especially civil society stakeholders representing public interests—which tends to strengthen the legitimacy of resulting standards.[99] Legitimacy here operates on two levels: there is what we might call *internal legitimacy*, which is concerned with the perception of participants at a particular venue or prospective adopters, and *external legitimacy*, which is concerned more broadly with whether

94. The EU AI Act's reliance on technical standards implicitly acknowledges the comparative advantage that specialized, multistakeholder bodies possess when it comes to translating broad regulatory goals into concrete technical practices. *See supra* note 39 and accompanying text.

95. *See* Lobel, *supra* note 6, at 382 (explaining that decentralized new governance approaches affirm the principle of subsidiarity: the idea that decisions are best made by those who possess superior information about a problem by virtue of their proximity to it); Abbott & Snidal, *supra* note 6, at 528–29 (noting the new governance approaches recognize "that expertise is often dispersed, and [seek] to harness a wide range of stakeholders who may have 'local' expertise otherwise unavailable to the state.").

96. *See* Walters & Wiseman, *supra* note 58, at 567 (noting that firms are particularly likely to possess superior knowledge about risks in a given industry during the early stages of its development).

97. *See* Meidinger, *supra* note 3, at 529.

98. *See* AYRES & BRAITHWAITE, *supra* note 79, at 113 ("If business is responsible for writing and enforcing its own code of conduct, the notion of regulation may become more palatable.").

99. *See infra* note 112 and accompanying text.

society at large perceives these governance arrangements as acceptable.[100] With AI standards, the latter type is of equal importance, if not even greater.

Why does the general public's perception of AI standards matter? As tempting as it may be to view AI standards as purely technical, this is not necessarily the case: the design of a technical standard can embed normative choices that have social, economic, and political consequences once implemented in the real world.[101] For example, benchmarks for evaluating model bias call for implicit value judgments about acceptable levels of discrimination risk, while transparency standards for disclosing information about a model's architecture or training data must determine how to weigh the benefits of information-sharing against the costs and risks that disclosure poses to the model developer.[102] Delegating highly consequential, value-laden choices to a narrow group of actors without any accountability mechanisms or stakeholder input is unlikely to be a politically tenable arrangement in the long term. Ensuring that diverse perspectives and interests are given meaningful weight in the deliberation process, however, helps mitigate these concerns and fosters legitimacy, making private standards-based governance a more sustainable approach.

2. Second-Best Considerations: Industry Capture and Participation Barriers

While private standards development can better leverage industry expertise in pursuit of more technically informed and effective governance mechanisms, the knowledge asymmetries that exist in such arrangements can be a double-edged sword. Since industry participants—or at least a particular subset of them—tend to possess a much more sophisticated understanding of both the underlying technology and the practical consequences of certain technical design choices, they might use their informational advantages in anticompetitive or self-serving ways.

---

100. *Cf.* Black, *supra* note 89, at 147 (explaining that a regulator's legitimacy depends on the perceptions of both those whose activities it directly governs and the broader group of actors on whose behalf it purports to govern).

101. For discussions of how internet standards can reflect political considerations, see generally Tarleton Gillespie, *Engineering a Principle: 'End-to-End' in the Design of the Internet*, 36 SOC. STUD. SCI. 427 (2006); LAURA DENARDIS, PROTOCOL POLITICS: THE GLOBALIZATION OF INTERNET GOVERNANCE (2009); Ian Brown, David D. Clark & Dirk Trossen, *Should Specific Values Be Embedded in the Internet Architecture?*, PROC. RE-ARCHITECTING INTERNET WORKSHOP, art. no. 10 (2010). For a discussion specifically focused on the political and normative dimensions of AI standards and their diffusion, see generally Alicia Solow-Niederman, *Can AI Standards Have Politics?*, 71 UCLA L. REV. DISC. 230 (2023).

102. Michael Veale & Frederik Zuiderveen Borgesius, *Demystifying the Draft EU Artificial Intelligence Act*, 4 COMPUT. L. REV. INT'L 97, 105 (2021).

A large firm, for instance, may seek to influence a particular standard's design to raise the costs of smaller rivals and thus further entrench its market position.[103] Intellectual property abuse has also historically been a significant concern in technical standard setting, manifesting through practices such as patent holdup, patent ambush, or royalty stacking, all of which can undermine the accessibility and adoption of standards.[104] Sometimes the problem may be as simple as industry stakeholders steering the design of a standard towards choices that privilege their own interests over public ones, insofar as the two diverge. AI's technical complexity only exacerbates these risks, as less technically sophisticated participants may struggle to detect the ways in which certain design choices could advance narrow industry priorities at the expense of broader societal ones.[105]

Technical knowledge is not the only area of asymmetry that may exist among participants. Disparities in resources for influencing the standards process can be equally problematic. Even when an SDO is nominally "open" and imposes no formal barriers to participation, meaningful engagement requires substantial time and financial investment. For civil society organizations and other non-industry stakeholders, this can serve as a de facto barrier that either severely limits their influence or excludes them from the process altogether.[106] Meanwhile, it is common for larger corporate

103. *See* Olia Kanevskaia, *Governance of ICT Standardization: Due Process in Technocratic Decision-Making*, 45 N.C. J. INT'L L. 549, 551 (2020) ("ICT standards may sometimes result in economic or administrative burdens by pushing up compliance costs for companies."); James J. Anton & Dennis A. Yao, *Standard-Setting Consortia, Antitrust, and High-Technology Industries*, 64 ANTITRUST L.J. 247, 249–50 (1995) (identifying firms' incentives to manipulate standards to raise rivals' costs despite there being "no technical rationale for creating such a disadvantage"); *see also* BREYER, *supra* note 48, at 115 ("While individual standards may not raise barriers significantly, a series of several standards . . . may well raise costs to the point where new firms will find it difficult to assemble sufficient capital to enter.").

104. *See, e.g.*, Mark A. Lemley & Carl Shapiro, *Patent Holdup and Royalty Stacking*, 85 TEX. L. REV. 1991 (2007); Janice M. Mueller, *Patent Misuse Through the Capture of Industry Standards*, 17 BERKELEY TECH. L.J. 623 (2002); Joseph Farrell, John Hayes, Carl Shapiro & Theresa Sullivan, *Standard Setting, Patents, and Hold-Up*, 74 ANTITRUST L.J. 603 (2007).

105. A similar phenomenon has also been observed in the administrative law context. *See* Wendy E. Wagner, *Administrative Law, Filter Failure, and Information Capture*, 59 DUKE L.J. 1321, 1333 (2010) ("[T]he ability to gain control of the rulemaking process through the use of excessive information may even be turned into a strategic advantage. Using technical terms and frames of reference that require a high level of background information and technical expertise, and relying heavily on 'particularized knowledge and specialized conventions,' these fully engaged stakeholders can deliberately hijack the proceedings.").

106. ADA LOVELACE INST., INCLUSIVE AI GOVERNANCE: CIVIL SOCIETY PARTICIPATION IN STANDARDS DEVELOPMENT 26 (Mar. 2023) (identifying the biggest barriers to civil society participation as "restrictive eligibility criteria . . . burdensome time

stakeholders to employ specialists who focus exclusively on standards work, permitting them to monitor and respond to every incoming standards proposal as well as to submit detailed proposals of their own at high volume.[107] This organizational capacity allows them to punch above their weight in the standards development process, which is often well worth the investment due to the financial stakes involved.[108] The upshot of punching above one's weight, however, is that it contributes to the perception that a particular SDO is captured by corporate interests, undermining the legitimacy of both the venue itself and any standards it produces.

That said, because this is a *comparative* second-best analysis, it is important to recognize that many of the challenges described above are not exclusive to standards setting or to private governance more generally. Resource disparities between different interest groups inevitably lead to differences in ability to shape governance outcomes regardless of the institutional arrangement. In the context of more state-centric modalities, these same dynamics express themselves in different forms, whether that be large firms dominating the notice and comment process in informal agency rulemaking or leveraging deep lobbying networks to track and influence legislative developments across all fifty states.[109] Indeed, a vast body of legal literature on regulatory capture documents how well-resourced actors often shape public governance mechanisms in ways that advantage their interests.[110]

To be sure, one could argue that corporate influence may be more pronounced in the context of private standards because commercial actors play a *direct* role in determining outcomes rather than merely influencing the process from the outside. At the same time, there is reason to believe the threat of corporate dominance may be somewhat exaggerated. The commercial actors involved in the standards process are not a monolith with perfectly aligned interests. The AI ecosystem consists of many types of firms operating at different layers of the AI stack (or different stages of the AI supply chain,

commitments, an inability to navigate complicated standardisation processes, industry dominance and a lack of awareness and interest.").

107. *See* BÜTHE & MATTLI, *supra* note 50, at 47.

108. *See id.* at 47–48.

109. *See, e.g.*, Wagner, *supra* note 105, at 1336–37; Elizabeth Warren, *Corporate Capture of the Rulemaking Process*, REGUL. REV. (June 14, 2016), https://www.theregreview.org/2016/06/14/warren-corporate-capture-of-the-rulemaking-process/ (citing an EPA study which found that "industry groups engaged in 170 times more informal communications with EPA than public interest players.").

110. The foundational work that gave rise to the theory of regulatory capture (despite not mentioning it by name) and that influenced decades of scholarship on the relationship between regulation and industry is George J. Stigler, *The Theory of Economic Regulation*, 2 BELL J. ECON. & MGMT. SCI. 3 (1971).

depending on how one conceptualizes it). This includes hardware vendors (e.g., GPU manufacturers), cloud infrastructure providers, model developers, and developers of AI "wrapper" applications—each with distinct and sometimes competing priorities. The heterogeneity of commercial interests can prevent industry stakeholders from speaking with a single voice and serve as a check against any one particular group dominating the entire process.[111] Of course, this alone does not prevent *non-industry* stakeholders from having their voices drowned out, nor does it guarantee that outcomes will align perfectly with public interests. Ensuring balanced representation of interests will also depend heavily on the institutional configuration of the standard-setting venue.

Where SDOs and other technology-focused private governance institutions have historically succeeded in maintaining their legitimacy, it has been due in large part to the openness—both in the participatory and informational sense—of their structures and processes.[112] Within SDOs specifically, openness typically manifests through transparent consensus-based mechanisms that allow for *meaningful* input and engagement from diverse interest groups (i.e., allowing participants to overcome any de facto barriers), thereby preventing any single constituency from dominating the process.[113] This helps ensure that standards are seen as the product of fair and democratic processes that are responsive to the concerns and interests of a broad range of stakeholders.

Thoughtful design of an SDO's structures and processes can go a long way to enhance openness while helping to mitigate concerns about corporate capture. This is likely to include substantive policies and procedural safeguards

111. *Cf.* David D. Clark, John Wroclawski, Karen R. Sollins & Robert Braden, *Tussle in Cyberspace: Defining Tomorrow's Internet*, 13 IEEE/ACM TRANSACTIONS NETWORKING 462 (2005) (arguing that the internet architecture reflects ongoing economic and political "tussles" among stakeholders with competing interests, and that system design should anticipate and balance such tensions).

112. *See* Mulligan & Bamberger, *supra* note 4, at 771 (explaining that past multistakeholder processes "have consistently focused on totems of participation . . . and transparency . . . for procedural legitimacy."); Kanevskaia, *supra* note 103, at 557 (maintaining that the standards produced by industry consortia who limit stakeholder participation and afford limited due process are more likely to lack legitimacy); *see also, e.g.*, A. Michael Froomkin, *Habermas@discourse.net: Toward a Critical Theory of Cyberspace*, 116 HARV. L. REV. 749, 798–805 (2003) (analyzing the IETF's governance model and concluding its open, participatory structure and processes satisfy the procedural requirements for legitimacy under Habermasian discourse theory).

113. *See, e.g.*, AYRES & BRAITHWAITE, *supra* note 79, at 57 (arguing that the concept of "contestability," which can serve as a countervailing force against the threat of regulatory capture, demands transparency and open access to information so that all interested groups can participate meaningfully).

in areas such as standards deliberation, consensus formation, representation in SDO leadership positions, and conflict of interest management (e.g., ex ante disclosure and licensing obligations for IP rights). Designing a system that successfully balances technical expertise, diverse stakeholder interests, practical implementability, and—as the following Section discusses—speed is an undeniably difficult task. The Internet Corporation for Assigned Names and Numbers (ICANN), the multistakeholder body that develops policies governing the technical management and distribution of internet name and number resources, illustrates this point. ICANN has created a complex institutional architecture aimed at ensuring balanced representation of stakeholder interests, including internal policymaking bodies organized and subdivided by stakeholder type,[114] a ruling Board of Directors composed of members drawn from diverse constituencies,[115] and various transparency, accountability, and review mechanisms.[116] This intricate structure notwithstanding, it is not uncommon to hear complaints about a small group of commercial players—particularly domain registries and registrars—wielding disproportionate influence over ICANN's activities.

Nevertheless, ICANN's continued functioning as a global internet governance body, despite its many imperfections, demonstrates that private governance can succeed even in areas where competing commercial, political, and societal interests exist.[117] For AI governance, the key lies in learning from these experiences and adapting institutional designs to the specific challenges of the field. AI standards bodies can indeed create governance frameworks that harness industry expertise while maintaining broader legitimacy. However, this will likely necessitate, among other things, the careful calibration of consensus thresholds (e.g., requiring consensus *within* each stakeholder group rather than across all participants),[118] the creation of targeted support

114. *See* Bylaws for Internet Corporation for Assigned Names and Numbers arts. III-VI (amended Jan. 9, 2025) [hereinafter ICANN Bylaws].

115. *See id.* arts. VII-VIII.

116. *See id.* arts. III-VI.

117. *See generally* Hortense Jongen & Jan Aart Scholte, *Legitimacy in Multistakeholder Global Governance at ICANN*, 27 GLOB. GOVERNANCE 298, 320 (2021) ("ICANN's experience shows what extent of legitimacy a global multistakeholder arrangement can realize in the early twenty-first century, particularly if it undertakes sustained intensive efforts to build support.").

118. It is crucial to understand that in the context of standards development, "consensus" does not equate to "unanimous consent." While there is no universal bright-line rule for determining when consensus exists, it is typically understood as "general agreement, characterized by the absence of sustained opposition to substantial issues by any important part of the concerned interests and by a process that involves seeking to take into account the views of all parties concerned and to reconcile any conflicting arguments." INT'L ORG. FOR STANDARDIZATION & INT'L ELECTROTECHNICAL COMM'N, *ISO/IEC Guide 2:2004, Standardization and Related Activities—General Vocabulary* ¶ 1.7 (2004). To better illustrate the

mechanisms for underrepresented stakeholders, and the establishment of mechanisms for enhanced transparency and managing conflicts of interest.

C. AGILITY

1. Greater Speed and Adaptability

One of the greatest challenges traditional regulation faces in governing emerging technologies is keeping pace with their rapid development.[119] There are at least two reasons for this. First, regulatory processes are often encumbered by procedural frictions. The rulemaking process typically involves multiple mandatory steps—such as notice and comment periods, impact analyses, and interagency reviews—that can take years to complete.[120] Even after rules are enacted, updating them to account for new developments may require going through the same time-consuming procedures. And that is not even to mention how partisan politics can further exacerbate these delays, making regulatory updates even slower and less responsive.

The second reason is that there is an information lag inherent to the regulatory process. Due to the fact that regulators are typically not "on the ground" where the innovation is happening, they often become aware of new advancements well after they have already occurred. Regulators are neither first-hand observers of what goes on in the R&D labs of major technology companies, nor are they privy to the conversations taking place in the hallways of industry conferences and trade shows. It is only once information about these developments eventually trickles down that they recognize the need to act, putting them perpetually behind the curve.[121]

AI bears all of the same regulatory challenges historically faced by other fast-changing emerging technologies and more, as it is not only evolving

---

implications of this definition, consider a scenario where civil society stakeholders constitute 15% of the total participants in an SDO and a large majority of these stakeholders are in objection to a proposed decision. It is still theoretically possible for "consensus" to be reached provided that virtually all of the remaining 85% of participants support the decision and there has been a genuine effort to consider civil society's objections and resolve any disagreements. Such a scenario could be prevented, however, if an SDO had rules defining different classes of stakeholders and requiring consensus to be present in each of them in order for a decision to be approved.

119. This is sometimes referred to as the "pacing problem." *See* Gary E. Marchant, *The Growing Gap Between Emerging Technologies and the Law*, *in* THE GROWING GAP BETWEEN EMERGING TECHNOLOGIES AND LEGAL-ETHICAL OVERSIGHT: THE PACING PROBLEM 19, 22–23 (Gary E. Marchant, Braden Allenby & Joseph Herkert eds., 2011).

120. *See* Lobel, *supra* note 6, at 390.

121. *See* Gervais, *supra* note 47, at 702 ("The speed of technological development also means that once it enters into force, regulation may, in fact, be outdated.").

quickly but doing so in unpredictable directions.[122] Two recent examples perfectly illustrate how the rapid pace of AI innovation is already straining traditional regulatory approaches.

The first can be seen in the initial draft of the EU AI Act. European regulators spent over a year devising a comprehensive framework built around the assumption that AI systems would be developed to serve individual use cases.[123] This assumption was turned on its head after the public release of ChatGPT, which became one of the most rapidly adopted technologies in history, surpassing one hundred million users within just two months of its introduction.[124] Prior to this point, large language models (LLMs) and other general-purpose AI did not appear to have been on the radar of European regulators despite OpenAI having already released multiple iterations of its GPT model in the preceding years (albeit to much less fanfare).[125] The sudden prominence of ChatGPT forced lawmakers to overhaul their draft legislation to account for general-purpose AI, delaying the regulatory process.[126]

A second example can be found both in the EU AI Act as well as an unsuccessful piece of AI safety legislation at the state level, California's proposed SB 1047. Both measures define "compute thresholds"—specific levels of computational power used during training that trigger regulatory obligations—to determine the responsibilities of AI model developers.[127] In the EU AI Act, models trained using at least 10^{25} floating point operations (FLOPS) of computation are presumed to be classified as "general purpose AI with systemic risk," a category that carries heightened regulatory obligations.[128] Similarly, SB 1047, a since-vetoed bill which aimed to mitigate the most serious

122. *See* Marchant & Gutierrez, *supra* note 90, at 384; *see also* Gervais, *supra* note 47, at 687, 701–02 (highlighting the dangers of regulatory interventions directed at "inchoate technologies" that evolve quickly along unpredictable trajectories).

123. *See* Gian Volpicelli, *ChatGPT Broke the EU Plan to Regulate AI*, POLITICO EU (Mar. 3, 2023), https://www.politico.eu/article/eu-plan-regulate-chatgpt-openai-artificial-intelligence-act/; Natali Helberger & Nicholas Diakopoulos, *ChatGPT and the AI Act*, 12 INTERNET POL'Y REV. 1 (2023), https://doi.org/10.14763/2023.1.1682.

124. Dan Milmo, *ChatGPT Reaches 100 Million Users Two Months After Launch*, GUARDIAN (Feb. 2, 2023), https://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app.

125. *See* Volpicelli, *supra* note 123.

126. *See id.*

127. *See* Artificial Intelligence Act, *supra* note 13, art. 51(2); S.B. 1047, § 3, 2023–2024 Leg., Reg. Sess. (Cal. 2024) (as enrolled Sep. 3, 2024), https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202320240SB1047 [hereinafter S.B. 1047].

128. *See* Artificial Intelligence Act, *supra* note 13, arts. 51, 55.

risks posed by the largest and most advanced AI models, would have imposed obligations on developers of models trained using more than 10^{26} FLOPS.[129]

Although both the EU AI Act and SB 1047 provided for mechanisms through which these thresholds could be amended over time, such updates are still subject to formal processes that inevitably introduce delay.[130] Indeed, in just the few months since California Governor Gavin Newsom vetoed SB 1047, major advancements in training techniques have already begun to challenge the assumption that achieving frontier-level capabilities requires such massive computational resources. Most notably, a Chinese hedge fund released an open-source model called DeepSeek R1 that reportedly matched the performance of advanced Western models despite using considerably less training compute.[131] This development led multiple commentators in the AI policy community to point out that, had SB 1047 been signed into law, its compute thresholds would have already been rendered obsolete before they even had a chance to take effect.[132] Beyond DeepSeek, similar risks of obsolescence loom with the rise of techniques like chain-of-thought reasoning and retrieval-augmented generation, which shift much of the computational burden to "inference time" and can produce sharp post-training jumps in model capability without increasing training compute.[133] In short, unforeseen technical developments—far from uncommon in an area as fluid as AI—can quickly frustrate even the most forward-thinking hard law schemes.

By comparison, private standards-based governance offers a nimbler alternative that can better adapt to external changes in the technological landscape. Since it is not *formally* burdened by the same constraints traditional regulation faces, standards bodies can potentially respond more promptly to

129. *See* S.B. 1047, *supra* note 127, § 3; *see also* Letter from Gavin Newsom, Governor of Cal., to Members of the Cal. State Sen. (Sep. 29, 2024), https://www.gov.ca.gov/wp-content/uploads/2024/09/SB-1047-Veto-Message.pdf.

130. *See* Artificial Intelligence Act, *supra* note 13, art. 51(3); S.B. 1047, *supra* note 127, § 3.

131. *See* Caiwei Chen, *How a Top Chinese AI Model Overcame U.S. Sanctions*, MIT TECH. REV. (Jan. 24, 2025), https://www.technologyreview.com/2025/01/24/1110526/china-deepseek-top-ai-despite-sanctions/; Sarosh Nagar & David Eaves, *AI's Efficiency Wars Have Begun: The DeepSeek Shock May Reshape a Global Race*, FOREIGN POL'Y (Feb. 5, 2025), https://foreignpolicy.com/2025/02/05/deep-seek-china-us-artificial-intelligence-ai-arms-race/. *But see* Dylan Patel, AJ Kourabi, Doug O'Laughlin & Reyk Knuhtsen, *DeepSeek Debates: Chinese Leadership on Cost, True Training Cost, Closed Model Margin Impacts*, SEMIANALYSIS (Jan. 31, 2025), https://semianalysis.com/2025/01/31/deepseek-debates/ (contesting widely reported estimates of DeepSeek's training costs).

132. *See, e.g.*, Timothy B. Lee (@binarybits), X, *Imagine how dumb the California legislature would look today if Newsom had signed SB 1047 last fall instead of vetoing it.* (Jan. 27, 2025, at 03:15 PM), https://x.com/binarybits/status/1751710571690037674.

133. *See* Sara Hooker, *On the Limitations of Compute Thresholds as a Governance Strategy*, ARXIV:2407.05694, at 5, 13–14 (July 8, 2024), https://arxiv.org/abs/2407.05694.

evolving technologies and emerging risks. It is worth noting that SDOs do end up recreating some of these constraints anyway, as many of the aforementioned "frictions" are in fact the product of procedural devices that help preserve legitimacy and ensure alignment with certain democratic ideals.[134] Still, private standards bodies have more room for institutional and process innovations, leaving them better equipped to strike an optimal balance between speed and legitimacy.[135] New initiatives and groups can also quickly convene without needing to worry about underlying legal authorities or possible judicial challenges like an administrative agency would.

As for the information lag, many of the private actors who participate in standards development are far less susceptible to it. Given their direct involvement in the ecosystem, these private actors are better positioned to see what is happening on the ground and gain a better sense of the technology's trajectory as it continues to develop. It is difficult to imagine these actors being caught off guard by major technological advancements that they themselves are responsible for. This closer proximity to technological development allows private standards bodies to more quickly identify when existing standards need revision or when entirely new approaches are warranted.

2. Second-Best Considerations: Prematurity and Tension with Legitimacy

Despite the potential for private standards processes to move faster than traditional regulation, there are pitfalls associated with moving *too* quickly and standardizing prematurely. Taking an anticipatory approach to standardization—attempting to get ahead of technological developments by establishing standards before a technology can be fully built and tested—requires that developers design standards without seeing what actually works in practice, forcing them to make assumptions about future needs and technological conditions.[136] Admittedly, this anticipatory approach may be seen as attractive for developing AI safety standards because it would appear better positioned to proactively address risks of harm before they materialize. According to this line of reasoning, permitting experimentation during the early stages and waiting for measures and techniques with a proven track record to emerge may needlessly subject the public to dangers that could

134. *See* Kanevskaia, *supra* note 103, at 554–55 (explaining that transnational private regulators, a category that can include technical SDOs, will often address legitimacy concerns through procedural frameworks rooted in administrative law principles such as due process, transparency, participation, and review).

135. *See* Lobel, *supra* note 6, at 390–91.

136. This challenge is exacerbated when the object being standardized is a component of a larger, complex system and thus has interdependencies (even if minimal) with other parts. *See* Christopher S. Yoo, *Modularity Theory and Internet Regulation*, 2016 U. ILL. L. REV. 1, 8–9.

otherwise be avoided by standardizing up front. However, standardizing prematurely can result in unproven standards that either fail to gain adoption, or even if they do, prove inefficient and ineffective—giving only the illusion of safety or security.[137]

The difficulty of developing effective standards through an anticipatory approach was one of the main lessons of the OSI-IETF internet standards war of the late 1980s and early 1990s. ISO's Open Systems Interconnect (OSI) architecture, which had the backing of many governments around the world, was carefully planned out in advance through a forward-thinking, bureaucratic process.[138] However, ISO's proactive standardization resulted in an over-engineered architecture that proved far more complex than was practical; the designers of OSI failed to account for the type of network capabilities for which there was legitimate demand and that had been proven to work in practice.[139] Meanwhile, the IETF's "rough consensus and running code" approach, which prioritized judging functional, real-world implementations of a prospective standard on its technical merits before approving it, produced the suite of protocols that ultimately prevailed.[140] Hence, even though SDOs may feel pressure to produce standards quickly, they must carefully consider the tradeoffs between speed and effectiveness.[141]

A closely related challenge and one that this Article briefly alluded to earlier is that moving fast often comes at the expense of legitimacy. Designing the structure and rules of an SDO to prioritize speed can compromise the perceived inclusiveness and impartiality of the standards process. For instance,

137. *See* NIST AI Standards Plan, *supra* note 29, at 5 ("Conversely, a standard that would attempt to get ahead of the underpinning science and engineering may be built on less rigorous technical foundations; it may prove unhelpful, counterproductive, or even technically incoherent.").

138. *See* JANET ABBATE, INVENTING THE INTERNET 168 (1999) ("The network standards effort was a departure from ISO's usual practice, in that it represented an attempt to standardize a technology that was still new and had not had a chance to stabilize . . . . But in the case of networks, some ISO members felt that formal standards should be outlined proactively.").

139. *See id.* at 176 (noting that many IT professionals saw OSI as unnecessarily complex and inefficient); *see also* Andrew L. Russell, *'Rough Consensus and Running Code' and the Internet-OSI Standards War*, 28 IEEE ANNALS HIST. COMPUTING 48, 53–54 (2006) (noting the internet engineers tended to view the OSI's design approach as "out of touch with existing networks and computers.").

140. *See* Russell, *supra* note 139, at 55.

141. A recent report on AI security standards published by the Alan Turing Institute's Centre for Emerging Technology and Security found that many SDOs who typically err on the side of technical maturity have reportedly been "facing pressure to standardize now." *See* Rosamund Powell, Sam Stockwell, Nalanda Sharadjaya & Hugh Boyes, *Towards Secure AI: How Far Can International Standards Take Us?*, CTR. FOR EMERGING TECH. & SEC. 31, 45 (Mar. 2024), https://cetas.turing.ac.uk/publications/towards-secure-ai.

providing for an override mechanism that allows quick resolution when disagreements arise—such as resorting to majority voting instead of continuing attempts at consensus-building—may expedite the process but leave stakeholders in the minority feeling marginalized.[142] Similarly, imposing narrow time windows where participants can provide input or voice objections to standards proposals may streamline standardization but at the cost of shutting out certain stakeholders, particularly those who lack the resources to respond promptly.[143] These process design choices, while facilitating agility, can thus undermine the very legitimacy that private standards-based governance requires to remain viable.

Conversely, the same participatory openness that lends legitimacy to SDOs can also hinder their progress by dramatically prolonging the process. As the number of participants grows, so too does the complexity of coordinating their input and resolving their often-conflicting interests. Discussions become unwieldy, decision-making more cumbersome, and consensus increasingly difficult to achieve. Notably, one of the core assumptions underlying Tiebout's foot-voting model is that communities have an optimal size; when this optimum is exceeded, a community can no longer provide local public goods at the same level and cost-efficiency that attracted residents in the first place.[144] Similarly, there comes a threshold beyond which adding more participants to an SDO does little to add to legitimacy and only contributes to crowdedness and gridlock, potentially negating one of the key advantages of standards-based governance. This tension adds an additional layer of complexity to the institutional and process design puzzle: SDOs must not only consider how to balance competing interests but also how to weigh speed and legitimacy when determining who gets a seat at the table and how decisions in the standardization process are made. This itself does not render private standards-based governance unworkable or inferior to alternatives—only demanding of careful institutional design.

D. SCALE

1. Better Positioning to Scale Across Borders

A final key advantage that private standards-based governance maintains over traditional regulation is its ability to facilitate coordination on a global scale. Whereas traditional regulation is territorially delimited—legal commands

142. *See id.* at 32 (reporting that interviewed SDO participants suggested that agility could be increased by "moving away from consensus-based agreements and towards majority-based voting and introducing streamlined approaches to submitting comments on draft standards.").

143. *Id.*

144. *See* Tiebout, *supra* note 54, at 419.

issued by sovereigns typically have no effect outside their jurisdiction—private standards more easily transcend borders. When standards are developed through open, multistakeholder arrangements and the final publications are made widely and freely accessible, they allow for broad participation and adoption from stakeholders regardless of their geographic location. In turn, this can help establish common or interoperable governance frameworks that extend across heterogenous legal systems and market contexts.

Global coordination may be less imperative for AI than for networked technologies and industries whose value depends heavily on widespread compatibility and interconnection. Take the internet, for instance: in order to function as a universal communication infrastructure—a single network that connects thousands of smaller networks from around the world—there must be some degree of global coordination around its core technical protocols as well as its naming and numbering systems.[145] The basic functionality of AI systems, by contrast, is not nearly as dependent on global uniformity: discrete AI systems can be built to conform to certain requirements in one jurisdiction without necessarily affecting those in other jurisdictions. That said, transnational cooperation around AI standards still offers significant advantages. Globally coordinated standards allow firms to develop a single AI system aligned with a widely accepted framework rather than incur the expensive burden of building multiple systems tailored to different local AI ecosystems. These standards can thus facilitate cross-border commerce and improve the ease of doing business for multinational firms that might otherwise face a complex patchwork of conflicting national requirements. They also support knowledge transfer between regions, defining globally recognized best practices that can help developing countries expand their domestic AI capacity and participate more meaningfully in the broader AI economy.

Governments, of course, have their own means of achieving cross-border regulatory harmonization, primarily through supranational organizations and agreements whereby member states develop and then independently enact uniform regulations. In fact, this approach would arguably be more effective at promoting global harmonization than private standards. Whereas the success of the latter is contingent on the absence of conflicting national regulations that would effectively preempt private standards, a multilateral regime aimed at regulatory harmonization would provide a clear roadmap for eliminating conflicting national regulations. However, this approach also has several drawbacks.

145. *See* LAURA DENARDIS, THE GLOBAL WAR FOR INTERNET GOVERNANCE 16–18 (2014).

First, it inherits many of the same weaknesses of national regulation—it is slow, has a top-down orientation, and entrusts policymaking to government representatives who are less likely to possess sufficient technical expertise. Second, it introduces several new complexities, such as conflicts between national values, divergent risk tolerances, and geopolitical rivalries. Multilateral talks around digital public policy issues such as privacy and international data flows have persistently stalled due to fundamental disagreements between major powers, giving little reason to expect that AI regulation would fare better.[146] Ultimately, if meaningful global coordination on AI governance is to be achieved, private standards-based governance represents a much more viable path forward.

2. Second-Best Considerations: Geopolitical Competition

Insofar as cooperation around AI standards does take on a transnational character, it tends to invite state involvement even when standards venues are nominally private. Though governments may not directly control the process, the potential political, economic, and social stakes give them strong incentives to try to influence their development.[147] Countries further recognize that leadership in setting AI standards can yield strategic advantages, conferring soft power, national prestige, and other economic advantages for domestic firms. Larger world powers may be tempted to influence private standards development if for no other reason than to prevent their adversaries from doing the same. If AI is indeed the transformative technology that many believe it to be, concerns will inevitably arise about rival states writing the rules

146. Because cross-border data flows have become an integral part of modern international commerce, attempts at multilateral regulatory cooperation have largely taken place within the context of negotiations over binding trade rules. *See, e.g.*, Anupam Chander & Paul M. Schwartz, *Privacy and/or Trade*, 90 U. CHI. L. REV. 49, 56–60, 65–69 (2023) (recalling the efforts to address transnational privacy issues during the Uruguay Round of negotiations that led to the creation of the WTO and how, rather than find a permanent solution that resolved underlying disagreements, these issues were ultimately "bracketed" through the inclusion of an open-ended GATS exception under which Members could justify trade-restrictive domestic privacy regulations). The most recent opportunity to make meaningful progress—plurilateral negotiations at the World Trade Organization under the Joint Statement Initiative on e-commerce—collapsed in 2023 following persistent disagreements among major economies and the abrupt withdrawal of the United States. *See* Alex Mueller, *One Step Forward, Two Steps Back: The United States' New Direction on Digital Trade*, 26 MINN. J.L. SCI. & TECH. 116 (2025).

147. *See* Hadrien Pouget, *What Will the Role of Standards Be in AI Governance?*, ADA LOVELACE INST. (Apr. 5, 2023), https://www.adalovelaceinstitute.org/blog/role-of-standards-in-ai-governance/ ("Since standards can carry some regulatory influence and relate to issues that matter to governments, governments are invested in their content and disagreements can arise.").

for AI to align with their preferences.[148] Traces of this can already be see in the United States, where AI development is now commonly framed as a race against China to determine whose vision for the technology prevails.[149]

The problem here is that the indirect involvement of states and infiltration of geopolitical rivalry into the standards arena threatens the integrity and effectiveness of the standards development process. It provides a new source of conflict and disagreement that can impede progress and significantly prolong standardization timelines. Similarly, distortions in the standards process arise when a country attempts to influence standards by placing its finger on the scale, potentially leading to suboptimal outcomes such as the adoption of technically inferior standards. There is already limited evidence of such scale tipping activity taking place at private SDOs. For example, participants at various technical SDOs have reported observing highly coordinated behavior by representatives of Chinese firms.[150] The *Wall Street Journal* even reported that, during a leadership election at the 3rd Generation Partnership Project (3GPP) that involved a candidate from Huawei, representatives from other Chinese companies were expected to capture proof that they cast their ballots for the preferred candidate.[151]

148. *See, e.g.*, Katherine Golden, *If the US and EU Don't Set AI Standards, China Will First, Say Gina Raimondo and Margrethe Vestager*, ATL. COUNCIL (Jan. 31, 2024), https://www.atlanticcouncil.org/blogs/new-atlanticist/if-the-us-and-eu-dont-set-ai-standards-china-will-first-say-gina-raimondo-and-margrethe-vestager/ (quoting former U.S. Commerce Secretary Gina Raimondo as warning that "if the US and EU don't show up [to set AI standards], China will, [and] autocracies will.").

149. *See* R. David Edelman, Diana Fu, Ryan Hass, Patricia M. Kim, Ying Lin, Yingyi Ma, Michael E. O'Hanlon, Melanie W. Sisson, Elham Tabassi & Nicol Turner Lee, *How Will AI Influence US-China Relations in the Next 5 Years?*, BROOKINGS INST. (June 18, 2025), https://www.brookings.edu/articles/how-will-ai-influence-us-china-relations-in-the-next-5-years/ (explaining how voices ranging from OpenAI CEO Sam Altman to former National Security Advisor Jake Sullivan have adopted this arms-race-with-China framing); *see also* Alexandra Alper & Jody Godoy, *AI Execs Say U.S. Must Increase Exports, Improve Infrastructure to Beat China*, REUTERS (May 8, 2025), https://www.reuters.com/world/us/us-ai-execs-give-congress-policy-wishlist-beating-china-2025-05-08/ (quoting Microsoft President Brad Smith as saying "the number-one factor that will define whether the U.S. or China wins this race is whose technology is most broadly adopted in the rest of the world.").

150. Emily de la Bruyère, *Setting the Standards: Locking in China's Technological Influence*, *in* CHINA'S DIGITAL AMBITIONS: A GLOBAL STRATEGY TO SUPPLANT THE LIBERAL ORDER 49, 57 (Nat'l Bureau Asian Rsch., NBR Special Rep. No. 97, Emily de la Bruyère et al. eds., 2022); Daniel R. Russel & Blake H. Berger, *Stacking the Deck: China's Influence in International Technology Standards Setting*, ASIA SOC'Y POL'Y INST. 12 (2021), https://asiasociety.org/sites/default/files/2021-11/ASPI_StacktheDeckreport_final.pdf.

151. Valentina Pop, Sha Hua & Daniel Michaels, *From Lightbulbs to 5G, China Battles West for Control of Vital Technology Standards*, WALL ST. J. (Feb. 8, 2021), https://www.wsj.com/articles/from-lightbulbs-to-5g-china-battles-west-for-control-of-vital-technology-standards-11612722698.

It is important to keep in mind, however, that these potential challenges around indirect government influence are far less severe than those that would arise in a more state-centric alternative where governments formally play a controlling role in setting transnational standards. Furthermore, just like many of the other challenges discussed throughout the Article, institutional design has an important role to play. Clear rules for achieving consensus, transparency and oversight mechanisms, and other procedural safeguards can help mitigate the risks of state interference. One specific option that SDOs may want to consider is permitting governments to participate in a limited advisory role—similar to ICANN's Government Advisory Committee (GAC)—where government representatives would have an opportunity to voice concerns over proposed standards without directly influencing consensus-based decision-making.[152] By providing a structured and transparent channel for governments to engage with SDOs, it could reduce the likelihood that they attempt to exert influence through indirect means. Of course, this is just an example, but it illustrates the type of design choices available to SDOs as they work to preserve the advantages of private standards-based governance while minimizing its vulnerabilities.

IV. CONCLUSION: THE PATH FORWARD

While no model for governing AI is perfect, private standards-based governance represents the most viable and compelling path forward. When implemented effectively, standards offer a powerful means of governing AI that can embed measurable expectations and constraints directly into the design, deployment, and oversight of AI systems. This approach offers significant advantages over traditional regulation, which often struggles under the weight of centralized authority, limited technical capacity, and procedural rigidity. By contrast, private standards-based governance stands to benefit from a bottom-up, multistakeholder architecture that enables experimentation and faster iteration, draws on deep field-level expertise and experience, and scales more readily across national borders. Its advantages lie not in its flawlessness but rather in its ability to navigate the realities of governing a complex, fast-moving, emerging technology. That said, this conclusion should

152. *See* ICANN Bylaws, *supra* note 114, art. XII, § 2(a). However, it should also be noted that the GAC has earned a fair share of criticism for the way it pulls an extraordinary group of actors (national governments) into the institutional fold as if it were any other stakeholder group, muddying the waters of power and authority over ICANN's activities. *See, e.g.*, Jonathan Weinberg, *Governments, Privatization, and "Privatization": ICANN and the GAC*, 18 MICH. TELECOMM. & TECH. L. REV. 189 (2011); MILTON MUELLER, NETWORKS AND STATES: THE GLOBAL POLITICS OF INTERNET GOVERNANCE 242–44 (William J. Drake & Ernest J. Wilson III eds., 2010).

not be mistaken for complacency, as there is still substantial work that remains if these advantages are to be realized.

Although many AI standardization initiatives are already underway, the task of building a robust standards-based regime is nowhere near complete. This reality is perhaps lost on some. In the United States, for example, most of the conversation around AI standards revolves heavily around NIST's AI RMF, which has led to the tendency to treat this framework as if it were intended as an exhaustive, turnkey solution to all of the AI governance challenges we currently face. To be sure, the AI RMF plays a valuable role in helping establish a common language for thinking about the organizational dimensions of AI governance. And NIST's supporting materials—such as its context-specific "profiles" and implementation "playbooks"—have made the framework more practical and usable for a variety of stakeholders.[153] But this body of work is, at most, only a first step. It does not (and was never meant to) address many of the technical dimensions of AI governance and cannot substitute for the full suite of standards needed to manage AI risk at scale across sectors and system types.

As AI standards work continues to unfold, several questions related to both the substantive and procedural aspects of AI standardization remain unanswered and will need to be confronted if this work is to fulfill the governance ambitions outlined throughout the Article. This includes the many questions related to SDO structure and process design that have been posed throughout this Article, such as how to mitigate the risks of industry capture or how to manage the tension between speed and legitimacy. There is also the "timing" question discussed in the previous Part: how to balance the mounting pressure to develop AI standards now with the practical reality that higher-quality standards often require greater technical maturity and iterative learning. Beyond these challenges, there are several more intricate questions about the appropriate scope and structure of standards themselves. A well-developed understanding of when and how standards should intervene across different stages of the AI system lifecycle remains lacking. For example, does pre-deployment testing and validation require different benchmarks and procedures than post-deployment? How many different modes of testing or validation are needed within a given stage, and do they require discrete standards?

153. *See, e.g.*, NAT'L INST. OF STANDARDS & TECH., AI RISK MANAGEMENT FRAMEWORK PLAYBOOK, *supra* note 38; NAT'L INST. OF STANDARDS & TECH., ARTIFICIAL INTELLIGENCE RISK MANAGEMENT FRAMEWORK: GENERATIVE ARTIFICIAL INTELLIGENCE PROFILE (NIST AI 600-1, 2024), https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf.

It would be a mistake to assume that all of these challenges and questions will inevitably resolve themselves with time. This Article does *not* argue that private standards-based governance is inherently self-correcting and that no further actions are necessary. Instead, as repeatedly underscored, the success of private standards-based governance will be heavily contingent on institutional design, or more specifically, configuring SDOs and their processes to make them more effective, inclusive, and legitimate. Sadly, several of the existing SDOs heavily involved in ongoing AI standards development leave much to be desired in this regard. Take ISO/IEC, for instance, which represents the earliest and perhaps most prominent AI standardization initiative to date. Its participation model requires stakeholders to engage indirectly through nationally designated standards bodies—entities that set their own rules for membership, engagement, and decision-making.[154] This structure leaves no formal avenues for direct participation by independent experts or civil society organizations (especially those who are not members of their respective national bodies), often limiting meaningful access to the standard-setting process.[155] Compounding these challenges, the resulting standards are typically locked behind paywalls and cost several hundred dollars to purchase, making it difficult for many stakeholders to access the very specifications intended to guide responsible AI development.[156] These major shortcomings in accessibility and transparency should give all stakeholders pause about whether current standardization efforts are on the right trajectory, and just as importantly, prompt renewed attention to their role in shaping these processes from the inside.

At the same time, responsibility for the success of private standards-based governance cannot fall solely on private actors. Traditional policymakers also have a role to play—not as top-down regulators, but as facilitators of effective and inclusive governance. In addition to protecting the integrity of the standards process by enforcing commitments and policing unfair or anticompetitive conduct within SDOs, they can serve as benign but ever-present backstops, strategically wielding the threat of formal regulation to steer private actors toward greater cooperation and a willingness to adopt more

154. *See* Kanevskaia, *supra* note 103, at 613.

155. *See id.* (noting that consumer and civil society groups can apply for "liaison" status in ISO/IEC technical committees, but the limited voting rights means they must resort to lobbying if they want to influence outcomes).

156. For example, ISO/IEC 42001, the foundational AI management standard, is priced at 199 Swiss francs (approximately 250 USD as of July 2025). *See* ISO, *ISO/IEC 42001:2023 Information Technology — Artificial Intelligence — Management System*, https://www.iso.org/standard/42001.

rigorous standards.[157] Policymakers should also consider the recommendations of those who have advocated for correcting imbalances by investing in the capacities of marginalized stakeholders, such as through stipends, technical assistance, or other focused interventions that reduce resource and knowledge gaps that influence standards process outcomes.[158]

In the end, the question is not whether standards will govern AI; indeed, they are already starting to do so.[159] The real question is whether policymakers and stakeholders will recognize their emergence for what it is—not just an inevitability but an opportunity—and respond with the seriousness it demands. For standards to serve as a legitimate and sustainable form of governance, there must be active investment in their design and the building of institutional foundations necessary for accountability, inclusivity, and public trust. Private standards-based governance offers great promise: it is agile, expert-driven, and globally scalable in ways that traditional regulation is usually not. But there is nothing inevitable about realizing this promise. Whether it becomes a tool for public good or merely the manifestation of particular interests will depend entirely on the actions taken going forward.

157. *Cf.* AYRES & BRAITHWAITE, *supra* note 79, at 38–41 (discussing the idea of the "benign big gun," wherein regulators can "speak softly" while maintaining a credible willingness to escalate up a regulatory pyramid toward more coercive strategies if voluntary cooperation fails).

158. *See* Margot E. Kaminski, *Voices In, Voices Out: Impacted Stakeholders and the Governance of AI*, 71 UCLA L. REV. DISC. 176, 194–95 (2024) (arguing that support for capacity-building efforts is necessary if stakeholders are going to meaningfully participate in AI governance); Mulligan & Bamberger, *supra* note 4, at 775 ("Meaningful participation in design debates further requires resources and strategies to bolster the uneven technological expertise among stakeholders.").

159. *See supra* text accompanying notes 33–39.