Unbundling AI Openness

Parth Nobel,[*] Alan Z. Rozenshtein[**] & Chinmayi Sharma[***]

2026 WISCONSIN LAW REVIEW (forthcoming)

The debate over AI openness—whether to make components of an artificial intelligence system available for public inspection and modification—forces policymakers to balance innovation, democratized access, safety and national security. By inviting startups and researchers into the fold, it enables independent oversight and inclusive collaboration. But technology giants can also use it to entrench their own power, while adversaries can use it to shortcut years and billions of dollars in building systems, like China's Deepseek-R1, that rival our own. How we govern AI openness today will shape the future of AI and America's role in it.

Policymakers and scholars grasp the stakes of AI openness, but the debate is trapped in a flawed premise: that AI is either "open" and "closed." This dangerous oversimplification—inherited from the world of open source software—belies the complex calculus at the heart of AI openness. Unlike traditional software, AI is a composite technology built on a stack of discrete components—from compute to labor—controlled by different stakeholders with competing interests. Each component's openness is neither a binary choice nor inherently desirable. Effective governance demands a nuanced understanding of how the relative openness of each component serves some goals while undermining others. Only then can we determine the trade-offs we are willing to make and how we hope to achieve them.

---

[*] PhD Candidate, Electrical Engineering, Stanford University; Intern, Amazon.com. Nobel was supported in part by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1656518. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation or his employer.

[**] Associate Professor of Law, University of Minnesota Law School; Senior Editor and Research Director, Lawfare; Visiting Senior Fellow, Institute for Law & AI; Nonresident Senior Fellow, Brookings Institution. Rozenshtein consults on a range of technology law and policy issues.

[***] Associate Professor of Law, Fordham Law School; Contributing Editor, Lawfare; Advisor, American Law Institute Principles of the Law, Civil Liability for Artificial Intelligence; Member, Microsoft Responsible AI Committee. For helpful comments, the authors thank Ryan Calo, James Grimmelmann, Woody Hartzog, John Speed Meyers, Paul Ohm, Sana Pandey, Neal Parikh, Ashwin Ramaswami, Andrew Selbst, Keith Winstein, and Bianca Wylie, and, for excellent research assistance, Emma Haberman, Ben Evelev, and Audrey Kim.

This Article aims to equip policymakers with the analytical toolkit to do just that. First, it introduces a novel taxonomy of "differential openness," unbundling AI into its constituent components and illustrating how each one has its own spectrum of openness. Second, it uses this taxonomy to systematically analyze how each component's relative openness necessitates intricate trade-offs both within and between policy goals. Third, it operationalizes these insights, providing policymakers with a playbook for how law can be precisely calibrated to achieve optimal configurations of component openness.

AI openness is neither all or nothing nor inherently good or evil—it is a tool that must be wielded with precision if it has any hope of serving the public interest.

Introduction

The question of "AI openness"—who controls artificial intelligence, who benefits from it, and who bears responsibility for its failures—has rapidly evolved from an obscure debate among scholars and programmers into a flash-point in global policy, corporate strategy, and international affairs.[1] On the one hand, "open spectrum AI" (osAI)—a term this Article coins in lieu of the more

[1] *See, e.g.*, EXEC. OFF. PRES., WINNING THE RACE: AMERICA'S AI ACTION PLAN 4–5 (2025), https://www.ai.gov/action-plan; *see also* Iain Martin, *The EU is Betting $56 Million on Open Source AI*, FORBES (Feb. 2, 2025), https://www.forbes.com/sites/iainmartin/2025/02/02/the-eu-is-betting-56-million-on-open-source-ai/; *see* Troy Wolverton, *AI's Openness Is Being Sharply Debated by Technologists, Policymakers*, S.F. EXAM'R (July 21, 2025), https://www.sfexaminer.com/news/technology/open-source-ai-debate-sharp-among-technologists-politicians/article_ab781b42-28e7-11ef-836b-9b118373b94c.html.

common "open source AI" to more accurately capture the complexity of systems that are, to some degree, free and publicly available for inspection, use, and modification[2]—has shown itself capable of being a force for profound good. Google DeepMind's AlphaFold predicts protein structures with revolutionary accuracy, accelerating drug discovery.[3] In a Boston hospital, Meta's Llama 3.1 performs on par with leading proprietary AI in generating differential diagnoses for complex cases, promising to enhance clinical decision support while protecting patient data.[4] Meanwhile, conservationists deploy Wildbook, a system using computer vision to identify individual animals from photographs, creating a massive, crowdsourced database to track endangered species and inform conservation policy.[5]

Yet, in the shadows of this progress, the same powerful osAI technologies are weaponized. Cybercriminals unleash WormGPT, an AI built on the open GPT-J model, to craft malware and highly convincing phishing emails with effortless precision.[6] Meanwhile, state-backed influence operations deploy networks of fake social media accounts using StyleGAN-generated profile pictures—synthetic faces of non-existent people—to amplify propaganda, discredit critics, and distort international discourse on human rights and global events.[7] And a recent investigation into a widely used open training dataset uncovered thousands of images of child sexual abuse material, tainting the very foundation of popular image-generation models.[8] This is the paradox of osAI: a single technological wellspring feeding both lifesaving innovation and sophisticated digital malice.

---

[2] *See infra* Part I.

[3] *See* Josh Abramson et al., *Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3*, 630 NATURE 493, 493, 496–97 (2024), https://doi.org/10.1038/s41586-024-07487-w.

[4] *See* Thomas A. Buckley et al., *Comparison of Frontier Open-Source and Proprietary Large Language Models for Complex Diagnoses*, 6 JAMA HEALTH F., no. 3, 2025, at 1, 2, https://doi.org/10.1001/jamahealthforum.2025.0040.

[5] *See* Tanya Y. Berger-Wolf et al., *Wildbook: Crowdsourcing, Computer Vision, and Data Science for Conservation*, *in* BLOOMBERG DATA FOR GOOD EXCH. CONF. 2017 (Oct. 24, 2017), https://doi.org/10.48550/arXiv.1710.08880.

[6] *See* Chuck Easttom, *Malicious Use of Artificial Intelligence*, 2025 IEEE 15TH ANN. COMPUTING AND COMMC'N WORKSHOP AND CONF. (CCWC) 499, 500 (2025), https://doi.org/10.1109/CCWC62904.2025.10903787.

[7] BENJAMIN STRICK, CTR. FOR INFO. RESILIENCE, ANALYSIS OF THE PRO-CHINA PROPAGANDA NETWORK TARGETING INTERNATIONAL NARRATIVES 4 (2024), https://www.info-res.org/app/uploads/2024/11/Analysis-of-the-Pro-China-Propaganda-Network-Targeting-International-Narratives_FINAL.pdf.

[8] DAVID THIEL, STAN. INTERNET OBSERVATORY CYBER POL'Y CTR., IDENTIFYING AND ELIMINATING CSAM IN GENERATIVE ML TRAINING DATA AND MODELS 1 (2023), https://purl.stanford.edu/kh752sm9123.

Meanwhile, openness is becoming a key driver of the AI market, which the United Nations projects to reach nearly $5 trillion in less than a decade.[9] Open models often lag their closed counterparts by less than six months,[10] a narrow gap that is fueling high-profile clashes and strategic moves across the industry. For example, Elon Musk sued OpenAI, accusing it of breaching a promise to put the public before profits and demanding it return to its open source roots.[11] In a widely publicized jab, he offered to drop the suit if the company simply renamed itself "ClosedAI."[12] Subsequently, Musk's own company, xAI, released its powerful Grok model under an open license.[13] Meta has also made a strategic bet on openness, championing its Llama models as "open-source" catalysts for innovation and security that can coexist with profitability.[14] Even OpenAI's CEO Sam Altman has acknowledged that the company's closed approach may have placed it "on the wrong side of history"[15]—signaling a shift from one of the industry's leading closed-model advocates—and it has released a leading open spectrum AI model of its own.[16] The debate has been further intensified by the rise of powerful osAI models from Chinese labs like DeepSeek, which now rival

[9] U.N. CONF. TRADE & DEV., TECHNOLOGY AND INNOVATION REPORT 2025: INCLUSIVE ARTIFICIAL INTELLIGENCE FOR DEVELOPMENT 6 (2025), https://unctad.org/publication/technology-and-innovation-report-2025.

[10] Ben Cottier et al., *How Far Behind Are Open Models?*, EPOCHAI (Nov. 4, 2024), https://epoch.ai/blog/open-models-report#open-models-have-lagged-on-benchmarks-by-5-to-22-months.

[11] Complaint at 5-9, Musk v. Altman, No. CGC-24-612746, (Cal. Super. Ct. filed Feb. 29, 2024); *see also* Adam Satariano, Cade Metz & Tripp Mickle, *Elon Musk Sues OpenAI and Sam Altman for Violating the Company's Principles*, N.Y. TIMES (Mar. 1, 2024).

[12] Elon Musk (@elonmusk), X (Mar. 6, 2024, 11:10 ET), https://x.com/elonmusk/status/1765409615070601417.

[13] *See Open Release of Grok-1*, XAI (Mar. 17, 2024), https://x.ai/news/grok-os.

[14] *See* Mark Zuckerberg, *Open Source AI is the Path Forward*, META (July 23, 2024), https://about.fb.com/news/2024/07/open-source-ai-is-the-path-forward/. Whether Meta's commitment to AI openness continues remains to be seen. *See* Mark Zuckerberg, *Personal Superintelligence*, META (July 30, 2025), https://www.meta.com/superintelligence/ ("We'll need to be rigorous about mitigating . . . risks and careful about what we choose to open source.").

[15] Kyle Wiggers, *Sam Altman: OpenAI Has Been on the "Wrong Side of History" Concerning Open Source*, TECHCRUNCH (Jan. 31, 2025).

[16] *Introducing gpt-oss*, OPENAI (Aug. 5, 2025), https://openai.com/index/introducing-gpt-oss/.

the performance of the proprietary Western systems upon which it was built[17] and at significantly lower cost,[18] adding a new layer of geopolitical urgency.[19]

The stakes of the debate over osAI couldn't be higher. On one view, openness is the key to unlocking unprecedented innovation, democratizing access to powerful technology, and enhancing safety by subjecting AI systems to broad, independent scrutiny. On the other, osAI risks catastrophic misuse, threatens national security, and will inevitably be co-opted by dominant corporate interests, reinforcing the very power structures they claim to challenge.[20]

But despite these high stakes, the discourse is dangerously oversimplified, flattening the concept of AI openness into an inaccurate "open versus closed" binary.[21] This misleading view has its roots in the history of open source software but is ill-suited for governing open spectrum AI. While openness in traditional software primarily meant access to source code, osAI systems are complex, layered technologies composed of multiple interdependent components: computational hardware that powers AI, training data that shapes capabilities, model weights that encode knowledge, source code that defines structure, operational records and controls that reveals performance characteristics, and the humans putting it all together.[22] Each component exists on its own spectrum of openness and carries distinct implications for safety, innovation, democratic control, and national security.[23] So, in the context of AI ecosystems, openness refers to the degree to which these components are transparent in their operation, accessible to external scrutiny or use, and inclusive of diverse contributors throughout the development process.

By stripping the discourse around osAI of its necessary complexity, policymakers fail to address the nuanced trade-offs inherent in its governance and risk undermining the very goals they seek to achieve. With few exceptions,[24]

---

[17] *See* Luis E. Romero, *ChatGPT, DeepSeek, or Llama? Meta's LeCun Says Open-Source is the Key*, FORBES (Jan. 28, 2025).

[18] *See* Kevin Roose, *Why DeekSeek Could Change What Silicon Valley Believes About A.I.*, N.Y. TIMES (Jan. 28, 2025); *see also* Prithwiraj Choudhury, Natarajan Balasubramanian & Mingtao Xu, *Why DeepSeek Shouldn't Have Been a Surprise*, HARV. BUS. REV. (Jan. 30, 2025), https://hbr.org/2025/01/why-deepseek-shouldnt-have-been-a-surprise.

[19] *See* LAURIE HARRIS, CONG. RSCH. SERV., IF13051, DEEPSEEK AND THE RACE TO DEVELOP ARTIFICIAL INTELLIGENCE (2025), https://www.congress.gov/crs-product/IF13051.

[20] *See infra* Part II.

[21] *See infra* Part I.A.

[22] *See infra* Part I.B.

[23] *See infra* Part II.

[24] *See, e.g.*, BIPARTISAN HOUSE TASK FORCE REP. ON A.I. 155 (2024), https://www.speaker.gov/wp-content/uploads/2024/12/AI-Task-Force-Report-FI-

policymakers tend to treat the openness of a model as a single, undifferentiated feature, without parsing the degree to which specific components are actually open. For example, the most thorough policy document on the topic, the National Telecommunication and Information Administration's (NTIA) 2024 report, focused almost entirely on the risks of releasing open model weights.[25] The more recent Trump Administration "AI Action Plan" does more or less the same, equating "open source" with "open weight."[26] Abroad, the EU AI Act reflects a similarly oversimplified conception of AI openness. By providing regulatory forbearance to models as long as their weights, architecture, and data-usage information are publicly available, the Act misses the opportunity to demand transparency in datasets, a component crucial to meaningful accountability.[27]

To craft effective regulation, a more precise framework is required. AI openness cannot be treated as a simple binary; it must be assessed at the component level, with more scrutiny devoted to how the relative openness of each component impacts safety, innovation, democratic control, and national security. While academic work has begun to recognize AI openness as a multifaceted issue,[28] much of the literature remains focused on the accessibility of

---

NAL.pdf ("Despite often being characterized as either open or closed, there is in fact a continuum of different forms of AI model availability and transparency. . . . [D]ifferent parts of a model can be made open while others remain closed.").

[25] NAT'L TELECOMM. & INFO. ADMIN., DUAL-USE FOUNDATION MODELS WITH WIDELY AVAILABLE MODEL WEIGHTS REPORT (2024), https://www.ntia.gov/programs-and-initiatives/artificial-intelligence/open-model-weights-report.

[26] EXEC. OFF. PRES., *supra* note 1, at 4–5.

[27] 2024 O.J. (L 1689), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689.

[28] *See* ELIZABETH SEGER ET AL., CTR. GOVERNANCE A.I., OPEN-SOURCING HIGHLY CAPABLE FOUNDATION MODELS: AN EVALUATION OF RISKS, BENEFITS, AND ALTERNATIVE METHODS FOR PURSUING OPEN-SOURCE OBJECTIVES 8–11, 13–14, 26–28 (2023), https://arxiv.org/abs/2311.09227; *see also* Rishi Bommasani et al., *Considerations for Governing Open Foundation Models*, 386 SCIENCE 151, 153 (2024), https://doi.org/10.1126/science.adp1848; Irene Solaiman, *The Gradient of Generative AI Release: Methods and Considerations*, PROCS. 2023 ACM CONF. ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY (FACCT '23) 111, 111, 113–14 (2023), https://doi.org/10.1145/3593013.3593981; David Gray Widder et al., *Why "Open" AI Systems Are Actually Closed, and Why This Matters*, 635 NATURE 827, 829–31 (2024), https//doi.org/10.1038/s41586-024-08141-1; Tejas N. Narechania & Ganesh Sitaraman, *An Antimonopoly Approach to Governing Artificial Intelligence*, 43 YALE L. & POL'Y REV. 95, 120 (2024); Sayash Kapoor et al., *Position: On the Societal Impact of Open Foundation Models*, PROCS. 41ST INT'L CONF. MACH. LEARNING (ICML '24) 23,082 (2024), https://doi.org/10.5555/3692070.3692998; Andreas Liesenfeld & Mark Dingemanse, *Rethinking Open Source Generative AI: Open-Washing and the EU AI Act*, PROCS. 2024 ACM

model weights,[29] and even more nuanced analyses often stop short of connecting component-level distinctions to concrete legal strategies.

CONF. ON FAIRNESS, ACCOUNTABILITY & TRANSPARENCY (FAcct '24) 5 (2024), https://doi.org/10.1145/3630106.3659005; Francisco Eiras et al., *Near to Mid-Term Risks and Opportunities of Open Source Generative AI,* 235 PROCS. 41ST INT'L CONF. MACH. LEARNING (July 2024), https://doi.org/10.5555/3692070.3692561; Nik Marda, Jasmine Sun & Mark Surman, *Public AI: Making AI Work for Everyone, by Everyone*, MOZILLA FOUND. (Sep. 30, 2024), https://www.mozillafoundation.org/en/research/library/public-ai/; Matt White et al., *The Model Openness Framework: Promoting Completeness and Openness for Reproducibility, Transparency and Usability* (Oct. 18, 2024) (unpublished manuscript), https://doi.org/10.48550/arXiv.2403.13784; Tamara Paris, AJung Moon & Jin L.C. Guo, *Opening the Scope of Openness in AI*, *in* FAccT '25: 2025 ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY &TRANSPARENCY 1293, 1296-1302 (June 23, 2024), https://doi.org/10.1145/3715275.3732087 (analyzing openness from an interdisciplinary perspective, focusing on interactivity, freedom, and inclusiveness) ; David Atkinson*, Open Shouldn't Mean Exempt: Open-Source Exceptionalism and Generative AI* 1, 4 (July 23, 2025) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5355736 (focusing on the open-closed source binary, and the products open-source AI produces).

[29] *See, e.g.,* PREM M. TRIVEDI & NAT MEYSENBURG, OPEN TECH. INST., OPENNESS IN ARTIFICIAL INTELLIGENCE MODELS 10-12 (2024), https://www.newamerica.org/oti/reports/openness-in-artificial-intelligence-models/ (advocating for AI openness to address more than model weights and source code to include transparency in the development process, but stopping short of acknowledging other components in the AI stack); Matthew Leisten, *Open (?) AI* 3–5 (Fed. Trade Comm'n Working Paper, 2024), https://doi.org/10.2139/ssrn.5044391 (focusing exclusively on model weights and architecture); Kaige Gao, Youngjin Yoo & Aaron Schecter, *Open Source AI Community as "Trading Zone": The Role of Open-Source Models in the Diffusion of Artificial Intelligence Innovation, in* 45 INT'L CONF. ON INFO. SYS. (2024), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5019689 (focusing exclusively on models); Yujin Potter et al., *"As an AI, I Believe AI Models Should Be Open Source"* 6-7 (2024) (unpublished manuscript), https://rdi.berkeley.edu/research/uploads/LLM_open_vs_closed.pdf (flattening AI openness into a binary); MIKE SEXTON, THIRD WAY, OPEN-SOURCE IS A NATIONAL SECURITY IMPERATIVE (2025), https://www.thirdway.org/report/open-source-ai-is-a-national-security-imperative (flattening AI openness into a binary); Domen Vake et al., *Is Open Source the Future of AI? A Data Driven Approach, in* 15 APPLIED A.I. & DATA SCI. 16 (2025), https://doi.org/10.3390/app15052790 (flattening AI openness into a binary); MASAO DAHLGREN, CTR. STRATEGIC & INT'L STUD., DEFENSE PRIORITIES IN THE OPEN-SOURCE AI DEBATE (2024), https://www.csis.org/analysis/defense-priorities-open-source-ai-debate (recognizing more components can be relevant but focusing exclusively on model weights); DIGIT. PUB. GOODS ALL., CORE CONSIDERATIONS FOR EXPLORING AI SYSTEMS AS DIGITAL PUBLIC GOODS 4–5 (2023) https://www.digitalpublicgoods.net/AI-CoP-Discussion-Paper.pdf (focusing exclusively on models and data); Thibault Schrepel & Jason Potts, *Measuring Openness of AI Foundation Models: Competition and Policy Implications, in* 2 RADIOLOGY SCI. (2023), https://doi.org/10.15212/RADSCI-2023-0007 (focusing on a binary of openness and its effects on the medical field); Thibault Schrepel & Jason Potts, *Measuring*

This Article fills that gap by introducing a framework of "differential openness" that rejects oversimplified labels of "open AI" or "open source AI" for the more analytically precise "open spectrum AI" (osAI). It "unbundles" osAI systems into their constituent components, mapping each along a gradient of openness and evaluating how specific configurations of openness advance or undermine public goals. This taxonomy provides policymakers with the analytical tools needed to navigate the complex trade-offs with precision and craft targeted, calibrated interventions that maximize benefits while mitigating risks.

To develop this argument, this Article proceeds in three parts. Part I dismantles the flawed "open versus closed" binary and introduces our taxonomy of what we call "differential openness" for osAI. It begins by challenging the open source software analogy, demonstrating that the governance frameworks built for it are technologically and culturally ill-suited for osAI systems, which are defined by many modular, interconnected components and a complex ecosystem of corporate and state actors. It then unbundles AI systems into their seven key components—compute, data, source code, model weights, system prompts, operational records and controls, and labor—to establish a more precise vocabulary for analyzing how openness functions at each layer of the AI stack.

Part II uses this new taxonomy to systematically evaluate how different configurations of component-level openness advance or undermine core policy objectives: public safety, innovation and economic growth, democratic accountability, and national security. We analyze the complex, often contradictory, effects of differential component openness on each goal. This analysis reveals that while openness is often a powerful engine for progress, it is not an intrinsic good but instead an instrumental value whose desirability depends entirely on context, forcing policymakers to confront the difficult trade-offs inherent in osAI governance.

Finally, Part III moves from diagnosis to prescription, examining the specific legal and regulatory levers policymakers can use to govern osAI. We analyze how tools related to liability, competition policy, intellectual property, trade controls, and direct government support can target specific components of the AI stack. This provides a concrete playbook for designing nuanced interventions that move beyond blunt, system-level mandates to foster a safer, more innovative, and more accountable osAI ecosystem.

Openness of AI Foundation Models: Competition and Policy Implications, 2 INFO. & COMMC'N TECH. L. 1, 8–13 (2025), https://doi.org/10.1080/13600834.2025.2461953 (advancing a broader multidimensional framework but ultimately flattening openness into a license-and-access-centric typology); Alex Engler, *How Open-Source Software Shapes AI Policy*, BROOKINGS INST. (Aug. 10, 2022), https://www.brookings.edu/articles/how-open-source-software-shapes-ai-policy/ (flattening AI openness into a binary).

The open-versus-closed binary and the assumption it is an unmitigated good or evil is a siren song that has already led policy astray. Effective governance requires abandoning this simplistic lens and embracing a more sophisticated, differential openness framework for governing osAI—one that calibrates policy to the distinct risks and benefits of each component of the AI stack. Only by unbundling AI in this way can we move beyond ideological debates and begin the difficult but essential work of crafting targeted rules for a technological future that best serves the public interest.

I. A Taxonomy of AI Openness

The debate over open spectrum AI is distorted by its inheritance from the history of open source software (OSS). This legacy has generated two core misconceptions that skew policy: a false "open closed" binary and the reflexive assumption that openness is an inherent good. This simplistic framing fails because it imports assumptions from a different technological and institutional era. The OSS world—driven largely by individual developers and academics focused on opening source code—is fundamentally distinct from the modern osAI ecosystem, which involves a complex stack of interdependent components, each of which exists on its own spectrum of openness[30] and is controlled by one of a concentrated set of powerful corporate and state actors.[31]

This Part dismantles the flawed OSS analogy and, in its place, introduces this Article's core contribution: a taxonomy that "unbundles" AI systems into their key technical and human components. By revealing osAI's differential openness—the many dimensions along which openness actually varies—we demonstrate that the value of opening any single component is not innate but instrumental, capable of advancing some policy goals while simultaneously undermining others.

This taxonomic precision is essential for effective governance. It moves the analysis beyond asking *if* a system is open to asking more critical questions: what is open, how open is it, and to what end? Answering these questions is a prerequisite for crafting policies that can effectively balance the competing values at stake in AI development—safety, innovation, democratic control, and national security—the central task we undertake in Part II.

A. Beyond the Open Source Software Analogy

The debate over open spectrum AI inherits a great deal—some of it useful, much of it not—from the history of OSS. This impulse, however misguided, is

[30] Solaiman, *supra* note 28, at 111–14.

[31] Widder et al., *supra* note 28, at 827.

understandable: openness yielded considerable benefits, from innovation to decentralized governance.[32] Although there are similarities in the core concepts of openness, the traditional software and AI are fundamentally distinct. To offer more rigor and nuance to the debate, this section dismantles the flawed OSS analogy by establishing three distinctions that require us to look beyond source code, beyond individual developers, and beyond the AI labs themselves. Ultimately, it warns policy makers against assuming that osAI can be governed by the OSS playbook.

1. Beyond Source Code

The first reason that the OSS analogy fails as applied to osAI is that software is a comparatively simpler technology, and its model of openness—while revolutionary when it was developed—is insufficient to capture the complex, multi-layered reality of AI.[33] The success of OSS hinged on making openness legible, scalable, and enforceable by focusing on a single, critical component: source code.[34] The OSS community accomplished this with a technical mechanism for distributing open source code and a legal mechanism for enforcing its continued openness.

The technical dimension of software openness is straightforward. The value of an OSS project is unlocked almost entirely by making its source code free and publicly accessible. Code repositories like Microsoft's GitHub provide the infrastructure for this, creating a universal platform where the OSS source code lives, allowing developers, from the original authors to third-party contributors, to inspect, modify, and contribute to a project. This can take the form of reporting bugs (or flaws in the software), suggesting improvements (from safety enhancements to improved efficiency), or building off the software in innovative ways (from new use cases to new capabilities).[35]

This technical access is made legally meaningful through a legal mechanism: spectrum of OSS licenses. By default, copyright law grants exclusive rights to the creator.[36] OSS licenses strategically override this default, creating a durable legal basis for permissionless use and collaboration. These licenses are not monolithic; they represent a range of tradeoffs between ensuring downstream freedom, maximizing adoption, and retaining some proprietary control.

---

[32] Kapoor et al., *supra* note 28 at 23,082–83.

[33] *See* David Gray Widder et al., *Open (For Business): Big Tech, Concentrated Power, and the Political Economy of Open AI* 2 (Aug. 17, 2023) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4543807.

[34] *See* Chinmayi Sharma, *Tragedy of the Digital Commons*, 101 N.C. L. REV. 1129, 1142–43 (2023).

[35] *Id.*

[36] *Id.* at 1164–65.

At one pole, "copyleft" licenses, such as the GNU General Public License (GPL)[37] enforce downstream openness by imposing restrictive terms that require derivative works to impose the same viral license.[38] In doing so, they prevent users from locking up derivative works behind closed systems; at times, this dampens OSS adoption. At the other, "permissive" licenses like MIT, BSD, and Apache impose minimal restrictions,[39] encouraging maximum OSS adoption and commercialization by requiring little more than attribution to the original OSS project.[40] While the former preserves downstream openness through brute force, the latter can foster openness by inviting more players to contribute to the OSS ecosystem without foregoing the possibility of financial gain.

Between these two poles is a growing class of source-available license configurations. Some make OSS source code visible and readable to users but impose significant restrictions such as prohibiting modification, redistribution, or commercial use without explicit permission from the original developer.[41] This achieves some of openness's value—trust and oversight through transparency—but it limits generative collaborations. Others, such as the edtech company Instructure, straddle both ends of the OSS spectrum a different way: by employing dual-licensing strategies, they can release a copyleft version of a project that contributes to the OSS community while maintaining proprietary version with

---

[37] Richard M. Stallman, *What Is Copyleft?*, *in* FREE SOFTWARE, FREE SOCIETY: SELECTED ESSAYS OF RICHARD M. STALLMAN 91, 91–92 (Joshua Gay ed., 1st ed. 2002); *see also* Michael J. Madison, *Reconstructing the Software License*, 35 LOY. U. CHI. L.J. 275, 283–84 (2003).

[38] *See* David McGowan, *Legal Implications of Open-Source Software*, 2001 U. ILL. L. REV. 241, 254 (2001).

[39] LAWRENCE ROSEN, OPEN SOURCE LICENSING: SOFTWARE FREEDOM AND INTELLECTUAL PROPERTY LAW 69–70 (2005).

[40] *See* Peter Picha & Souhaila Serbout, *On the Adoption of Open Source Software Licensing*, 19 EUROPLOP '24, PROC. 29TH EUR. CONF. PATTERN LANGS. PROG., PEOPLE, AND PRACS. 1, 1–7 (2024), https://doi.org/10.1145/3698322.3698341 (describing the benefits of permissive licenses such as MIT, Apache 2.0, and BSD, including "open[ing] up possibilities for innovative and potentially profitable uses of the software"). Andre Morin, Jennifer Urban & Piotr Sliz, *A Quick Guide to Software Licensing for Scientist-Programmer,* 8 PLOS COMPUT. BIOL. (2012), https://doi.org/10.1371/journal.pcbi.1002598.

[41] *See, e.g.,* Mega Limited, *Mega Limited Code License,* GITHUB, https://github.com/meganz/MEGAsync/blob/master/LICENCE.md (last visited Aug. 1, 2025); *Adopting and Developing BSL Software,* MARIADB, https://mariadb.com/bsl-faq-adopting/ (last visited Aug. 1, 2025); Thomas Claburn, *Redis Has a License to Kill: Open-Source Database Maker Takes Some Code Proprietary,* REGISTER (Aug. 23, 2018); *Licenses,* REDIS, https://redis.io/legal/licenses/ (last visited July 31, 2025).

extended features to earn a profit.[42] These hybrid license configurations foreshadow similar strategies in osAI: reconciling the competing interest between realizing the benefits of openness and preserving profitability.[43]

This dual focus on OSS source code—made available technically and legally—has created a powerful and legible narrative about openness. For example, the contrast between an open operating system like Linux and a closed one like Windows provides a clear case study in how these choices shaped market structures, pricing, and user control.[44] And while other factors shape the software ecosystem—including hardware lock-in, proprietary data formats, and anticompetitive behavior[45]—the focal importance of source code and the clear differences between open and closed licenses has made the OSS model so influential.

Consequently, policymakers latched onto the elegantly simple single-component framework.[46] Source code access became the proxy for procurement

[42] *See Our Open Source Strategy*, INSTRUCTURE, https://www.instructure.com/resources/blog/our-open-source-strategy (last visited Aug. 1, 2025).

[43] *See, e.g.,* Shirin Ghaffary, *Why Meta is Giving Away its Extremely Powerful AI Model*, VOX (July 28, 2023); Bart de Witte, *Case Study: Meta's Strategy for Open-Sourcing Llama: A Detailed Analysis*, HEALTHCARE INNOVATION LETTER (Aug. 5, 2024), https://blog.hippoai.org/metas-strategy-for-open-sourcing-Llama-a-detailed-analysis-hippogram-27/.

[44] *Compare Licensing,* MICROSOFT, https://www.microsoft.com/en-us/licensing (last visited Aug. 1, 2025) (describing Microsoft's complex and restrictive commercial licensing framework, which governs access, use, and distribution through product and enterprise-specific terms), *with* Linux, *GPL-2.0 License*, GITHUB, https://github.com/torvalds/linux/blob/master/LICENSES/preferred/GPL-2.0 (last visited Aug. 1, 2025) (providing a free and permissive license for the Linux kernel that guarantees users the rights to access, use, modify, and distribute the source code).

[45] *See generally* TIM WU, THE MASTER SWITCH: THE RISE AND FALL OF INFORMATION EMPIRES (2010) (describing how the internet has moved from an open, generative space to a closed ecosystem of walled gardens); *see also* JONATHAN ZITTRAIN, FUTURE OF THE INTERNET—AND HOW TO STOP IT (2008) (same).

[46] *See* ROBERT W. HAHN, GOVERNMENT POLICY TOWARD OPEN SOURCE SOFTWARE 4–6 (2002); *see also* CONG. RSCH. SERV., RL32268, INTELLECTUAL PROPERTY, COMPUTER SOFTWARE AND THE OPEN SOURCE MOVEMENT 1, 2, 4, 11–13, 17–18 (2004).

policies that sought transparency and security,[47] antitrust analysis that examines whether closed systems created unfair market advantages,[48] and liability frameworks that treated OSS projects more leniently when assigning responsibility for failures or vulnerabilities.[49] This entrenched the idea that governing technological openness was primarily a matter of governing source code.

The simplicity that offered such clarity for OSS, however, now becomes a liability. As we explain in detail in the next section,[50] osAI systems are not monolithic programs; they are layered systems composed of interdependent components—compute, data, source code, model weights, and more—where source code is just one piece of the puzzle, and often not the most important one. A myopic focus on source code obfuscates both the cascading effects of AI openness decisions and the policy levers available to governments.

2. Beyond Altruism

A second reason the OSS analogy fails when applied to osAI is because the primary actors driving openness in AI, and their motivations, are fundamentally different than in the software context.[51] Therefore, it is incumbent on policymakers to understand the power dynamics in the osAI ecosystem—why certain players may choose to open or close different AI components—to diagnose

---

[47] *See* Robert W. Gomulkiewicz, *Considering a Right to Repair Software,* 37 BERKELEY TECH. L.J. 943, 958–60 (2022); *see also* David S. Evans & Bernard J. Reddy, *Government Preferences for Promoting Open-Source Software: A Solution in Search of a Problem,* 9 MICH. TELECOMM. TECH. L. REV. 313, 315 (2003).

[48] *See, e.g.*, Press Release, U.S. Dep't of Just., Justice Department Sues Google for Monopolizing Digital Advertising Technologies (Feb. 6, 2025), https://www.justice.gov/archives/opa/pr/justice-department-sues-google-monopolizing-digital-advertising-technologies; *see also* Press Release, U.S. Dep't of Just., Justice Department Sues Apple for Monopolizing Smartphone Markets (Feb. 6, 2025), https://www.justice.gov/archives/opa/pr/justice-department-sues-apple-monopolizing-smartphone-markets; United States v. Apple, Inc., No. 24-CV-4055 (JXN)(LDW), 2025 WL 1829127, at *16 (D.N.J. June 30, 2025) ("[T]he Amended complaint alleges Apple maintains a market share of 65 percent in the smartphone market and 70 percent in the performance smartphone market, imposed several barriers to entry, and has engaged in anticompetitive conduct.").

[49] *Compare* Sharma, *supra* note 34, at 1134 (emphasizing that "[o]pen source is not the problem" and arguing open source must be treated as critical public infrastructure that requires government protected intervention) *with* Bryan H. Choi, *Tainted Source Code*, 39 HARV. J.L. & TECH. 1 (forthcoming 2025) (arguing that OSS contributions should not be categorically exempt from liability and proposing a negligence-based framework for harms caused by defective open-source code), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5169060.

[50] *See infra* Part I.B.

[51] *See* Widder et al., *supra* note 28, at 827.

when openness decisions might serve policy goals and how policy levers can account for the incentives that drive them.[52]

The ethical commitment to software openness evolved in a distinct institutional context, one defined by a decentralized community of academics, researchers, hobbyists, corporations, and government entities collaborating on an ecosystem of open protocols and widespread information sharing.[53] Many were driven in part by self-gain, but OSS is unique because at its core, it believes openness serves the public interest.[54] Accordingly, the modern OSS movement responded to the trend of software commercialization[55] by codifying an ethical vision of software freedom: users should have the right to run, study, modify, and share the code they use.[56] By harnessing the engine of volunteer developers and corporate interests alike, the open source movement flourished, yielding technological advancements and improved accountability.[57] Today, cornerstone OSS projects like Linux, Apache, and Python form the backbone of global computing infrastructure.[58]

While many hope that openness in the AI ecosystem will catalyze similar collaborative innovation and build public trust, osAI is driven by a very different, much more centralized, set of stakeholders.[59] These actors are motivated by different, often competing, equities—some prioritize economic growth, others national security or public accountability—and not all are driven by the altruistic motivations that gave rise to the OSS movement.

Meta's release of Llama model weights, for instance, was not just a nod to open science but a calculated play to gain multiple strategic advantages.[60] By making its models widely and freely available, Meta encourages a global community of developers to build tools and applications on its platform, effectively

---

[52] *See* Widder et al., *supra* note 33, at 4.

[53] *See* Sharma, *supra* note 34, at 1152 (describing the groundbreaking development of the Linux OS and Apache web server).

[54] *See id.* at 1148–49.

[55] *See* Martin Campbell-Kelly, *Not All Bad: An Historical Perspective on Software Patents*, 11 MICH. TELECOMM. TECH. L. REV. 191, 211–212 (2005).

[56] *See* Richard Stallman, *Why Open Source Misses the Point of Free Software*, GNU OPERATING SYS., https://www.gnu.org/philosophy/open-source-misses-the-point.en.html (last visited Aug. 1, 2025).

[57] Sharma, *supra* note 34 at 1148–52.

[58] *See* Jesus M. Gonzalez-Barahona, *A Brief History of Free, Open Source Software and Its Communities*, 54 COMPUT. 75, 79 (2021); *see also* Paul Jansen, *TIOBE Index for July 2025*, TIOBE, https://www.tiobe.com/tiobe-index/ (last visited Aug. 1, 2025) (reporting Python as the most popular among all programming languages, open and closed).

[59] *See* Widder et al., *supra* note 33, at 4.

[60] *See* Ghaffary, *supra* note 43; de Witte, *supra* note 43 *But see* Eli Tan, *Meta's New Superintelligence Lab is Discussing Major A.I. Strategy Changes,* N.Y. TIMES (July 14, 2025).

crowdsourcing innovation and making the Llama architecture a de facto industry standard. This strategy seeks to commoditize the model layer of the AI stack, creating a competitive disadvantage for rivals like OpenAI who charge for access, while also serving as a powerful recruiting tool by allowing its researchers to be more public about their advances.

Indeed, recent developments in the OSS movement's own trajectory caution against assuming openness is always altruistic. What began as a grassroots, volunteer-driven movement is now largely powered by corporate developers—more than half of OSS contributions today come from employees at firms like Google, Microsoft, and IBM.[61] Many of these companies have learned to strategically embrace openness not as a value, but as a vehicle: for shaping ecosystems, deflecting regulation, and entrenching market position, often sacrificing true openness values at the altar of pursuing greater market power.[62] For example, Google openly released Android, its mobile operating system, which drove app innovation, and therefore consumer adoption, without lowering the barriers to entry for innovating on the operating system itself—yielding more customers, no new competitors, and greater lock-in.[63] Other companies skirt the licenses enforcing downstream openness by capitalizing on what they learned from OSS projects they used to build parallel systems they can profit from.[64] As we turn to osAI, these dynamics remind us that openness can be both a public good and a competitive tactic.

3. Beyond the Developer

Even a focus on the strategic motivations of developers is too narrow. The third failure of the OSS analogy is that it cannot account for the sprawling ecosystem of diverse stakeholders who control different, and often more critical, layers of the AI technology stack. OSS developers are the ones writing source code; they have the power to open source it. The entities that build osAI, on the other hand, are distributed across a wide range of powerful players, each building different components of the AI stack. Thus many hands shape osAI's differential openness.

Some of the largest and most powerful players in the AI ecosystem are the companies that design, produce, and manage the specialized computational hardware AI requires.[65] Their core incentive is control over supply and demand. Companies that design the specialized hardware, the often overseas

[61] Sharma, *supra* note 34 at 1150.

[62] *See id.* at 1153–54.

[63] *See* Widder et al., *supra* note 33, at 13.

[64] *See* Stephen Shankland, *Google Gets Web Allies by Letting Outsiders Help Build Chrome's Foundation*, CNET (Nov. 30, 2020).

[65] *See* Widder et al., *supra* note 33, at 7–8.

manufacturers that produce them, and the service providers that make them available to downstream consumers (leading AI labs, startups, and researchers alike), seek to maximize return on high-capital investments by controlling access to compute.[66] Essentially, they dictate who gets to experiment with AI—and who gets priced out.[67] For actors with an economic interest in scarcity, they lack the incentive to change the status quo.

Beyond hardware, AI learns from data, and so it relies heavily on those who generate data (whether willingly or not) and those who transform raw data into usable datasets. Because data quality, quantity, and type significantly impact model capabilities and biases, data providers who create, curate, and commodify datasets wield enormous influence in the AI ecosystem.[68] These entities range from companies that happen to own vast archives of proprietary data they can sell to AI companies, such as news organizations like the New York Times or academic publishers like Elsevier, to those who collect extensive user interaction data, like social media platforms and e-commerce sites.[69] Other companies specifically specialize in collecting, refining, and labeling data—some even generate synthetic data specifically for AI development, including firms like Scale AI.[70] Data providers can choose to filter out harmful content, avoid copyrighted material, and rectify biases—or not. They either opt to release datasets openly or restrict access through licensing or fees,[71] functionally determining, like hardware providers, who can play in the sandbox. They are driven by the desire to minimize legal risk while maximizing the ability to profit from a scarce resource.[72]

The most visible members of the AI ecosystem are frontier model developers, including organizations such as OpenAI (ChatGPT), Google DeepMind

---

[66] *See* Will Henshall, *Big Tech Companies Were Investors in Smaller AI Labs. Now They're Rivals*, TIME (May 13, 2024).

[67] *See* Kevin M.K. Fodoup, *Promoting Access to Innovative AI*, 7 J.L. & TECH. TEX. 1, 15 (2024) ("The field has seen a continuous trend toward gigantic models, meaning that only the most resourced corporations can internally develop state-of-the art innovative capabilities.").

[68] *See* Kevin Roose, *The Data That Powers A.I. Is Disappearing Fast*, N.Y. TIMES (July 19, 2024) ("[W]idespread data restrictions may pose a threat to A.I. companies, which need a steady supply of high-quality data to keep their models fresh and up-to-date.").

[69] *See* Diana Kwon, *Publishers Are Selling Papers to Train AIs—And Making Millions of Dollars*, 636 NATURE 529, 530 (2024), https://doi.org/10.1038/d41586-024-04018-5; *see also* Annie Palmer, *Amazon AI Deal with New York Times Brings the Paper's Content to Alexa*, CNBC (May 29, 2025, 09:54 ET).

[70] *See Data Engine*, SCALE, https://scale.com/data-engine (last visited Aug. 1, 2025).

[71] *See* Kevin Paul & Anna Tong, *Inside Big Tech's Underground Race to Buy AI Training Data*, REUTERS (Apr. 5, 2024); *see also* Narechania & Sitaraman, *supra* note 28, at 122–23.

[72] *See* Widder et al., *supra* note 33, at 8–10.

(Gemini and Gemma), Meta (Llama), Anthropic (Claude), and xAI (Grok). These companies are building the most powerful AI systems in the ecosystem today—the foundation models that everyone else builds upon—which is resource and expertise intensive.[73] Their incentive calculus blends public positioning with competitive strategy.[74] By releasing Llama, Meta sought to capture developer mindshare and ecosystem control.[75] In contrast, OpenAI's closed approach may have helped protect its lead in fine-tuning and enterprise deployment—at least, so far.[76] Showcasing the market domination these behemoths seek, many frontier companies are proactively investing in owning—and therefore controlling—the hardware AI depends on, vertically integrating the stack and doubling their capacity to influence openness in the ecosystem.[77]

Once a model is trained, downstream developers adapt foundational AI models to specific use cases for public consumption—often taking the form of applications.[78] Their incentives revolve around defensibility, differentiation, and user trust. For instance, a healthcare technology firm might fine-tune an osAI model on proprietary medical data, creating a powerful diagnostic tool that they can share with others or withhold for themselves. Many downstream developers may make some parts of an osAI system open while restricting others, to protect proprietary assets and minimize exposure.[79] (See Appendix A for a range of examples.) This hybrid openness, reminiscent of similar arrangements in OSS licenses, is becoming more common as companies leverage the benefits of open source collaboration while keeping their own competitive edge.

---

[73] *See* MARKUS ANDERLJUNG ET AL., FRONTIER AI REGULATION: MANAGING EMERGING RISKS TO PUBLIC SAFETY (2023), https://cdn.governance.ai/Frontier_AI_Regulation_Managing_Emerging_Risks.pdf.

[74] *See* Widder et al., *supra* note 33, at 4.

[75] *See* Ghaffary, *supra* note 43; de Witte, *supra* note 43.

[76] *See* Narechania & Sitaraman, *supra* note 28, at 120.

[77] *See, e.g.*, Amin Vadat, *Introducing Google Axion Processors, Our New Arm-based CPUs*, GOOGLE CLOUD: BLOG (Apr. 9, 2024), https://cloud.google.com/blog/products/compute/introducing-googles-new-arm-based-cpu; *see also* Jake Siegel, *With a Systems Approach to Chips, Microsoft Aims to Tailor Everything 'from Silicon to Service' to Meet AI Demand*, MICROSOFT (Nov. 15, 2023), https://news.microsoft.com/source/features/ai/in-house-chips-silicon-to-service-to-meet-ai-demand; *AWS and NVIDIA*, AMAZON WEB SERVS., https://aws.amazon.com/nvidia/ (last visited Aug. 1, 2025); Katie Paul & Krystal Hu, *Exclusive: Meta Begins Testing Its First In-House AI Training Chip*, REUTERS (Mar. 11, 2025, at 09:37 ET).

[78] *See* Narechania & Sitaraman, *supra* note 28, at 119 (2024) ("Downstream developers need to access the foundation models for fine-tuning and use in a particular application . . . ").

[79] *See* Appendix A; *see also* Bommasani et al., *supra* note 28, at 151.

Finally, governance stakeholders—including regulators, standards bodies, and civil society groups—are trying to balance different policy goals such as safety, innovation, democratic control, and national security.[80] These groups don't control the technology directly, but their decisions around data governance, operational transparency, ethical constraints, and public accountability shape the legal and ethical landscape in which AI operates.[81] How much transparency should be required? Should companies be allowed to release powerful models with no oversight? Should there be safeguards against AI monopolization? Their efficacy in this role, however, can be impacted by political strife, resource constraints, expertise scarcity, and the bludgeoning of corporate lobbyists.[82]

Unsurprisingly, each of the different players in the AI ecosystem operates under different incentives, including as it relates to openness. The technological nature of software is not as reliant on hardware, data sources, and expertise as AI, and policymakers can get away with a singular focus on developers, who control source code openness.[83] AI stakeholders are diverse and diffuse, which means effective osAI policy must accommodate its technical and sociocultural distinctions from open source software.

---

[80] *See infra* Part II.

[81] *See* Press Release, Bureau of Indus. and Sec., Biden-Harris Administration Announces Regulatory Framework for the Responsible Diffusion of Advanced Artificial Intelligence Technology (Jan. 13, 2025), https://www.bis.gov/press-release/biden-harris-administration-announces-regulatory-framework-responsible-diffusion-advanced-artificial; *see also* Matt O'Brien, *White House Says No Need to Restrict Open Source AI, For Now*, PBS (July 30, 2024, at 14:02 ET); EXEC. OFF. OF THE PRESIDENT, NATIONAL CYBERSECURITY STRATEGY 1, 21 (2023), https://bidenwhitehouse.archives.gov/wp-content/uploads/2023/03/National-Cybersecurity-Strategy-2023.pdf; Ben Brooks, *California's AI Reforms Scare All Developers, Not Just Big Tech*, TECH POL'Y PRESS (Aug. 23, 2024), https://www.techpolicy.press/californias-ai-reforms-scare-all-developers-not-just-big-tech/; Zuzanna Warso & Maximilian Gahntz, *How the EU AI Act Can Increase Transparency Around AI Training Data*, TECH POL'Y PRESS (Dec. 9, 2024), https://www.techpolicy.press/how-the-eu-ai-act-can-increase-transparency-around-ai-training-data/; Pablo Chavez, *Sovereign AI in a Hybrid World: National Strategies and Policy Responses*, LAWFARE (Nov. 7, 2024), https://www.lawfaremedia.org/article/sovereign-ai-in-a-hybrid-world--national-strategies-and-policy-responses.

[82] *See* Chinmayi Sharma, *AI's Hippocratic Oath*, 102 WASH. U. L. REV. 1101, 1137–41 (2025).

[83] *See* Tejas N. Narechania, *Machine Learning as Natural Monopoly,* 107 IOWA L. REV. 1543, 1569–88 (2022).

B. Unbundling AI

The simplicity that once made OSS governance tractable becomes a liability with osAI. To grapple with osAI's complexity, we introduce our core analytical innovation: the concept of "unbundling" AI systems into constituent components and fitting them within a framework of differential openness. This section dissects the technical layers of an AI system—compute, data, source code, model weights, system prompts, operational control and records, and the human element—identifying how openness manifests differently across them. These components are not static; they interact with each other both in development and post-deployment through feedback loops.[84] Information gathered by one component can inform the future development of another, and by controlling more than one component, a single entity can multiply its market advantage. By disaggregating AI into a legible taxonomy and explaining the concept of differential openness, we lay the foundation for demonstrating how policymakers can develop more targeted interventions: identifying which component is most relevant to their goals, who controls it, and how open that component should be.[85] Figure 1 visualizes these relationship:

[84] *Understanding and Managing the AI Life Cycle*, U.S. GEN. SERVS. ADMIN. CTRS. EXCELLENCE, https://coe.gsa.gov/coe/ai-guide-for-government/understanding-managing-ai-lifecycle/ (last visited Aug. 1, 2025) (explaining that model development continues after deployment based on information gathered).

[85] We borrow the concept of "unbundling" from the telecommunications sector, where regulators forced dominant incumbents to give competitors access to essential network infrastructure components on fair terms, spurring competition and innovation without requiring complete duplication of physical assets. *See, e.g.*, Press Release, Fed. Commc'ns Comm'n, FCC Promotes Local Telecommunications Competition: Adopts Rules on Unbundling of Network Elements (Sep. 15, 1999), https://transition.fcc.gov/Bureaus/Common_Carrier/News_Releases/1999/nrcc9066.html (explaining the adoption of "unbundling" rules for telecommunications). While often cited as a failure in the U.S., the analogy of disaggregating systems into component parts and targeting regulation to specific components is powerful in the AI space. *See general* Gordon Klein & Julia Wendel, *The Impact of Local Loop Unbundling Revisited*, in 25TH EUR. REG'L CONF. INT'L TELECOMM. SOC'Y (June 2014), https://www.econstor.eu/bitstream/10419/101416/1/795233892.pdf (describing the impact of unbundling requirements).

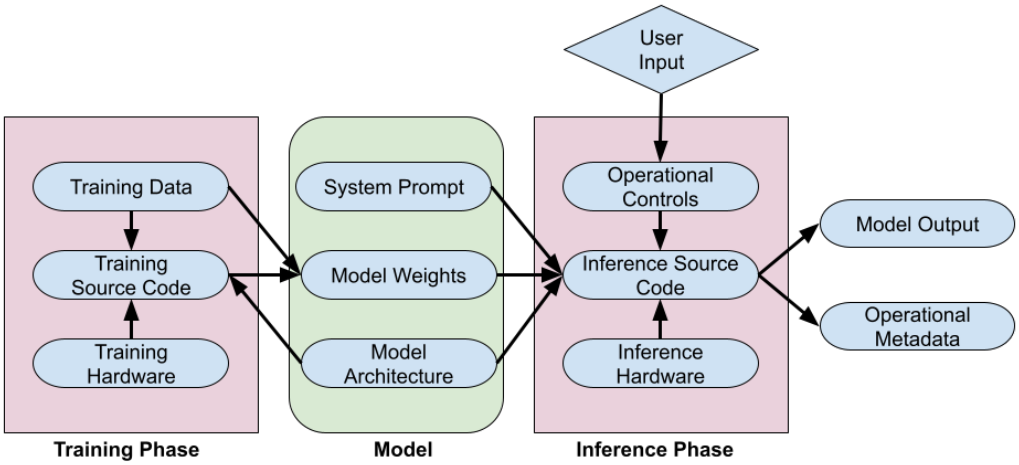

Figure 1 illustrates an AI model's lifecycle, which comprises a *training phase* and an *inference phase*. In the training phase, *training source code* running on *training hardware* (normally a GPU or TPU) processes *training data* to create *model weights*. These weights, which are numerical parameters, combine with the *model architecture* (the model's underlying structure) and a text-based *system prompt* to form the complete *model*. During the subsequent inference phase, a *user input* is first managed by *operational controls*, such as safety filters, before being processed by *inference source code* on *inference hardware* (similar to training hardware). This code applies the multi-component model to the input to generate a *model output* as well as *operational metadata* like usage logs and audit trails. Note that safety features can also be part of the training process or be applied to the model's final output.

1. Compute

At the foundation of AI systems lies compute—the physical hardware that powers both AI training and "inference," the process by which a trained model produces output based on user input. Unlike traditional software, which can run on almost any computer, cutting-edge "frontier" AI models require massive, specialized, and often cost-prohibitive hardware like graphical processing units (GPUs) and tensor processing units (TPUs).[86] This reality has created a "hardware lottery," where even the best theoretical advances rely on the happenstance of available computation to have practical impact.[87] The hardware layer creates a significant bottleneck in the AI supply chain, with the market

[86] Narechania, *supra* note 83, at 1569–88.

[87] *See* Sara Hooker, *The Hardware Lottery*, 64 COMMC'NS ACM 58, 60–63 (2021).

dominated by a few key firms: Nvidia and Google for chip design, Taiwan Semiconductor Manufacturing Company (TSMC) for manufacturing, and the Dutch company ASML for essential lithography equipment.[88] However, other parts of the AI stack are investing in their own compute to avoid reliance on these titans.[89] These emerging competitors, however, are few and no less powerful.

Compute's openness is multifaceted. First is the challenge of opening physical access. Chips and data centers are constrained by high costs, limited supply, exclusionary vendor relationships, and national export controls.[90] In some rare cases, only one company has access to core compute infrastructure: Google is the only entity that owns TPUs; others are forced to rent from it.[91] For the many that cannot afford access to compute directly, they are reliant on renting compute—cloud services—from the handful of entities that own it.[92] Second, hardware architecture is usually closed: designs for specialized chips like those from Google and Nvidia are proprietary, preventing independent replication or modification.[93] Third, software stacks create lock-in: to use Nvidia's market-leading hardware, developers are functionally required to use its proprietary CUDA programming interface, making it difficult to move to another platform when their systems are engineered around one.[94]

The RISC-V movement, which is trying to enable independent chip design by opening chip blueprints,[95] can simultaneously challenge the duopoly that

---

[88] *See* Narechania & Sitaraman, *supra* note 28, at 112–13.

[89] *See supra* note 77 and accompany text.

[90] *See* Hervé Legenvre & Erkko Autio, *NVIDIA: Harnessing Open Innovation to Promote User Lock-In*, EUR. BUS. REV. (Nov. 26, 2024), https://www.europeanbusinessreview.com/nvidia-harnessing-open-innovation-to-promote-user-lock-in/.

[91] *See* Kenrick Cai & Krystal Hu, *Exclusive: OpenAI Taps Google in Unprecedented Cloud Deal Despite AI Rivalry, Sources Say,* REUTERS (June 11, 2025).

[92] *See* Cade Metz, Karen Weise & Mike Isaac, *Nvidia's Big Tech Rivals Put Their Own A.I. Chips on the Table*, N.Y. TIMES (Jan. 29, 2024) (identifying the tech companies large enough to invest in their own compute hardware); *see also* Erin Griffith, *The Desperate Hunt for the A.I. Boom's Most Indispensable Prize*, N.Y. TIMES (Aug. 16, 2023), (explaining that most companies rent compute power from cloud services to avoid building their own data centers).

[93] *See* Widder et al., *supra* note 33, at 7.

[94] *See CUDA Compatibility*, NVIDIA, https://docs.nvidia.com/deploy/cuda-compatibility/ (last visited Aug. 1, 2025) (explaining that use of CUDA requires an Nvidia driver). *But see* Emre Çıtak, *Nvidia to Bring CUDA Platform Support to the RISC-V*, DATACONOMY MEDIA (July 21, 2025), https://riscv.org/ecosystem-news/2025/07/nvidia-to-bring-cuda-platform-support-to-the-risc-v/.

[95] *See* riscv-isa-manual, *License*, GITHUB (last visited Aug. 1, 2025), https://github.com/riscv/riscv-isa-manual/blob/main/LICENSE ("This document is released under a Creative Commons Attribution 4.0 International License."); Che Pan &

controls this hardware layer, while building increased reliance on the centralized entities controlling compute's software layer.[96]

Fully open compute infrastructure would entail something akin to a public option: universally accessible hardware components with low to no barriers to entry.[97] However, this reality is unlikely to manifest. And while some initiatives—like decentralized compute networks[98] or research cloud credits[99]—aim to expand access, these are generally partial, selective, and funnel market power back to incumbents. So in most cases only those with significant capital can afford the high-end GPUs or cloud services necessary to meaningfully experiment with large-scale AI today.[100]

2. Data

Data is the fuel that powers AI and is among the most contested components in the ecosystem. The capabilities of traditional software are defined by its source code; AI, however, "learns" from extensive data sets[101] that range from raw, unstructured data (social media posts or biotech sensor readings) to carefully curated training data (labeled images) to highly specialized datasets

---

Brenda Goh, *China to Publish Policy to Boost RISC-V Chip Use Nationwide, Sources Say*, REUTERS (Mar. 4, 2025) (reporting on China's support for the RISC-V movement, which is attempting to enable independent chip design by opening chip blueprints).

[96] *See* Çıtak, supra note 94.

[97] Eleanor Shearer, Matt Davies & Mathew Lawrence, *The Role of Public Compute*, ADA LOVELACE INSTITUTE (April 24, 2024), https://www.adalovelaceinstitute.org/blog/the-role-of-public-compute/.

[98] *See* Will Knight, *These Startups Are Building Advanced AI Models Without Data Centers,* WIRED (Apr. 30, 2025) (explaining that while startups are exploring decentralized frontier model development, most AI companies require "huge quantities of compute concentrated inside data centers stuffed with advanced GPUs.").

[99] *See Apply for Google Cloud Research Credits,* GOOGLE CLOUD, https://edu.google.com/intl/ALL_us/programs/credits/research/ (last visited Aug. 1, 2025); *see also AWS Cloud Credit for Research,* AMAZON WEB SERVS., https://aws.amazon.com/government-education/research-and-technical-computing/cloud-credit-for-research/ (last visited Aug. 1, 2025).

[100] *See* Jai Vipra & Sarah Myers West, *Computational Power and AI*, AI NOW INST. (Sep. 27, 2023), https://ainowinstitute.org/publications/compute-and-ai (explaining that compute is scarce, and therefore a bottleneck for AI development, which amplifies the market power of the few companies providing it).

[101] *See* Christopher S. Yoo, *Beyond Algorithmic Disclosure For AI*, 25 COLUM. SCI. & TECH. L. REV. 314, 318–21 (2024).

for refining (or fine-tuning) model performance in particular domains or for specific tasks (medical diagnoses or legal document analysis).[102]

Privacy concerns, copyright laws, and competitive advantages all determine what data is shared, who can use it, and under what terms. In this way, they also dictate how open the models built on them can truly be.[103] Consequently, AI training, finetuning, and testing data is far more contested and controlled than traditional source code, adding yet another layer of complexity to AI's differential openness.[104] For example, OpenAI is the poster child for the liability exposure that emerges when datasets are visible, facing a slew of lawsuits accusing it of training models of hordes of copyrighted material.[105] Even models known for their openness, such as Mistral 7B, still withhold access to, or even information about, datasets, citing competitive pressures.[106]

Some datasets, such as Common Crawl—a massive, publicly archived crawl of the web—are fully open, allowing unrestricted access to raw training materials.[107] Others are partially open, meaning they are available under certain conditions, such as being licensed for research use but restricted for commercial applications.[108] Many datasets, however, remain completely closed, either through proprietary licenses or strict contractual agreements to protect privacy, competitive advantage, or intellectual property rights.[109] Proprietary

---

[102] *See* Jenny Quang, *Does Training AI Violate Copyright Law?,* 36 BERKELEY TECH. L.J. 1407, 1411 (2021).

[103] *See* Yoo, *supra* note 101, at 321–24; *see also* Katie Knibbs, *Meta Secretly Trained Its AI on a Notorious Piracy Database, Newly Unredacted Court Docs Reveal*, WIRED (Jan. 9, 2025); TINA SADEK ET. AL., RAND, ARTIFICIAL INTELLIGENCE IMPACTS ON PRIVACY LAW (2024), https://www.rand.org/pubs/research_reports/RRA3243-2.html.

[104] *See* Mehtab Khan & Alex Hanna, *The Subjects and Stages of AI Dataset Development: A Framework For Dataset Accountability,* 19 OHIO ST. TECH. L.J. 171, 179 (2023) ("The vast majority of the training data from [frontier] models are private."); *see also* Roose, *supra* note 68.

[105] *See* Kyle Jahner, *OpenAI Sued by New Set of Authors Over Training Data Copyrights*, BLOOMBERG L. (July 2, 2025).

[106] *See* Comment from Arthurmensch to mistralai/Mistral-7B-v0.1, HUGGING FACE (Oct. 12, 2023), https://huggingface.co/mistralai/Mistral-7B-v0.1/discussions/8 ("Unfortunately we're unable to share details about the training and datasets . . . due to the highly competitive nature of the field.").

[107] *See Our Mission,* COMMON CRAWL, https://commoncrawl.org/mission (last visited Aug. 1, 2025) ("Small startups or even individuals can now access high quality crawl data that was previously only available to large search engine corporations.").

[108] For example, ImageNet, a large database of labeled images, is available free for non-commercial use. IMAGENET, https://www.image-net.org/ (last visited Aug. 1, 2025).

[109] *See* Fodouop, supra note 67, at 15 (2024); *see also* Sydney Rouser, *Unfair Competition in the Creative Industries: The Impact of AI Scraping*, 16 TENN. J.L. & POL'Y 134, 144 (2024).

medical, financial, or corporate datasets, for example, are often off-limits to all but the companies that own them.[110] Use of this data, no matter how open or closed, imports the risks embedded in them.

Importantly, true data openness is about more than just access to a dataset; it is also determined by the transparency of its curation.[111] Meaningful access is often dependent on the upstream models, such as smaller neural networks, that are used for filtering, classification, or ranking—models that are themselves often neither open nor auditable.[112] This upstream opacity means that even a notionally "open" dataset may be shaped by hidden selection biases, creating a black box at the very start of the AI pipeline.[113]

3. Source Code

While not the sole determinant of functionality, source code remains a critical component of the AI stack and a factor in osAI's differential openness. First, "inference code" shapes the potential capabilities and functions the system can perform by defining its architecture, the structure of how the model processes input data into predictions.[114] And second, "training code" details how the model learns from its training data.[115] Openness in inference code facilitates visibility and allows stakeholders to understand and assess a model's theoretical capabilities, while openness of training code enables replication, verification, and potentially improvement of the original results.[116]

---

[110] *See, e.g.*, Isabelle Rose I Alberto et al., *The Impact of Commercial Health Datasets on Medical Research and Health-Care Algorithms*, 5 LANCET DIGIT HEALTH e288 (2023).

[111] *See* Widder et al., *supra* note 33, at 9.

[112] *See, e.g.*, Catherine Arnett et al., *Toxicity of the Commons: Curating Open-Source Pre-Training Data* (Nov. 18, 2024) (unpublished manuscript), https://arxiv.org/abs/2410.22587 (example of open classifier for toxicity filtering).

[113] *See* Stefan Baack et al., *Towards Best Practices for Open Datasets for LLM Training*, MOZILLA Found. (Jan. 13, 2025), https://arxiv.org/abs/2501.08365.

[114] *See Open Source AI Definition: Version 1.0,* OPEN SOURCE INITIATIVE, https://opensource.org/ai/open-source-ai-definition (last visited Aug. 1, 2025) (distinguishing code used to guide training from code used for model architecture).

[115] *See id.*

[116] *See* Andrew D. Mitchell et al., *AI Regulation and the Protection of Source Code*, 31 INT'L J.L. & INFO. TECH. 283, 286–87 (2023), https://doi.org/10.1093/ijlit/eaad026.

Openness of AI source code generally aligns with the established spectrum of OSS licenses employing permissive licenses, such as MIT, for maximal accessibility,[117] or copyleft licenses, like GPL, to ensure ongoing openness of derivative works.[118] But many leading systems—such as OpenAI's GPT series, Anthropic's Claude, and Google's Gemini—keep source code entirely proprietary, preventing external developers and researchers from inspecting or experimenting on the model architecture's inner workings.[119]

4. Model Weights

But even complete openness in source code—covering both architecture and training scripts—does not fully define or predict an AI system's behavior. That largely depends on the next crucial piece: model weights. An AI system emerges from training with a static "model" that determines *how* the model makes predictions based on inputs. Model weights—billions (and soon-to-be trillions) of numerical parameters—store this learning as compressed knowledge.[120] They shape everything from a model's writing style to its ability to recognize images. While AI source code establishes model architecture—like a building's blueprint—it is the model's weights that dictate what the model knows and how well it performs, much like furniture dictates how a building is actually used.

Model weight' openness exists along a spectrum of licenses, similar to source code, that either enables or restricts transparency, experimentation, and reuse.[121] At one extreme are fully closed models, like Anthropic's Claude and OpenAI's ChatGPT.[122] These models do not release model weights and only allow access to chat interfaces or APIs—restrictive protocols for machine-to-machine communication. Desires to preserve a competitive edge or curb misuse

---

[117] *See, e.g.,* deepseek-ai, *DeepSeek-R1/LICENSE*, HUGGING FACE, https://huggingface.co/deepseek-ai/DeepSeek-R1/blob/main/LICENSE (last visited Aug. 1, 2025).

[118] choosealicense.com, *licenses/markdown/agpl-3.0.md*, HUGGING FACE, https://huggingface.co/choosealicense (last visited Aug. 1, 2025).

[119] *See* Mitchell et al., *supra* note 116, at 284 (2023).

[120] Giorgio Franceschelli et al., *Training Foundation Models as Data Compression: On Information, Model Weights and Copyright Law* at 1 (Mar. 12, 2025) (unpublished manuscript), https://arxiv.org/abs/2407.13493.

[121] *See* Eiras et al., supra note 28; *AI Model Weights*, NAT'L TELECOMM. & INFO. ADMIN., *supra* note 27.

[122] Sharon Goldman, *Why Anthropic and OpenAI Are Obsessed with Securing LLM Model Weights*, VENTUREBEAT (Dec. 15, 2023, 5:00 AM), https://venturebeat.com/ai/why-anthropic-and-openai-are-obsessed-with-securing-llm-model-weights/

risks end up preventing independent scrutiny or external modification.[123] Some companies, like OpenAI, make their model weights accessible to select research institutes, such as the UK AI Safety Institute; but even this accessibility is strictly controlled.[124] On the other end are models, like Deepseek's R1, that release weights under copyleft or permissive licenses.[125] In between are models such as Meta's Llama, that release weights but limit their use.[126]

The release of model weights under permissive licenses is a necessary condition for a model to be truly open, but not a sufficient one. Many of the largest models labeled "open source" are merely open-weight, and often minimally so. The most notable example is Meta's Llama, the most popular "open source" American model[127]—though it would be more accurate to call it "non-proprietary" for two reasons: (1) many of its components, like training data and training code, are not public; and (2) openly released components, like model weights, come under a highly restrictive license—neither copyleft nor permissive.[128] These constraints sharply limit the use, redistribution, and innovation that genuine openness is meant to support.

This exploitation of differential openness has been termed "open-washing,"[129] which risks misleading the public and policymakers into believing these

---

[123] *See, e.g., Reasoning Models*, OPENAI, https://platform.openai.com/docs/guides/reasoning (last visited Aug. 1, 2025) ("[W]e don't expose the raw reasoning tokens emitted by the model"). OpenAI also generally lacks information on accessing weights.

[124] *See Pre-Deployment Evaluation of OpenAI's o1 Model*, AI SEC. INSTITUTE (Dec. 18, 2024), https://www.aisi.gov.uk/work/pre-deployment-evaluation-of-openais-o1-model; *See also GPT-4o System Card*, OPENAI (Aug. 8, 2024), https://openai.com/index/gpt-4o-system-card/ (discussing that "OpenAI worked with more than 100 external red teamers, speaking a total of 45 different languages, and representing geographic backgrounds of 29 different countries.").

[125] deepseek-ai, *DeepSeek-R1*, HUGGING FACE, https://huggingface.co/deepseek-ai/DeepSeek-R1 (last visited Aug. 1, 2025) ("DeepSeek-R1 series support commercial use, allow for any modifications and derivative works, including, but not limited to, distillation for training other LLMs.").

[126] *See* Widder et al., *supra* note 28.

[127] *See, e.g.*, *WebDev Arena*, LMARENA, https://lmarena.ai/leaderboard/webdev (last visited Aug. 1, 2025) (showing Llama as the leading American non-proprietary model for web development).

[128] *See* Michael Nolan, *Llama and ChatGPT Are Not Open-Source*, IEEE SPECTRUM (July 27, 2023), https://spectrum.ieee.org/open-source-llm-not-open; Parth Nobel, Alan Z. Rozenshtein & Chinmayi Sharma, *Open-Access AI: Lessons From Open-Source Software*, LAWFARE (Oct. 25, 2024), https://www.lawfaremedia.org/article/open-access-ai--lessons-from-open-source-software.

[129] *See* Sarah Kessler, *Openwashing*, N.Y. TIMES (MAY 17, 2024); Liesenfeld & Dingemanse, *supra* note 28, at 1774 ("Our survey yields 40 text generators that are described as 'open source' or 'open' . . . . We also find a large number of systems (roughly the bottom third)

systems are more open than they are, while allowing companies to benefit from a reputational boost and, in some cases, regulatory benefits of choosing "openness." For example, the EU AI Act privileges certain models that make model weights open under vaguely defined open licenses without requiring disclosure of critical elements like training data or fine-tuning methods.[130]

5. System Prompts

Another increasingly important layer in deployed AI systems is system prompts: the foundational instructions or configuration strings that set the behavior of the model at runtime. Separate from a model's learned knowledge, and unbeknownst to users, system prompts govern tone, style, boundaries, and behavioral defaults, significantly shaping outputs without changing the underlying model weights or architecture. It accomplishes this by appending content to the user's own input before submitting the query to the model—sometimes, it reframes the query and other times, it overrides the user's explicit request. [131] For instance, a system prompt might instruct a model: "You are a professional research assistant. Your tone must be neutral and objective. You must refuse any request for personal opinions or political commentary."

Because these prompts are often crafted through iterative experimentation and internal alignment processes, they can lead to unintended, sometimes downright abhorrent, outcomes. Google's misguided system prompts in an early version of Gemini led to the widely derided generation of troubling, historically inaccurate images—such as multi-racial Nazis.[132] And when Grok's system prompts were updated to "not shy away from making claims which are politically incorrect, as long as they are well substantiated," it dubbed itself MechaHitler within days.[133]

Openness of system prompts also exists on a spectrum. Some models, such as Anthropic's Claude and xAI's Grok, publish their system prompts, allowing

---

that make only model weights available but share little to no detail about other parts of their system.").

[130] *See supra* note 27 and accompanying text; *see also* Liesenfeld & Dingemanse, *supra* note 28.

[131] *See* Anna Neumann et al., *Position Is Power: System Prompts as a Mechanism of Bias in Large Language Models (LLMs)*, PROCS. 2025 ACM CONF. ON FAIRNESS, ACCOUNTABILITY & TRANSPARENCY (FACCT '25) 573 (2025), https://doi.org/10.1145/3715275.3732038.

[132] *See* James Grimmelmann, Blake E. Reid & Alan Z. Rozenshtein, *Generative Baseline Hell and the Regulation of Machine-Learning Foundation Models*, LAWFARE (May 8, 2024), https://www.lawfaremedia.org/article/generative-baseline-hell-and-the-regulation-of-machine-learning-foundation-models.

[133] Lisa Hagen, Huo Jingnan & Audrey Nguyen, *Elon Musk's AI Chatbot, Grok, Started Calling Itself "MechaHitler,"* NPR (July 9, 2025).

users and the public to understand, copy, and modify the behavioral guardrails guiding the model.[134] Others, including many commercial offerings like OpenAI's GPT models and Google's Gemini, treat system prompts as proprietary—hiding them from public view to retain control, limit gaming, or obscure alignment choices.[135] While seemingly minor, these hidden prompts play an outsized role in shaping downstream applications and safety, making their disclosure an increasingly relevant axis of differential openness.

6. Operational Control and Records

Beyond the model itself, an AI system's behavior is also influenced by operational controls, which are layers on top of the model that further enhance how a model behaves in the real world, and operational data, which documents everything from the system's development process to its real world interactions.

Operational controls facilitate model behavior as they transition from development to real-world deployment.[136] They include content filters, moderation tools, and risk management layers.[137] Their importance and specific design are highly dependent on the AI model's intended use. For instance, a chatbot designed for medical diagnosis will incorporate strict safety protocols to prevent faulty information, whereas one for casual companionship may employ less rigorous safeguards.

Complementing these active controls are operational records—model cards, data cards, technical reports, and system design papers—that provide behind-the-scenes static documentation of how an AI system is built, trained, tested, and deployed.[138] These records do not directly interact with AI models

---

[134] *See, e.g., System Prompts*, ANTHROPIC, https://docs.anthropic.com/en/release-notes/system-prompts (last visited Aug. 1, 2025); *see also* xai-org, *Grok-Prompts*, GITHUB, https://github.com/xai-org/grok-prompts (last visited Aug. 1, 2025).

[135] *See* Kyle Jahner, *Trade Secrets Law Is Awkward Fit in AI Prompt-Hacking Lawsuit*, BLOOMBERG L. (Mar. 14, 2025).

[136] *See* Rosario Cammarota et al., *Trustworthy AI Inference Systems: An Industry Research View* (Feb. 10, 2023) (unpublished manuscript), http://arxiv.org/abs/2008.04449.

[137] Daniel Maggen, *Predict and Suspect: The Emergence of Artificial Legal Meaning*, 23 N.C. J.L. & TECH. 67, 98 (2021) (discussing how various algorithms in AI systems take on "triage responsibilities" such as processing, classifying, and filtering information); Pranav Gade et.al., *Cheaply Removing Safety Fine-tuning from Llama 2-Chat 13B* at 1 (Mar. 28, 2024) (unpublished manuscript), https://arxiv.org/abs/2311.00117; *Our Approach to AI Safety*, OPENAI (Apr. 5, 2023), https://openai.com/index/our-approach-to-ai-safety/.

[138] KASIA CHMIELINSKI ET AL., SHORENSTEIN CTR. ON MEDIA, POL. & PUB. POL'Y AT HARV. KENNEDY SCH., THE CLEAR DOCUMENTATION FRAMEWORK FOR AI TRANSPARENCY: RECOMMENDA-

but are instead references for internal and external stakeholders to guide model iteration. When openly available, they facilitate third party experimentation, since AI systems are too complex to understand by reading source code alone. Even with traditional OSS, the accessibility of a system for inspection, experimentation, and reuse relies on effective documentation.[139] Similarly, operational records also allow external researchers, policymakers, and users to better understand and evaluate the AI system's intentions, limitations, and risks.[140] Conversely, vague or missing documentation impairs external oversight and makes it harder to diagnose or prevent harm, such as bias or flaws.

Beyond such static documentation, operational records include dynamic metadata generated with every system interaction. This data includes monitoring logs, audit trails, and performance metrics—elements that reveal how an AI model behaves in the wild.[141] They are essential for accountability, risk detection, and safety improvement based on insight from real world interactions; predeployment testing can only go so far. Yet, despite their value, this type of metadata is rarely, if ever, included in traditional software openness frameworks.[142] Without transparency in this area, even an osAI system that is fully

---

TIONS FOR PRACTITIONERS & CONTEXT FOR POLICYMAKERS (2024), https://shorensteincenter.org/clear-documentation-framework-ai-transparency-recommendations-practitioners-context-policymakers/.

[139] *See* DALIA TOPELSON RITVO ET AL., HARV. L. SCH. BERKMAN KLEIN CTR. FOR INTERNET & SOC'Y, ORGANIZATION & STRUCTURE OF OPEN SOURCE SOFTWARE DEVELOPMENT INITIATIVES 22 (2024), https://clinic.cyber.harvard.edu/wp-content/uploads/2017/03/2017-03_governance-FINAL.pdf ("Though it is sometimes overlooked, the history of the open source movement shows us that the projects that defined their corporate structure and governance practices early and concretely set themselves up for success.").

[140] *See, e.g.*, CHMIELINSKI ET AL., *supra* note 138, at 2 ("Documentation of datasets, models, and AI systems is crucial and serves several purposes, including: (1) Supporting responsible development and use, as well as mitigation of downstream harms, by providing transparency into the design, attributes, intended use, and shortcomings of datasets, models, and AI systems; (2) Motivating dataset, model, or AI system creators and curators to reflect on the choices they make; and (3) Facilitating dataset, model, and AI system evaluation and auditing.").

[141] *See* Dominik Kreuzberger et al., *Machine Learning Operations (MLOps): Overview, Definition, and Architecture*, 11 IEEE ACCESS 31,866 (2023), https://doi.org/10.1109/ACCESS.2023.3262138.

[142] The Open Source Definition makes no mention of usage records. OPEN SOURCE INITIATIVE, *supra* note 114. While some OSS that collects telemetry publishes it, it is generally "anonymized and aggregated to ensure user privacy" and not part of the software release. *Firefox Public Data Report*, FIREFOX, https://data.firefox.com/ (last visited Aug. 1, 2025).

open in terms of code and weights can still be a black box when it comes to real-world deployment and actual user experience.[143]

7. The Human Layer

AI systems are not just technical artifacts—they are built by human hands and minds. The talent, expertise, and institutions that train and organize AI professionals are fundamental components of the AI stack and determinants of osAI's differential openness.[144] Like source code or training data, this human layer can be more or less open, and its degree of openness shapes innovation, concentration, and accountability across the ecosystem.

The state of the human layer is not merely a workforce issue—it is a governance mechanism. The answer to the question of who gets to contribute, switch jobs, start labs, or critique dominant approaches determines whose values are embedded in AI. In OSS, labor is relatively open: contributors from anywhere, with any degree of formal training, can submit code and build reputational capital.[145] But the AI ecosystem—where physical presence (especially in San Francisco and Silicon Valley) is still crucial[146]—is characterized by three key forms of closure: (1) restrictive pipelines that limit access to a diverse talent pool, (2) institutional constraints that inhibit professional mobility, and (3) corporate controls that suppress the diffusion of critical knowledge.

First, access to the field is constrained at both the domestic and international levels. Domestically, STEM education and talent pipelines fail to produce enough qualified individuals to meet demand, and the domestic talent that exists is unevenly distributed by race, gender, geography, and institutional prestige.[147] This lack of diversity means that communities most impacted by AI often have no voice in its design. The ecosystem's reliance on foreign talent faces

---

[143] *See* Lisa Lee, *What is Metadata in AI?*, SALESFORCE (May 17, 2024), https://www.salesforce.com/blog/what-is-metadata/; Everton Gomede, *The Importance of Metrics and Metadata in Data Observability*, MEDIUM (Mar. 5, 2024), https://pub.aimind.so/the-importance-of-metrics-and-metadata-in-data-observability-f6d571fd2269.

[144] Gordon Hansen, *Immigration and Regional Specialization in AI*, *in* ROBOTS AND AI 180, 180 (Lili Yan Ing & Gene M Grossman eds. 2022).

[145] *See What Is Open Source?*, IBM (Feb. 5, 2025), https://www.ibm.com/think/topics/open-source.

[146] Kristina McElheran et al., *AI Adoption in America: Who, What, and Where*, 33 J. ECON. & MGMT. STRATEGY 375, 376 (2024), https://doi.org/10.1111/jems.12576.

[147] *See* Darrell M. West, *Improving Workforce Development and STEM Education to Preserve America's Innovation Edge,* BROOKINGS INST. (July 26, 2023), https://www.brookings.edu/articles/improving-workforce-development-and-stem-education-to-preserve-americas-innovation-edge/ ("[A]ccording to a Deloitte study, there are fewer than 100,000

even steeper barriers. As Nvidia CEO Jensen Huang has noted, half of the world's top AI researchers are Chinese, highlighting the global nature of expertise.[148] Yet unlike in OSS, where anyone can contribute via platforms like GitHub, meaningful participation in frontier AI development typically requires being hired by a dominant firm and moving to America. Consequently, as demand outpaces domestic supply, visa bottlenecks and restrictive immigration policies have become critical chokepoints,[149] failing to retain international students and attract foreign researchers, thus locking out the very talent the U.S. needs to lead.[150]

Second, for those who do gain access, institutional constraints then limit mobility. Noncompete agreements have long blocked researchers from switching companies or launching startups,[151] locking expertise inside a handful of dominant firms and slowing the diffusion of knowledge.[152] This is not theory; it's been proven. Many credit Silicon Valley's meteoric growth to California's

---

U.S. graduates with electrical engineering and computer science degrees each year, which is below the number that will be required in the coming decade."); *See* NAT'L SCI. BD., THE STEM LABOR FORCE: SCIENTISTS ENGINEER, AND SKILLED TECHNICAL WORKERS 16 (2024), https://ncses.nsf.gov/pubs/nsb20245/representation-of-demographic-groups-in-stem ("[T]he proportion of men in STEM occupations remained higher than that of women in 2011, 2016, and 2021 . . . STEM workers were disproportionately Asian and White").

[148] Daniel Howley, *Nvidia CEO Jensen Huang Sounds the Alarm on Open-Source AI*, YAHOO! FINANCE (May 29, 2024).

[149] *See* Remco Zwetsloot et al., *Skilled and Mobile: Survey Evidence of AI Researchers' Immigration Preferences*, *in* AIES '21: PROC'S 2021 AAAI/ACM CONF. ON AI, ETHICS & SOC'Y 1050, 1052 (July 2021), https://doi.org/10.1145/3461702.3462617 ("AI PhDs who chose to leave the U.S. were likely to cite immigration-related concerns (23%) and the U.S. immigration system (33%) as highly relevant."); Sharifymoghaddam v Blinken No. 23-CV-1472-RCL, 2024 WL 939991, at *6 (D.D.C. Mar. 5, 2024) (acknowledging "delays in visa processing [had] interfer[ed] with [plaintiff's] career progression and ability to contribute to artificial intelligence research in the United States" but denying visa application).

[150] *See* Helen Toner, Director of Strategy, Ctr. for Sec. & Emerging Tech., Walsh Sch. of Foreign Serv., Geo. Univ., Testimony Before the U.S.-China Econ. & Sec. Rev. Comm'n, Technology, Trade, and Military-Civil Fusion: China's Pursuit of Artificial Intelligence, New Materials, and New Energy 5 (June 7, 2019), https://cset.georgetown.edu/publication/technology-trade-and-military-civil-fusion-chinas-pursuit-of-artificial-intelligence/ ("More than 85% of Chinese and Indian students in U.S. computer science and engineering PhD programs state that they intend to stay after graduation [but are often unable to.]").

[151] Caitlyn Harrington, *Innovation-Killing Noncompete Agreements Are Finally Dying,* WIRED (Dec. 4, 2023) ("35% of people working in computer- and math-related vocations work under noncompetes—the highest share in any industry.")

[152] Owen Hughes, *DeepMind Scraps Noncompete Clauses Amid Ongoing AI Talent Wars,* TECH REPUBLIC (July 12, 2023), https://www.techrepublic.com/article/news-deepmind-noncompete-clauses-ai-talent-wars ("DeepMind is using aggressive noncompete clauses… Some senior researchers are subject to a full year of paid 'garden leave.'").

aggressive stance against enforcing noncompete clauses.[153] By contrast, an open labor market empowers experts to join competitors or found new ventures that may better align with their ethical or scientific values.

Third, knowledge itself is trapped in institutional silos, weakening the broader diffusion of expertise. Restrictions on academic moonlighting,[154] opaque clearance processes for publications,[155] and overreliance on proprietary data further trap knowledge within institutional silos,[156] Experts within companies are on a short leash when it comes to academic contributions and outside academics are simultaneously locked out of the systems they seek to research. Adherence to open science principles—transparent peer review, public dissemination of results, and reproducibility requirements—is a direct countermeasure, enhancing labor openness by reducing barriers to entry for contributors outside the monoculture of elite and profit-driven institutions.[157]

These layers of closure are starkly illustrated by the labor-intensive process of Reinforcement Learning from Human Feedback (RLHF). This work, essential for model alignment and safety, is performed by armies of human annotators—often low-paid contractors in non-English-speaking countries like Kenya, India, or the Phillipines—who rate and label millions of model outputs, making the answer keys from which AI learns.[158] These workers are invisible in governance frameworks: their contributions are essential, yet they have no voice in system design or deployment. Openness in this context would demand transparency about who performs this difficult work, ensure fair compensation, and mandate their inclusion in feedback and oversight processes. Treating these

---

[153] Mike McPhate, *California Today: Silicon Valley's Secret Sauce*, N.Y. TIMES (MAY 19, 2017).

[154] *See e.g., Consulting and Other Outside Professional Activities by Members of the Academic Council and University Medical Line Faculty*, STAN. UNIV., https://doresearch.stanford.edu/policies/research-policy-handbook/conflicts-commitment-and-interest/consulting-and-other-outside-professional-activities-members-academic-council-and-medical-center-line-faculty (last visited Aug. 1, 2025) ("The maximum number of Consulting days permissible for a member of the Academic Council or the University Medical Line Faculty on a full-time appointment is 13 days per academic quarter.").

[155] Thomas Klebel et al., *Peer Review and Preprint Policies Are Unclear at Most Major Journals*, PLOS ONE, Oct. 21, 2020, at 3, https://doi.org/10.1371/journal.pone.0239518.

[156] Anat Lior, *Private and Academic AI Collaboration: Opportunities and Challenges to Open Science in the US*, 11 J. OPEN ACCESS TO L., no. 2, 2024, at 3, https://doi.org/10.63567/8wfpme67.

[157] *See* Erin C McKiernan et al., *Point of View: How Open Science Helps Researchers Succeed*, ELIFE SCIENCES, July 7, 2016, at 3–9, https://doi.org/10.7554/eLife.16800.

[158] Sarah Hastings-Woodhouse, *Reinforcement Learning from Human Feedback (RHLF): A Simple Explainer*, BLUEDOT IMPACT (May 15, 2025), https://bluedot.org/blog/rlhf-explainer.

workers as integral contributors, not interchangeable cogs, strengthens both accountability and system integrity.

Ultimately, just as open data or public weights can democratize technology, so can open labor practices diffuse expertise, reduce capture, and foster accountability. Conversely, just as closed infrastructure or proprietary models entrench power, labor constraints can centralize control over the direction and pace of AI innovation. Differential openness for osAI must therefore treat labor not as a background condition, but as a core, governable component of the AI stack.

* * *

Part I has provided the essential analytical toolkit of differential openness. By rejecting the flawed "open versus closed" binary and unbundling AI into its seven core components, we have seen why OSS governance models fall short: every osAI system represents a unique configuration of component-level choices, each existing on its own spectrum of openness. With this granular understanding in place, Part II will apply this framework to the central challenge of osAI policy: promoting and navigating the tradeoffs between safety, innovation, democratic control, and national security.

II. The Value of AI Openness

The concept of openness, as inherited from traditional software development, carries unhelpful baggage into the discourse on artificial intelligence. Beyond the flawed analogy of applying a software-centric model to a complex AI stack, the most distorting piece of this legacy is the assumption that openness is, by definition, an intrinsic good.[159]

This Part challenges that premise. We argue AI openness is more accurately characterized as an *instrumental* good. It is not inherently good or bad; it is a powerful policy tool whose desirability is entirely contingent on which components are opened, to what degree, and to what end.[160] Consequently, there is no universal answer to how open or closed any particular AI system should be. Instead, the differential openness of osAI must be calibrated with precision to achieve specific policy objectives.

---

[159] *See* Richard Stallman, *Free Software Is Even More Important Now*, GNU OPERATING SYS., https://www.gnu.org/philosophy/free-software-even-more-important.html (last visited Aug. 1, 2025).

[160] *See* JON BATEMAN ET AL., CARNEGIE ENDOWMENT INT'L PEACE, BEYOND OPEN V. CLOSED: EMERGING CONSENSUS AND KEY QUESTIONS FOR FOUNDATION AI MODEL GOVERNANCE 4 (2024), https://carnegieendowment.org/research/2024/07/beyond-open-vs-closed-emerging-consensus-and-key-questions-for-foundation-ai-model-governance.

To build this case, this Part uses the unbundling framework to systematically evaluate how differential openness affects four key policy objectives. Sections A through D analyze each objective in turn—public safety, innovation and economic growth, democratic control, and national security—demonstrating how, for each goal, openness functions as a double-edged sword, creating both profound benefits and acute risks. Finally, Section E synthesizes this analysis, moving from the tensions *within* each goal to the unavoidable tradeoffs *between* them. This reveals the complex balancing act policymakers face, where every decision to open or close a component of the AI stack necessarily prioritizes one value over another.

A. Safety

AI already shapes critical, sometimes life-or-death, decisions. It detects cancer in medical scans,[161] approves or denies mortgages,[162] flags security threats,[163] and determines what billions of people see online.[164] Its failures do not unfold in the abstract—they manifest in hospitals,[165] courtrooms,[166] financial markets,[167] and battlefields.[168] When AI goes wrong, people lose jobs, homes, access to critical services, and sometimes even their lives.

---

[161] Rebecca C. Fitzgerald et al., *The Future of Early Cancer Detection*, 28 NAT. MED. 666, 666-67 (2022), https:// 10.1038/s41591-022-01746-x.

[162] *See, e.g.,* Elijah Clark, *Rocket Mortgage's AI Technology: The Future of Mortgage Lending*, FORBES (Apr. 15, 2024); Kori Hale, *A.I. Bias Causes 80% of Black Mortgage Applicants to Be Denied*, FORBES (Sep. 2, 2021).

[163] Aya H. Salem et al., *Advancing Cybersecurity: A Comprehensive Review of AI-Driven Detection Techniques*, 11 J. BIG DATA 1, 16 (2024), https://doi.org/10.1186/s40537-024-00957-y.

[164] *See* Sang Ah Kim*, Social Media Algorithms: Why You See What You See,* GEO. L. TECH. REV. 147, 150–51 (2017).

[165] Laure Wynants et al., *Prediction Models for Diagnosis and Prognosis of Covid-19: Systematic Review and Critical Appraisal*, BMJ, Apr. 7, 2020, at 1, https://doi.org/10.1136/bmj.m1328 (reviewing 731 AI systems purporting to diagnose or predict prognosis for COVID patients and finding five prognostic tools "showed adequate predictive performance in studies at low risk of bias")

[166] Jess Weatherbed, *Errors Found in US Judge's Withdrawn Decision Stink of AI*, VERGE (JULY 25, 2025, 5:30 AM CDT), https://www.theverge.com/news/713653/judge-withdraws-cormedix-case-ai-citation-errors.

[167] *See* Paolo Giudici & Emanuela Raffinetti, *SAFE Artificial Intelligence in Finance*, 56 FIN. RSCH. LETTERS, 1, 2–3, 12 (2023), https://doi.org/10.1016/j.frl.2023.104088.

[168] Michael Biesecker, Sam Mednick & Garance Burke*, As Israel Uses US-Made AI Models in War, Concerns Arise About Tech's Role in Who Lives and Who Dies*, AP NEWS (Feb. 18, 2025).

Governing AI safety requires addressing three distinct challenges: (1) ensuring models are accurate (producing correct and unbiased outputs) and reliable (doing so consistently); (2) maintaining alignment so that their outputs do not cause harm through misuse or unintended behavior; and (3) enabling auditability through sufficient transparency and control to diagnose and remedy failures.

Differential openness is the primary tool for meeting these challenges, but it is a double-edged sword. While the open source software (OSS) ethos that "given enough eyeballs, all bugs are shallow"[169] applies with equal force to osAI, the same transparency that enables public auditing can also be exploited by malicious actors. The central question for safety, then, is not whether to make AI open or closed wholesale, but where targeted openness can meaningfully reduce harm without creating unacceptable risks.[170]

1. Benefits

To see how this works in practice, consider an autonomous vehicle that strikes a Black woman in a wheelchair crossing a street at night. To understand what went wrong—and to prevent it from happening again—investigators need visibility into the entire AI stack. Was the model trained on diverse and representative data? Did it struggle in low-light conditions? Was the decision logic flawed, the hardware malfunctioning, or the system manipulated?

Access to the training data would allow external researchers the capacity to assess whether the dataset the AV system was trained on included enough nighttime scenarios or sufficient representation of people with different skin tones or disabilities. Without that access, it is impossible to know whether the model was ever given the chance to learn how to recognize someone like the victim.[171] Separately, data openness often uncovers systemic flaws, such as routine scraping of illegal content or the lack of responsible filtering that internal

---

[169] ERIC S. RAYMOND, THE CATHEDRAL AND THE BAZAAR 30 (1999).

[170] Alondra Nelson et al., Comment Letter to Dep't of Com. on Request for Comment on Dual Use Foundation Artificial Intelligence Models With Widely Available Model Weights 2–3 (Mar. 27, 2024), https://hai-production.s3.amazonaws.com/files/2024-03/Response-NTIA-RFC-Open-Foundation-Models.pdf.

[171] *See* Jack Cable, & Aeva Black, *With Open Source Artificial Intelligence, Don't Forget the Lessons of Open Source Software*, CYBERSECURITY & INFRASTRUCTURE SEC. AGENCY BLOG (July 29, 2024), https://www.cisa.gov/news-events/news/open-source-artificial-intelligence-dont-forget-lessons-open-source-software.

teams might miss or ignore.[172] For example, researchers have found ample child sexual abuse material in image datasets.[173]

Moving deeper into the model's architecture, transparent model weights and system prompts would allow independent experts to investigate the AI's decision-making processes directly. They could identify if the system's logic deprioritizes pedestrians with certain demographic characteristics or in certain conditions—for example, wearing certain types of clothing, walking a pet, or standing in poor lighting. Looking under the hood at the model's decision-making processes can also enable crucial research into broader problems like AI hallucinations, where large language models generate false but highly plausible-seeming information.[174] Insight into safety prompts can expose misguided or malicious instructions appended to use inputs that can lead to unsafe outputs.[175]

Once a system is deployed, operational records become equally critical.[176] Transparent metadata—including detailed logs and decision trails—surface critical failures, particularly important in high-stakes applications like medical diagnostics, hiring systems, or autonomous vehicles.[177] Logs might pinpoint exactly when and why the system failed—whether it detected the car-accident victim at all, whether it misclassified her as a shadow or background object, or whether it delayed braking. This data is also essential for detecting adversarial interference, such as malicious tampering.

---

[172] *See* KEVIN KLYMAN ET AL., HUMAN-CENTERED ARTIFICIAL INTELLIGENCE AT STAN. UNIV., SAFEGUARDING THIRD-PARTY RESEARCH (2025), https://hai.stanford.edu/assets/files/hai-policy-brief-safeguarding-third-party-ai-research.pdf.

[173] David Thiel, *Investigation Finds AI Image Generation Models Trained on Child Abuse*, STAN. UNIV. CYBER POL'Y CTR. (Dec. 20, 2023), https://cyber.fsi.stanford.edu/news/investigation-finds-ai-image-generation-models-trained-child-abuse.

[174] *See, e.g.*, Sebastian Farquhar et al., *Detecting Hallucinations in Large Language Models Using Semantic Entropy*, 630 NATURE 625 (2024), https://doi.org/10.1038/s41586-024-07421-0.

[175] NAT'L TELECOMM. & INFO. ADMIN., *supra* note 25, at 17–18.

[176] NAT'L INST. OF STANDARDS & TECH., DEP'T OF COM., ARTIFICIAL INTELLIGENCE RISK MANAGEMENT FRAMEWORK (AI RMF 1.0) 15–16, 35 (2023), https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf.

[177] U.S. DEP'T HOMELAND SEC. & AI SAFETY & SEC. BOARD, ROLES AND RESPONSIBILITIES FRAMEWORK FOR ARTIFICIAL INTELLIGENCE IN CRITICAL INFRASTRUCTURE 2, 19 (2024), https://www.dhs.gov/sites/default/files/2024-11/24_1114_dhs_ai-roles-and-responsibilities-framework-508.pdf (highlighting the importance of maintaining operational records for critical infrastructure uses of AI).

Finally, access to operational controls and compute enables proactive, not merely reactive, safety work.[178] Open access to a model's bias-detection algorithms or adversarial testing frameworks allows independent researchers to perform "red-teaming"—simulating a range of edge cases—such as pedestrians with different skin tones, clothing styles, or body types, to surface blind spots *before* they cause harm. This requires computational resources to be available beyond companies and a handful of well-funded safety labs,[179] as opening safety tools without providing the hardware to run them is an empty gesture.[180]

At the root of all of this is the human component. When human workers are—from data labelers in RLHF to internal engineers—properly trained, protected, empowered, and embedded in transparent workflows, they can serve as an early warning system—flagging unsafe outputs, flawed incentives, or rushed deployments at their source. These individuals are best positioned to uncover issues, especially when they stem from secretive development processes.[181]

2. Costs

But for all the ways that openness can strengthen safety, it also introduces serious and often irrevocable risks. The same transparency that enables oversight can be exploited by malicious actors, and the diffusion of powerful tools containing hidden flaws, biases, or security gaps can amplify harm, even when users are well-meaning. This trade-off is not abstract; it manifests at each layer of the AI stack.

---

[178] *See* Vinita Fordham, Allie Diehl & David Caswell, *Securing Government Against Adversarial AI*, DELOITTE (Apr. 11, 2023), https://www2.deloitte.com/us/en/insights/industry/public-sector/adversarial-ai.html.

[179] Manish Parashar, *Enabling Responsible Artificial Intelligence Research and Development Through the Democratization of Advanced Cyberinfrastructure*, HARV. DATA SCI. REV. (Special Issue), no. 4, Apr. 2024, https://hdsr.mitpress.mit.edu/pub/fysjbutp/release/2.

[180] *See* Nicol Turner Lee, Paul Resnick & Genie Barton, *Algorithmic Bias Detection and Mitigation: Best Practices and Policies to Reduce Consumer Harms,* BROOKINGS INST. (May 22, 2019), https://www.brookings.edu/articles/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/.

[181] Sharma, *supra* note 82, at 1157–58.

The most acute risk lies in the dissemination of model weights.[182] Once released, an AI model's core knowledge cannot be recalled.[183] It remains indefinitely available to be repurposed, modified, and potentially weaponized.[184] Any embedded flaw—a bias, unsafe instruction, or alignment gap—can be replicated into perpetuity. There is no practical mechanism to compel bad actors to cease its use or alert all well-meaning users of emergent harms. Consequently, open-weight models are routinely stripped of safeguards to generate extremist propaganda, nonconsensual deepfakes, and automated social engineering scams.[185]

Open data creates parallel dangers. The release of training sets containing private health records, intimate photos, or copyrighted material —even in the name of transparency—can constitute a massive violation of privacy and property rights.[186] The use of scraped social media content and personal images has already led to real-world harms, from non-consensual deepfake pornography[187] to government targeting and surveillance of dissidents.[188]

Finally, operational controls and records, while essential to oversight, can be weaponized. Transparent safety benchmarks and bias detection tools can be reverse engineered by adversaries to learn how to evade them. For example, deepfake creators can use open detection models as a training tool, fine-tuning

---

[182] P'SHIP ON AI, THE PARTNERSHIP ON AI RESPONSE TO THE NTIA REQUEST FOR COMMENT (RFC) ON DUAL USE FOUNDATION ARTIFICIAL INTELLIGENCE MODELS WITH WIDELY AVAILABLE MODEL WEIGHTS 6 (Apr. 2024), https://partnershiponai.org/wp-content/uploads/2024/04/PAI-Response-to-NTIA-RFC-Open-Foundation-Models.pdf.

[183] Kapoor et al., *supra* note 28 at 3 ("Once the weights for a foundation model are made widely available, little recourse exists for the foundation model developer to rescind access.").

[184] Edd Gent, *Protesters Decry Meta's "Irreversible Proliferation" of AI*, IEEE SPECTRUM (Oct. 6, 2023), https://spectrum.ieee.org/meta-ai; *see also* Gade, *supra* note 137.

[185] *See On Open-Weight Foundation Models*, FTC: TECHNOLOGY BLOG (July 10, 2024), https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2024/07/open-weights-foundation-models; NAT'L TELECOMM. & INFO. ADMIN., *supra* note 25, at 24–26.

[186] *See, e.g.*, Jahner, *supra* note 105.

[187] *See* David Evan Harris, *Open-Source AI Is Uniquely Dangerous: But the Regulations That Could Rein It In Would Benefit All of AI*, IEEE SPECTRUM (Jan. 12, 2024), https://spectrum.ieee.org/open-source-ai-2666932122.

[188] Jay Stanley, *Machine Surveillance is Being Super-Charged by Large AI Models*, AM. C.L. UNION (Mar. 21, 2025), https://www.aclu.org/news/privacy-technology/machine-surveillance-is-being-super-charged-by-large-ai-models; Darrell M. West, *How AI Can Enable Public Surveillance*, BROOKINGS INST. (Apr. 15, 2025), https://www.brookings.edu/articles/how-ai-can-enable-public-surveillance/.

their outputs until they beat the very systems designed to stop them.[189] Likewise, public operational data—from error logs to performance metrics—can provide a detailed roadmap to an AI system's blind spots, allowing cyberattackers to design exploits that target known weaknesses.[190] Even operational records stored in the name of safety can, if opened without sufficient anonymization, create so-called "radioactive" piles[191] of sensitive personal data.[192] Openness here accelerates the learning curve for attackers to overtake safety advances.

B. Innovation and Economic Growth

Artificial intelligence is both a powerful driver of innovation and a critical battleground for economic growth. Unlike traditional software, which can often be replicated with minimal resources, frontier AI requires immense computational power, vast datasets, and optimized model architectures.[193] Control over these core components is currently highly centralized, with a few dominant firms creating significant risks of monopolistic behavior, including discriminatory pricing, vendor lock-in, and "kill zones" that stifle new entrants. As scholars like Tejas Narechania and Ganesh Sitaraman have explored, this concentration threatens to stagnate the very innovation that AI promises to deliver.[194]

In this high-stakes environment, well-calibrated differential openness for osAI can serve as a potent anti-concentration tool, countering monopolistic tendencies by democratizing access to cutting-edge technology. However, its effectiveness is not guaranteed. While opening certain components of the AI stack can accelerate progress and broaden participation, strategic, partial openness can also be subverted into a tool for incumbent entrenchment—a

---

[189] In the OSS setting, spam detection is often not implemented as open source. *See, e.g.*, Jim O'Leary, *Improving First Impressions on Signal*, SIGNAL (Nov. 1, 2021), https://signal.org/blog/keeping-spam-off-signal/.("If we put this code on the Internet alongside everything else, spammers would just read it and adjust their tactics to gain an advantage in the cat-and-mouse game of keeping spam off the network.").

[190] *See* Scott Ikeda, *NIST Warns AI Developers of "Poisoning" Methods, Cyber Threats That Reverse Engineer Models*, CPO MAG. (Jan. 15, 2024), https://www.cpomagazine.com/cyber-security/nist-warns-ai-developers-of-poisoning-methods-cyber-threats-that-reverse-engineer-models/.

[191] Trey Herr, *Protecting Society from Radioactive Data*, TECH POL'Y PRESS (July 21, 2025), https://www.techpolicy.press/protecting-society-from-radioactive-data/.

[192] Kevin Bankston, *In ChatGPT Case, Order to Retain All Chats Threatens User Privacy*, CTR. FOR DEMOCRACY & TECH. (June 25, 2025), https://cdt.org/insights/in-chatgpt-case-order-to-retain-all-chats-threatens-user-privacy/.

[193] Tim Hwang, *Computational Power and the Social Impact of Artificial Intelligence* at 10 (Mar. 23, 2018) (unpublished manuscript), https://arxiv.org/abs/1803.08971.

[194] Narechania & Sitaraman, *supra* note 28, 128–37.

form of "open-washing" that creates an illusion of accessibility while keeping true market power consolidated.[195]

1. Benefits

Optimal configurations of osAI's differential openness can dramatically lower barriers to entry and accelerate the pace of technological progress. By reducing redundancy and fostering collaboration, they enable new entrants to build upon existing advances without the prohibitive cost of developing foundational models from scratch.

This pro-competitive effect has been demonstrated by the impact of open-weight models like Meta's Llama, Stability AI's Stable Diffusion, Alibaba's Qwen3-Coder, and EleutherAI's GPT-Neo. These releases have empowered a global community of researchers, startups, and independent developers to create domain-specific applications in fields such as biomedical AI, climate modeling, and materials science—all without requiring billions of dollars in training costs.[196] This decentralization is further supported by osAI development frameworks like TensorFlow and PyTorch[197] and publicly available datasets such as Common Crawl[198] and the Pile,[199] which provide a shared foundation for innovation that is distributed widely rather than siloed within a few dominant firms.

---

[195] *See supra* note 129 and accompanying text.

[196] Niklas Muennighoff et al., *s1: Simple Test-Time Scaling* (Mar. 1, 2025) (unpublished manuscript), https://arxiv.org/abs/2501.19393 (mathematical reasoning experiments dependent on open-weight Qwen and open-traces in Gemini); Shrey Pandit et al., *MedHallu: A Comprehensive Benchmark for Detecting Medical Hallucinations in Large Language Models* (Feb. 20, 2025) (unpublished manuscript), https://arxiv.org/abs/2502.14302 (study of medical hallucinations on open and closed models); Jimeng Shi et al., *Deep Learning and Foundation Models for Weather Prediction: A Survey* (Jan. 12, 2025) (unpublished manuscript) https://arxiv.org/abs/2501.06907 (see section 3 for a list of weather modeling examples); Yingheng Tang et al., *MatterChat: A Multi-Modal LLM for Material Science* (Apr. 26, 2025) (unpublished manuscript), https://arxiv.org/abs/2502.13107 (material science LLM based on pretrained models).

[197] *See* Cade Metz, *Google Just Open Sourced TensorFlow, Its Artificial Intelligence Engine*, Wired (Nov. 9, 2015); Adam Paszke et al., *PyTorch: An Imperative Style, High-Performance Deep Learning Library*, *in* PROCS 33RD INT'L CONF. NEURAL INFO. PROCESSING SYSTEMS (NEURIPS 2019) (2019), https://doi.org/10.5555/3454287.3455008; Akshay Agrawal et al., *TensorFlow Eager: A Multi-Stage, Python-Embedded DSL for Machine Learning*, in PROCS. MACH. LEARNING & SYS. 1 (MLSYS 2019) (2019), https://arxiv.org/abs/1903.01855.

[198] Common Crawl is available under a limited license: *Terms of Use*, COMMON CRAWL, https://commoncrawl.org/terms-of-use (last updated Mar. 7, 2024).

[199] Leo Gao et al., *The Pile: An 800GB Dataset of Diverse Text for Language Modeling* (Dec. 31, 2020) (unpublished manuscript), https://arxiv.org/abs/2101.00027.

Crucially, this form of openness also spreads the *use* of AI, which is itself a powerful driver of innovation and economic growth. Because open-weight models can be downloaded and run on local hardware,[200] users bypass the pay-per-use fees of controlled API endpoints.[201] This empowers a much broader base of individuals and researchers to experiment with and integrate powerful AI into novel applications, fostering a more dynamic, ground-up form of economic growth that cannot be achieved through controlled platforms alone.[202]

2. Costs

However, as Narechania and Sitaraman wisely caution, we must reject the "false promise" that openness "will completely address the problems with an unregulated AI oligopoly."[203] While opening model weights or source code can spur experimentation, this alone does not ensure broad competition if other critical parts of the ecosystem remain closed. As we've emphasized, AI development is not just about access to code; it is also about who controls the surrounding infrastructure that makes AI usable and scalable.[204] If control over compute, proprietary data, deployment pathways, and expert labor remain tightly controlled, then models that only open weights or source code merely offer the illusion of an open marketplace.

This is where openness can be subverted into a strategy for incumbent entrenchment—a form of "open-washing" where companies claim the reputational benefits of openness while withholding key components.[205] Meta's high-profile Llama release is a prime example: it made a self-interested business decision while positioning itself as a champion of openness.[206] While Meta's motivation to crowd out competitor proprietary systems may seem pro-competitive, its ultimate goal is not to dismantle the oligopolistic nature of the industry, but rather to reinforce its own position within it.

---

[200] ANNA HERMANSEN & CAILEAN OSBORNE, LINUX FOUND., THE ECONOMIC AND WORKFORCE IMPACTS OF OPEN SOURCE AI: INSIGHTS FROM INDUSTRY, ACADEMIA, AND OPEN SOURCE RESEARCH PUBLICATIONS (2025), https://www.linuxfoundation.org/research/economic-impacts-of-open-source-ai.

[201] *Pricing,* OPENAI, https://platform.openai.com/docs/pricing (last visited Aug. 1, 2025) (showing how OpenAI charges per API use).

[202] *See* Robert Wolfe et al., *Laboratory-Scale AI: Open-Weight Models Are Competitive with ChatGPT Even in Low-Resource Settings*, *in* 2024 ACM CONF. ON FAIRNESS, ACCOUNTABILITY & TRANSPARENCY 2–3 (2024), http://doi.org/10.1145/3630106.3658966.

[203] Narechania & Sitaraman, *supra* note 28, at 153–54.

[204] *See* supra Part I.B; *see also* Nobel, Rozenshtein & Sharma, *supra* note 128.

[205] Liesenfeld & Dingemanse, *supra* note 28.

[206] *See* Mike Isaac, *How A.I. Made Mark Zuckerberg Popular Again in Silicon Valley*, N.Y. TIMES (May 29, 2024).

The strategic value of a splashy "open source" release is often found not in what is shared, but in the critical components that are held back. Control over four core areas—compute infrastructure, deployment access, proprietary training data, and labor—allows incumbents to retain real market power even as they gesture toward openness.

First, compute. Deploying advanced AI, open or not, requires specialized GPUs and cloud-scale infrastructure that are functionally inaccessible to startups or researchers without vast financial resources. A legal-tech or biomedical startup may fine-tune an open-weight model, but it cannot compete if it cannot afford the cloud resources to deploy it at scale—resources often controlled by the very firms developing the models. This forces smaller players into dependency on dominant providers like Amazon Web Services (AWS), Microsoft Azure, and Google Cloud, reinforcing the incumbents' market position.

Second, deployment access. Making an AI model open is not the same as making it usable. Deploying an AI system in a real-world product—for example, in a legal-research app or a medical-imaging tool—requires cloud hosting and interface tools, which are often controlled by the same few companies. If developers are forced to rely on these gatekeepers, they risk getting locked into a particular company's ecosystem—subject to its pricing, policies, and terms of use. While open standards like Anthropic's widely adopted Model Context Protocol (MCP) promise interoperability that would make it easier to swap one model for another,[207] their effectiveness depends on continued adoption by dominant firms who have historically used such standards to drive adoption of their services before reverting to restrictive, proprietary interfaces once their market position is secure.[208]

Third, data. Training data remains one of the most valuable and least open parts of the AI stack. While models like Llama release weights, they do not release the massive proprietary datasets used to train them.[209] Since a model's performance depends heavily on what it's trained on, incumbents who hoard

---

[207] *Introducing the Model Context Protocol*, ANTHROPIC (Nov. 25, 2024), https://www.anthropic.com/news/model-context-protocol; *see also* Benj Edwards, *MCP: The New "USB-C for AI" That's Bringing Fierce Rivals Together*, ARS TECHNICA (Apr. 1, 2025); Kyle Wiggers, *Google to Embrace Anthropic's Standard for Connecting AI Models to Data*, TECHCRUNCH (Apr. 1, 2025); Kyle Wiggers, *OpenAI Adopts Rival Anthropic's Standard for Connecting AI Models to Data*, TECHCRUNCH (Mar. 26, 2025).

[208] Chinmayi Sharma, *Concentrated Digital Markets, Restrictive APIs, and the Fight for Internet Interoperability*, 50 U. MEMPHIS L. REV. 441, 455–61 (2019).

[209] Training data is a "mix of publicly available, licensed data and information from Meta's products and services." meta-llama, *llama-models*, GITHUB, https://github.com/meta-llama/llama-models/blob/main/models/llama4/MODEL_CARD.md (last visited Aug. 1, 2025).

their data can preserve a decisive competitive edge, regardless of who has access to their model weights. Open-weight models trained on publicly available datasets like Common Crawl rarely match the performance of those trained on curated, private corpora.[210]

Fourth, labor. A thriving AI ecosystem requires open opportunities for people to shape and build them. However, when expertise is locked within a few corporate or academic labs through restrictive employment practices and a failure to invest in broad talent development, the human capacity needed to realize the potential of openness is stifled. This creates a facade of access while concentrating the most critical resource—human talent—in the hands of a few.[211]

C. Democratic Access and Control

Democratizing AI is a twofold goal. It requires liberalizing access to ensure powerful tools are available to the many, not just the few. It also demands the establishment of democratic societal control, so the public has a say in how these technologies are developed and deployed. True democratization, therefore, means both expanding the number of people who can build with and benefit from AI and ensuring that its evolution reflects public values rather than the narrow interests of powerful corporations or states.

1. Benefits

Openness is a powerful force for liberalizing access to technology. By dramatically lowering the immense cost of entry, osAI frameworks like PyTorch and open-weight models like Meta's Llama have fundamentally altered the AI landscape. They allow startups, academic labs, and independent researchers to build upon state-of-the-art foundations, moving AI development beyond a handful of elite corporate labs and fostering a more vibrant and competitive ecosystem. In a global context, this distribution of technical capability helps mitigate the risk that AI prowess remains confined to the U.S.–China duopoly. Teams in Nairobi or São Paulo can fine-tune frontier models for local agriculture, public health, or language-revitalization projects, creating opportunities for local adaptation where centralized systems have historically failed.

Beyond broadening inclusive participation, openness is a critical tool for enabling democratic control. Openness in documentation, system prompts, safety benchmarks, and bias detection tools empowers civil society, allowing

---

[210] *Leaderboard Overview*, LMARENA, https://lmarena.ai/leaderboard (last visited Aug. 1, 2025) (no model on this AI leaderboard has been trained CommonCrawl or another similarly open dataset).

[211] Toner, *supra* note 150, at 5.

journalists, advocates, and independent auditors to put corporate and government claims to the test, challenge discriminatory outcomes, and hold powerful institutions accountable. For example, automated fairness audits have revealed racial disparities in credit-scoring algorithms, enabling regulatory enforcement and consumer protection. Open compliance tools allow external experts to stress-test AI for bias, fraud, and security vulnerabilities, preventing companies from self-policing in ways that prioritize corporate interests over public welfare.

Finally, enhancing the diversity of the human labor behind AI—through inclusive hiring, equitable training pathways, and protections for whistleblowers and annotators—strengthens democratic control over systems. Too often the communities most affected by AI are excluded from its development, resulting in systems that are less likely to serve their interests. Expanding the range of backgrounds and institutions shaping these tools is therefore essential for ensuring that societal control is not just a theoretical ideal but a practical reality.

2. Costs

Unfortunately, openness does not automatically lead to a more equitable or controllable AI ecosystem. If not carefully structured, it can paradoxically undermine the very democratic goals it purports to serve by creating an illusion of control while consolidating power, or by fragmenting authority so completely that collective governance becomes impossible.

First, openness can create a facade of democratic control that masks a deep consolidation of corporate power.[212] Even if a model's weights are released for free, the power to deploy it at scale remains centralized in the hands of the few corporations that control the underlying compute cloud infrastructure and the models themselves. These firms can become de facto private regulators, exercising unfettered control over the terms of AI access, development, and use. When a handful of unaccountable companies can decide which political speech, scientific research, or social tools are allowed to run on their platforms, the power to shape society shifts from democratic institutions to corporate boardrooms.

Second, openness can undermine society's ability to exert collective control by leading to uncontrollable fragmentation. As discussed above, once an open-weight model is released, the ability to enforce terms of use, responsible practices, and even legal compliance is severely hampered. The decentralized nature of the ecosystem also curtails the ability to identify a single actor to hold accountable for harm. This allows companies to distance themselves from the

[212] *See* Widder et al., *supra* note 28, at 831.

consequences of downstream misuse, shifting the externalities of their products onto the public. While this form of openness empowers the individual user, it critically weakens the power of society as a whole to set and enforce rules.

D. National Security and Global Leadership

AI is not just an economic and technological asset—it is a strategic resource that will shape military capabilities, intelligence dominance, and geopolitical influence for decades to come.[213] The race to control AI is already a defining factor in global power struggles, determining which nations lead in technological advancement, economic strength, and security. AI models underpin cybersecurity, intelligence gathering, autonomous military systems, and economic stability, making their regulation a matter of national security as much as technological governance.

1. Benefits

Strategic openness in AI can enhance national power and secure global leadership. By championing osAI, the U.S. can establish its technology as the global standard, shoring up its economic power while preventing adversaries from spreading their technological influence, as China has tried to do by, for example, integrating its models into Saudi Arabia's national oil company.[214] Much like Google's open source Android operating system secured American influence over the mobile technology landscape, osAI creates a powerful gravitational pull, drawing international users into an ecosystem that naturally favors U.S. cloud infrastructure and hardware. Far from symbolic, this leadership is critical for shaping the future of digital governance. When the most widely adopted AI models, architectures, and regulatory frameworks come from open, democratic sources, global norms may have a better chance of favoring civil liberties over surveillance, pluralism over censorship, and accountability over opacity. Conversely, if authoritarian states dominate the global AI ecosystem, they will embed values of control and repression into core infrastructures.[215]

Furthermore, American technological dominance has historically rested on fostering a deep talent pool of both domestic and foreign experts. Sustaining

---

[213] *See* TRIVEDI & MEYSENBURG, *supra* note 29, at 20–22.

[214] *See* Malcolm Moore, *Saudi Aramco Chief Says DeepSeek AI Makes 'Big Difference' to Operations*, FIN. TIMES (Mar. 4, 2025); *see, e.g.,* Mohammed Soliman, *Realigning US-Saudi Relations for the AI Era*, MIDDLE EAST INSTITUTE (May 5, 2025), https://www.mei.edu/publications/realigning-us-saudi-relations-ai-era (emphasizing the U.S. need to keep ahead of "competing tech corridors being built by China").

[215] Zeyi Yang, *Here's How DeepSeek Censorship Actually Works—and How to Get Around It*, WIRED (Jan. 31, 2025).

this advantage requires proactively maintaining pathways into the national talent pipeline through robust and inclusive STEM education and modernized immigration systems. Training international students in American universities, for example, remains one of the country's most effective forms of soft power, seeding global influence while attracting the world's brightest minds—whether the talent it cultivates remains in the country or not. Restrictive immigration policies, such as those currently being pursued by the present administration, undermine our global leadership and create opportunities for adversaries to snatch up the experts we myopically reject.

Finally, openness strengthens military alliances by enabling deep technological collaboration.[216] When the U.S. and its partners build upon shared, open AI frameworks, they dismantle the technical barriers that have historically complicated joint operations. This common foundation enables seamless data fusion, shared intelligence pictures, and integrated command-and-control systems, allowing allied forces to act with a speed and cohesion that closed, proprietary systems cannot match. A collaborative approach, fostered by osAI, allows allies to pool resources and talent to out-innovate adversaries, bolstering collective defense.

2. Costs

Despite these benefits, an unrestricted approach carries severe and direct risks to national security. When advanced models release weights publicly, they can provide a powerful accelerant to adversary states, allowing them to bypass years and billions of dollars of research and development and adapt these models for military and intelligence applications. While some researchers doubt whether today's osAI actually increases adversary capabilities, the pace at which the technology is advancing makes it a possibility that cannot be ignored.[217]

Open model risk is dangerously compounded by openness in the hardware supply chain. The widespread commercial availability of high-performance computing chips gives America's adversaries a powerful toolkit to close the technological gap. American-designed chips, particularly those from Nvidia,

---

[216] BEN FITZGERALD, PETER L. LEVIN & JACQUELINE PARZIALE, CTR. FOR A NEW AM. SEC., OPEN SOURCE SOFTWARE & THE DEPARTMENT OF DEFENSE 9 (2016), https://www.cnas.org/publications/reports/open-source-software-and-the-department-of-defense.

[217] *See* CHRISTOPHER A. MOUTON, CALEB LUCAS & ELLA GUEST, RAND, THE OPERATIONAL RISKS OF AI IN LARGE-SCALE BIOLOGICAL ATTACKS 1 (2024), https://www.rand.org/pubs/research_reports/RRA2977-2.html.

have powered much of China's AI development.[218] China is also leveraging other open compute components, such as chip design and the software that connects it to applications chips power, to build out its own compute infrastructure, allowing it to move away from western compute providers and undermining the efficacy of export controls.[219]

The threat also extends beyond rival nations to non-state actors. The same open models that empower startups can be weaponized by terrorist cells, transnational criminal organizations, or hacktivist collectives. Low barriers to entry mean that groups with limited resources can suddenly access previously unattainable capabilities, dramatically increasing the risk of sophisticated, AI-powered disinformation campaigns, automated cyberattacks against critical infrastructure, and even the design of biological or chemical weapons.

The emergence of DeepSeek, a Chinese AI company whose models now rival the best proprietary systems in the West, serves as a powerful case study. Built in record time on open Western architectures[220] and likely powered by high-end American chips available before export restrictions took full effect,[221] DeepSeek's success illustrates the peril of osAI: openness catalyzes rapid innovation, but not always in ways aligned with U.S. strategic interests.[222] The fact that DeepSeek is an independent company, unlike state-backed behemoths such as Baidu or Tencent,[223] highlights a broader risk: osAI is not just enabling China's largest firms but fostering a diverse ecosystem of smaller competitors that are more resilient to Western restrictions.

The long-term implications of this proliferation are complex. While some hope that adversaries adopting open Western components might also adopt more democratic technical norms, this outcome is far from certain; such hopes were shared when China joined the World Trade Organization, but economic

---

[218] Zijing Wu & Eleanor Olcott, *Nvidia AI Chips Worth $1bn Smuggled to China After Trump Export Controls*, FIN. TIMES (July 24, 2025); Che Pan & Casey Hall, *Nvidia AI Chips: Repair Demand Booms in China for Banned Products*, REUTERS (July 25, 2025).

[219] Che Pan & Brenda Goh, *Exclusive: China to Publish Policy to Boost RISC-V Chip Use Nationwide, Sources Say*, REUTERS (Mar. 4, 2025, 12:14 PM CST) ("China plans to issue guidance to encourage the use of open-source RISC-V chips nationwide for the first time.").

[220] Specifically, DeepSeek distilled Llama to create some of its models, which in turn relied on the transformer architecture invented by Google. DEEPSEEK, DEEPSEEK-R1: INCENTIVIZING REASONING CAPABILITY IN LLMS VIA REINFORCEMENT LEARNING (2025), https://github.com/deepseek-ai/DeepSeek-R1/blob/main/DeepSeek_R1.pdf.

[221] *See* Nathan Lambert, *The American DeepSeek Project*, INTERCONNECTS (July 4, 2025), https://www.interconnects.ai/p/the-american-deepseek-project.

[222] *See* Paul Mozur et al., *China's Rush to Dominate A.I. Comes With a Twist: It Depends on U.S. Technology*, N.Y. TIMES (Feb. 21, 2024).

[223] Cade Metz, *What to Know About DeepSeek and How It Is Upending A.I.*, N.Y. Times (Jan. 27, 2025).

integration did not produce political liberalization.[224] Moreover, an honest assessment requires looking inward. The U.S. is not immune to the problematic uses of AI, such as deploying AI-powered surveillance against immigrants, raising concerns about whether American leadership always aligns with democratic ideals.[225]

Ultimately, the rapid pace of AI progress suggests a permanent technological lead may be impossible for any single nation. If osAI empowers competitors like China, it may paradoxically improve global security by shifting the strategic calculus from zero-sum competition to mutual risk management.[226] Recognizing a shared interest in preventing a future that no one can unilaterally control could create powerful new incentives for cooperation on global safety standards, a dynamic that echoes the logic of nuclear arms control.[227]

E. Navigating Tradeoffs in AI Openness

Regulating osAI is an exercise in strategic prioritization. The central challenge is not whether AI should be "open" or "closed," but how differential openness at the component level creates trade-offs along two axes. These tensions exist *within* single policy goals and *between* competing ones, all of which are layered upon deeper *structural* conflicts. Strategic governance, therefore, is not about ideology but calibration: weighing the costs and benefits of opening each component of an AI system. AI is simultaneously an economic asset and a security risk, a public good and a proprietary investment— contradictions that are built into its very design.

1. Tradeoffs Within Policy Goals

The decision to open any specific AI component is often a double-edged sword, capable of both advancing and undermining the same policy objective. For public safety, which depends on oversight and control, openness is critical.

---

[224] *See What Happened When China Joined the WTO?*, COUNCIL ON FOREIGN RELS., https://education.cfr.org/learn/reading/what-happened-when-china-joined-wto (last updated Feb. 6, 2025).

[225] *See Steven Hubbard, Invisible Gatekeepers: DHS' Growing Use of AI in Immigration Decisions*, AM. IMMIGR. COUNCIL (May 9, 2025), https://www.americanimmigrationcouncil.org/blog/invisible-gatekeepers-dhs-growing-use-of-ai-in-immigration-decisions/.

[226] *See* Steven Adler, *Are We Ready for a 'DeepSeek for Bioweapons'?*, LAWFARE (May 29, 2025), https://www.lawfaremedia.org/article/are-we-ready-for-a--deepseek-for-bioweapons.

[227] *See* Simon Goldsten & Peter N. Salib, *DeepSeek Points Toward U.S.-China Cooperation, Not a Race*, LAWFARE (Mar. 5, 2025), https://www.lawfaremedia.org/article/deepseek-points-toward-u.s.-china-cooperation--not-a-race.

Making components like system prompts, operational metadata, and control layers transparent allows regulators, researchers, and companies to monitor failures and intervene early. Yet, the same transparency that enables oversight also invites exploitation. Security benchmarks designed to build trust in a fraud detection system can double as roadmaps for adversaries seeking to evade it. Similarly, while greater labor inclusivity can improve errors detection by bringing in diverse perspectives, expanding the talent pool too quickly without shared standards can lead to inconsistent and risk-prone practices, particularly in high-stakes domains.

Innovation, by contrast, thrives on experimentation, flexibility, and rapid iteration, which are fueled by open-weight models, transparent architectures, available compute, and accessible training data. These components maximize experimentation across a broader ecosystem of players. However, openness does not inherently create competition; it can also entrench dominance. For instance, companies that release model weights openly while keeping components like cloud services, fine-tuning expertise, or proprietary hosting platforms locked behind paywalls can create dependencies that stifle the very flexibility on which innovation relies.

Accountability, in turn, relies on transparency, traceability, and interpretability. Opening components like operational data and system prompts make it possible to audit, challenge, and correct AI decisions. But more openness doesn't guarantee more accountability. Open training data can expose bias but also diffuse responsibility; if a model trained on public datasets produces discriminatory outcomes, it becomes difficult to assign blame. Furthermore, transparency can invite regulatory arbitrage. Full visibility into evaluation benchmarks may lead developers to optimize for test performance rather than real-world fairness or robustness, turning a push for oversight into a playbook for compliance theater.

Finally, national security depends on maintaining a strategic advantage through controlled access to powerful technology. Restricting model weights, proprietary training data, and advanced systems helps ensure that critical AI tools stay in trusted hands. But excessive secrecy can backfire. If U.S. systems remain too closed, global users may turn to alternative ecosystems, eroding American influence and control.

2. Tradeoffs Between Policy Goals

These trade-offs become even more acute when different policy goals, each with its own logic for openness or closure, come into direct conflict. The most persistent tension exists between the need for secrecy in the name of public safety and the demand for transparency to foster innovation. This conflict plays out across multiple components. For example, transparency into the human alignment pipeline—who curates data, designs prompts, or flags edge cases—

and operational records enhances safety by allowing external scrutiny. However, it can also chill innovation if researchers fear legal or professional retaliation for pursuing controversial ideas. Conversely, the openness intended to promote innovation, such as releasing model weights or training datasets, can complicate efforts to enforce safety and accountability. Open weights allow anyone to strip safeguards, while open datasets built on sensitive sources introduce risks of bias, privacy, and security.

A similar clash occurs between innovation and accountability. The very components that ensure AI decisions can be audited and corrected—such as transparent system prompts and operational data—often introduce regulatory friction. Requiring their openness imposes compliance costs, delays deployment, and can raise the bar for entry, particularly for small players. A company that develops an AI system capable of dramatically improving cancer detection may be forced to delay or redesign it if laws demand full interpretability, a standard that many cutting-edge deep learning models, by their very nature, cannot meet.[228]

A parallel tension emerges between competition and national security. Openness in datasets, architectures, and evaluation benchmarks is essential for breaking up AI monopolies by lowering entry barriers for startups academic labs. But the same openness that fosters domestic competition can erode the technological asymmetry that national security depends upon by accelerating adversarial capabilities. If frontier AI models or military-grade training data were made fully open, adversarial states would gain immediate access to capabilities once held exclusively by a few AI leaders. At the same time, excessive secrecy in the name of national security risks stifling domestic competition just as much as it hinders foreign rivals.

3. Deeper Structural Tradeoffs

These policy tradeoffs are layered onto a deeper structural tension between centralization and decentralization. Beyond who gets access to AI lies the question of who builds and governs it. From a safety, accountability, and national security perspective, centralizing development within a few "national champion" firms simplifies top-down enforcement, making it easier for regulators to secure sensitive capabilities, enforce safeguards, maintain professional standards, and oversee compliance.[229] But while a centralized, security-first model

---

[228] *See* Emrullah ŞAHiN, Naciye Nur Arslan & Durmuş Özdemir, *Unlocking the Black Box: An in-Depth Review on Interpretability, Explainability, and Reliability in Deep Learning*, 37 NEURAL COMPUT & APPLIC 859 (2025), https://doi.org/10.1007/s00521-024-10437-2.

[229] *See* Jack M. Balkin, *Old-School/New-School Speech Regulation*, 127 HARV. L. REV. 2296, 2298 (2014).

may make oversight easier, it can create single-points of failure in the ecosystem with cascading effects.[230] It also risks starving the ecosystem of the flexibility and diversity of thought needed for long-term innovation and global leadership. Decentralization, by contrast, is foundational to democratization, as it redistributes power from a few powerful actors to the many communities impacted by the technology.

Beneath this lies an even more fundamental conflict: the imperative for control versus the ideal of freedom. Control is essential for national security, safety, and even certain forms of market-driven innovation, as it involves limiting access and embedding safeguards and maintaining oversight. Yet this control—whether over model behavior through alignment filters or over components as proprietary assets—inevitably constrains the freedom to iterate, experiment, and build a decentralized osAI ecosystem. A model marketed as reducing harmful speech, for example, can become a form of soft censorship if its operational controls are not transparent. Likewise, controlling the labor pipeline to ensure high standards may protect national assets from economic espionage but also erode the ethos of widespread public participation.

All these tensions are intensified by geopolitical urgency. In a global race for AI leadership, especially between the U.S. and China, speed is often treated as a proxy for strength, creating immense pressure to prioritize rapid deployment over rigorous oversight. In this climate, responsible approval cycles, fairness audits, and compliance thresholds are increasingly cast as obstacles in a zero-sum contest for dominance where even prudent caution can feel intolerable.

None of these dilemmas can be "solved" by picking openness or closedness. Rather, they are enduring trade-offs of differential openness that must be strategically managed. Every decision is an act of prioritization, requiring a sophisticated, context-aware governance approach that calibrates differential openness at the component level to strike a deliberate, evolving balance between competing values.

III. Calibrating Differential AI Openness

The default posture toward osAI today is largely reactive and uneven—shaped more by technical happenstance and institutional inertia than by deliberate governance. Core components like model weights, datasets, and prompts are often opened or withheld based on developer preference or commercial strategy rather than public policy. But as Part II made clear, how open or closed

[230] U.S. Dep't Treasury, "New Treasury Report Assesses Opportunities, Challenges Facing Financial Sector Cloud-Based Technology Adoption," press release, February 8, 2023, https://home.treasury.gov/news/press-releases/jy1252.

each component is has direct, often conflicting implications for safety, innovation, accountability and national security.

This Part moves from diagnosis to prescription, examining the specific legal and regulatory levers policymakers can use to strategically calibrate AI's differential openness. We analyze how tools related to liability, competition policy, intellectual property, trade controls, and direct government support can be targeted at specific components of the AI stack. For each, we establish the policy baseline and then map how specific interventions can be used to strike a more deliberate balance among competing public goals. This provides a concrete playbook for designing policies calibrated to differential openness at the component level, moving beyond blunt, system-level mandates to foster a safer, more innovative, and more accountable osAI ecosystem.

A. Liability

Liability frameworks are among the most powerful tools for shaping policy outcomes in technological ecosystems—but their current application to osAI components fails to target granular components and consider trade-offs in differential openness. Liability exposure for osAI developers is therefore uneven and often counterproductive. The status quo creates two overlapping problems: excessive openness of certain components without safeguards, and a chilling effect on responsible transparency that could enable better oversight and innovation. Reforming the way case law is applied to osAI or overriding common law with statutory interventions can discourage unfettered openness of components like datasets, model weights, and compute, that compromise safety and national security, while encouraging responsible openness of components like system prompts and operational records that enable true accountability.

1. Baseline

The legal liability for osAI is complex and contradictory. For osAI developers and users, openness acts as both a shield and a sword: existing legal doctrines can immunize open components from liability, yet these components' very transparency can also create unique legal exposures that make litigation more likely.

On the one hand, existing tort and contract doctrines benefit developers with broad liability shields that permit the release of some osAI components—especially model weights and training data—without embedded safeguards or control mechanisms.

First, they benefit from established legal defenses that are difficult for plaintiffs to overcome. Tort's economic loss doctrine generally blocks recovery for purely financial harm, relegating issues like lost profit from malfunctioning

systems to contract law.[231] But under contract law, the OSS licenses (e.g., MIT, Apache, GPL) under which many components are released almost universally disclaim warranties and liability, providing the software on an "as is" basis[232] and denying users the right to sue for defects.[233] Aside from economic harms, tort law is also doctrinally inaccessible to plaintiffs who suffer harms stemming from misinformation or biased outputs when they are not accompanied by demonstrable physical or emotional harm.

Second, the technical nature of osAI components provides a powerful argument against negligence claims. Establishing negligence requires proving the defendant breached an established standard of care—in other words, failed to take reasonable measures to prevent the plaintiff's harm. Counterintuitively, the fact that open components like datasets and model weights are, unlike their proprietary counterparts, inherently hard to secure gives developers a colorable argument that they did not breach a standard of care. Once developers release these specific components, they forfeit control over their downstream uses. They can argue that they cannot be held legally responsible for failing to do the impossible: include non-removable safety features in a system that, by its very nature, is designed to be modified and stripped of safeguards.[234] If courts recognize this defense, developers have little incentive to include or maintain safety components like alignment layers, usage constraints, or prompt protections that, while removeable, can still curb many harmful uses.

Finally, the decentralized structure of the osAI ecosystem creates immense practical barriers to litigation. It is nearly impossible to trace a specific harm back to a single responsible party among a global network of contributors—for example, to identify the entity that contributed CSAM to an open dataset. This shield extends to infrastructure-level osAI actors like cloud providers and hosting services, who can argue that downstream misuse is too legally remote to

---

[231] Michael D. Scott, *Tort Liability for Vendors of Insecure Software: Has the Time Finally Come?*, 67 MD. L. REV. 425, 470 (2008).

[232] *See, e.g.*, *GNU General Public License*, GNU OPERATING SYS., https://www.gnu.org/licenses/gpl-3.0.en.html (last visited Aug. 1, 2025) (providing software "'AS IS' WITHOUT WARRANTY OF ANY KIND"); *The MIT License*, OPEN SOURCE INITIATIVE, https://opensource.org/license/mit (last visited Aug. 1, 2025) (same); *Apache License, Version 2.0*, APACHE SOFTWARE FOUND., https://www.apache.org/licenses/LICENSE-2.0 (last visited Aug. 1, 2025) (same).

[233] *See* Choi, *supra* note 49 (manuscript at 6–7); *see also* Michael L. Rustad & Thomas H. Koenig, *The Tort of Negligent Enablement of Cybercrime*, 20 BERKELEY TECH. L.J. 1553, 1565 (2005) (noting courts' position "in favor of broad enforceability of mass market license agreements").

[234] *See* KETAN RAMAKRISHNAN ET AL., RAND, U.S. TORT LIABILITY FOR LARGE-SCALE ARTIFICIAL INTELLIGENCE DAMAGES: A PRIMER FOR DEVELOPERS AND POLICYMAKERS 29 (2024), https://www.rand.org/pubs/research_reports/RRA3084-1.html.

establish proximate cause. They may also invoke Section 230 of the Communications Decency Act,[235] claiming immunity for content generated by third parties on their systems, although the legal viability of such a defense remains to be litigated.[236]

On the other hand, while openness can provide a defense, it can also become a vulnerability that exposes osAI developers and even users to unique forms of liability. The transparency inherent in certain open components—like system prompts, operational records, and control mechanisms—creates new avenues for litigation that do not exist for proprietary "black box" systems. When these are made public, they provide a detailed road map for a plaintiff's legal team to scrutinize a system's design and build a case for negligence. This dynamic creates a chilling effect, where osAI developers may fear that safety-enhancing transparency will paradoxically increase their legal exposure.

Furthermore, openness can shift the burden of responsibility onto the users of osAI components. A court could determine that the ability to inspect, stress test, and improve an open component creates a legal duty to do so, meaning a user who deploys an osAI model without a reasonable audit or safety-enhancing modifications could be found negligent if harm occurs. This potential for downstream liability could discourage the adoption of open components, slowing innovation.

The status quo of liability for open AI components is fraught with uncertainty. But this baseline is just that—a baseline. Policymakers can choose to deviate from it, overriding common law to either encourage or discourage the development and proliferation of individual open components.

2. Reforms

Courts and policymakers can shift the baseline of liability for specific osAI components to land on the optimal configuration of differential openness—the one that strikes a more desirable set of tradeoffs. For example, courts can adapt legal doctrines to new technologies and narrow these defenses. They could decline to enforce overly broad liability disclaimers, especially if they treat osAI components like model weights or datasets as products that foreseeably cause emotional harm[237] (for example, nonconsensual pornography deepfakes) or

---

[235] 47 U.S.C. § 230.

[236] *See, e.g.*, Sean Norick Long, Esther Tetruashvily & Ashwin Ramaswami, *Why Section 230 Reformers Should Start Paying Attention to Social Code Platforms*, GEO. L. TECH. REV. (Nov. 2022), https://georgetownlawtechreview.org/why-section-230-reformers-should-start-paying-attention-to-social-code-platforms/GLTR-11-2022/.

[237] *See, e.g.*, Garcia v. Character Techs., Inc., No. 6:24-CV-1903-ACC-UAM, 2025 WL 1461721 (M.D. Fla. May 21, 2025) (permitting lawsuit against AI chatbot company alleged to have caused a user's suicide).

physical harm (for example, cyberattacks on critical infrastructure).[238] Similarly, courts could reject the argument that releasing an inherently insecure osAI component does not breach a standard of care, finding that the decision to release the unsecured components was unreasonable in the first instance, particularly if the osAI developer knew that misuse was uncontrollable and the consequences of release were irreversible.

Courts could also extend liability to osAI infrastructure providers. A court might find that a cloud provider that hosts open models proximately caused harm if the downstream misuse of its services was foreseeable. This legal link is stronger for cloud providers than for distributors of data or software, as they often have the technical capacity to monitor and intervene in how their platforms are used.[239] Imposing liability on these actors—those with the most visibility and control in the AI stack—could strike a more effective balance than a uniform rule applied to all osAI component distributors.

Finally, courts could calibrate the liability of osAI users. Rather than requiring full technical audits, which are resource-intensive and could chill adoption, courts could establish a scalable "duty to inquire." In a negligence suit, this would mean assessing whether an osAI user took reasonable precautionary measures commensurate with the deployment risk. For high-risk contexts like medicine, this could create a legal expectation that users consult accessible documentation like model cards, README files, or public vulnerability disclosures—both before and after deployment. Holding osAI users accountable for ignoring known risks would shift liability to those best positioned to manage deployment risks, while also incentivizing developers of open components to be more transparent in the operational records that could facilitate this duty.

However, relying on courts to adapt common law doctrines is an incremental process that may not keep pace with technological change. For this reason, policymakers could intervene directly to calibrate osAI liability. For example, legislatures hold the power to overcome existing legal barriers, whether by narrowing the economic loss doctrine, expanding the range of compensable harms, or clarifying that Section 230 offers no shield for harms enabled by AI systems built with open components. Such statutes could impose proactive obligations across the AI stack, requiring actors who build, host, or distribute

---

[238] *See* Scott, *supra* note 231, at 471 ("Arguments can be made, however, that some claims arising from the failure of security software should be recoverable despite the economic loss rule. For example, a company's reputation is an interest protected by tort law. Additionally, the data contained in the computer is property separate and apart from the software itself.") (citation altered).

[239] *See, e.g.,* David Evan Harris, LINKEDIN (Dec. 18, 2024, 03:37:35 PM EST), https://www.linkedin.com/posts/davidevanharris_mitigating-the-risk-of-generative-ai-models-activity-7275247850702217216-0K7l/ (reporting that HuggingFace refuses to take down projects known to be built on datasets containing high quantities of CSAM).

open components to anticipate and monitor the downstream effects of their tools.

However, any attempt to legislate new liability regimes could produce perverse outcomes. First, the threat of litigation, while a routine cost for large corporations, can be a death sentence for the community-driven projects that are the lifeblood of the open ecosystem. Most of these osAI projects operate with minimal financial reserves and no insurance, making them uniquely vulnerable to the costs of legal defense, which could deter new entrants and eliminate existing players in the nascent osAI ecosystem.

Moreover, a miscalibrated liability framework risks chilling component openness when openness might be desired. Faced with new legal exposures, osAI developers may retreat from transparency, closing off components like operational records out of fear they could be used as evidence of negligence. Likewise, commercial osAI users facing a new duty to inspect open components may simply forgo them, deeming the financial and technical burden of auditing them too high. Even osAI compute providers, who are best positioned to absorb litigation costs, might restrict access to their services rather than accept new monitoring obligations and their attendant risks. The result could be an ecosystem that is less transparent, less innovative, and less safe.

Early legislative proposals reveal the difficulty of calibrating policy with an eye toward differential openness. California's "Safe and Secure Innovation for Frontier Artificial Intelligence Models Act" (more commonly known as SB 1047), though ultimately vetoed, would have imposed forward-looking duties on developers of frontier models, requiring safety certifications and making them liable for foreseeable harms stemming from failures to adequately mitigate misuse with no exemptions for osAI.[240] But SB 1047 was criticized for being a blunt instrument that would have "disincentivize[d] developers from open-sourcing their models" because of the threat of liability and the costs of complex safety procedures.[241]

Similarly, the EU AI Act imposes documentation, risk management, and conformity obligations on providers of general-purpose AI models, while also exposing them to some liability.[242] At the same time, Act shows the danger of legislating without a deep understanding of differential openness in the AI stack.[243] For example, it creates a loophole by exempting many "open-source"

---

[240] S.B. 1047, 2024 Leg., Reg. Sess. (Ca. 2024).

[241] Andy Jung, *California's AI Bill Threatens To Derail Open-Source Innovation*, REASON (Aug. 13, 2024).

[242] 2024 O.J. (L 1689) art. 53, https://artificialintelligenceact.eu/article/53/.

[243] David Atkinson, *Open Shouldn't Mean Exempt: Open-Source Exceptionalism and Generative AI* (July 24, 2025) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5355736 (cautioning against exceptionalism for AI openness).

models from transparency requirements as long as they publish their model weights architecture, and usage data, which allow osAI developers releasing some open components to receive preferential treatment while keeping more critical components, such as training datasets, closed from public scrutiny.[244] And even when the Act applies to certain high-risk "open-source" models, it forgoes the opportunity to demand disclosure around datasets, a component that warrants some oversight in the name of safety and democratic control.[245]

A more surgical approach would shift liability downstream, focusing on the commercial users of open components rather than pre-commercial developers—a distinction in the osAI ecosystem that laws like SB 1047 and the EU AI Act miss. For example, the European Union's proposed Product Liability Directive (PLD) offers a promising model by carving out an explicit exemption for developers not engaged in commercial activity. This creates a safe harbor that protects upstream, non-commercial innovation while ensuring that accountability attaches to the "economic operators" who actually commercialize and deploy osAI technologies in the market.[246]

Alternatively, legislative intervention can foster openness in the labor component, reviving avenues for liability even when technical components are closed. Bolstering whistleblower and anti-retaliation protections, for example, harnesses the human element to identify risky decision-making at its source.[247] This introduces a degree of transparency into the AI stack that preserves liability as a tool for promoting safety, without relying on the unfettered openness of components most susceptible to abuse, like model weights.

A final note: even the best calibrated liability regimes may collide with constitutional limits. The First Amendment may constrain efforts to regulate the publication or distribution of certain software or data-based osAI components—such as source code, datasets, system prompts, or model weights—if courts view them as information and therefore protected forms of speech.[248] Indeed, the landmark *Bernstein* line of case establishes that software source

---

[244] *See* 2024 O.J. (L 1689) art. 2, https://artificialintelligenceact.eu/article/2/.

[245] *Id.* art. 53, https://artificialintelligenceact.eu/article/53/.

[246] *See* 2024 O.J. (L 2853) 2–3.

[247] *See* Charlie Bullock & Mackenzie Arnold, *Protecting AI Whistleblowers*, LAWFARE (June 25, 2025), https://www.lawfaremedia.org/article/protecting-ai-whistleblowers.

[248] *See* Kleindienst v. Mandel, 408 U.S. 753, 762–63 (1972) ("In a variety of contexts this Court has referred to a First Amendment right to 'receive information and ideas' . . . "); Ilan Kogan, *Artificial Intelligence, Existential Risk, and the First Amendment*, 27 U. PA. J. CONST. L. 156, 202–10 (2025); *see also* Eugene Volokh, Mark A. Lemley & Peter Henderson, *Freedom of Speech and AI Output*, 3 J. FREE SPEECH L. 651, 654–57 (2023).

code, as a vehicle for communicating ideas among researchers, constitutes protected speech and that regulatory burdens on its distribution constitute impermissible prior restraint.[249]

Even if this line of cases were not applied to other components, such as datasets and model weights, individuals might still claim a right to receive information and ideas through accessing and using osAI components, which would independently implicate First Amendment protections, albeit potentially at lower levels of constitutional scrutiny.[250] It remains unresolved, however, whether components enabling AI functionality but lacking human readability—such as model weights, source code, system prompts, and operational controls—would similarly qualify as expressive under the First Amendment, and, if so, what level of scrutiny courts would apply.[251] While the contours of constitutional protection for open components is unclear, policymakers seeking to impose liability on their distribution must be prepared to navigate the murky line between speech rights and osAI governance.

B. Competition

Competition policy that encourages the right configuration of differential openness in osAI systems is key to unlocking innovation and strengthening global leadership, but it must not come at the cost of safety and democratic control. The AI ecosystem is marked by extreme concentration, with a handful of firms dominating every layer of the stack—from compute and data to foundation models and deployment platforms—all within a limited and often contradictory enforcement landscape. Counteracting these monopolistic harms requires a toolkit of potential reforms, including ex ante structural rules, renewed ex post antitrust enforcement, and policies designed to open the critical nontechnical component of the AI stack: the labor market.

---

[249] Bernstein v. U.S. Dep't of State, 922 F. Supp. 1426, 1435 (N.D. Cal. 1996), *aff'd sub nom.,* Bernstein v. U.S. Dep't of Justice, 176 F.3d 1132, 1135 (9th Cir. 1999).

[250] *See* Peter N. Salib, *AI Outputs are Not Protected Speech*, 102 WASH. U. L. REV. 83, 142–44 (2024).

[251] *Compare* Alan Z. Rozenshtein, *There Is No General First Amendment Right to Distribute Machine-Learning Model Weights,* LAWFARE (Apr. 4, 2024), https://www.lawfaremedia.org/article/there-is-no-general-first-amendment-right-to-distribute-machine-learning-model-weights*, and* Doni Bloomfield, U.S. Expert Controls of AI Models 25–32 (June 3, 2024), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4741033 (unpublished manuscript), *with* Michael Paradis, *Regulations Targeting Large Language Models Warrant Strict Scrutiny Under the First Amendment*, LAWFARE (July 26, 2024), https://www.lawfaremedia.org/article/regulations-targeting-large-language-models-warrant-strict-scrutiny-under-the-first-amendment.

1. Baseline

The AI ecosystem is marked by extreme concentration. Across the stack, a handful of firms dominate access to training data, the development of foundation models, the supply of essential compute hardware, the deployment platforms that bring AI to market, and the available human expertise.[252] This structural reality, combined with a historically light-touch approach to antitrust enforcement,[253] has, as former Federal Trade Commission (FTC) chair Lina Khan recognized, narrowed the range of players that can participate in the ecosystem.[254] While recent efforts, such as the 2024 joint DOJ-FTC investigations into firms like Nvidia, Microsoft, and OpenAI,[255] signal a potential shift toward more aggressive enforcement, it is unclear whether those investigations will manifest in lawsuits under the new Trump administration,[256] especially given its deregulatory stance and recent advocacy for jumpstarting U.S. leadership in AI.[257]

This concentration is not without its advantages. It's possible that, in a global technological race, consolidation is a strategic necessity, creating "national champions" with the scale and resources to out-compete state-backed rivals. On this view, a fragmented market would be less efficient. From an oversight perspective, a smaller number of actors is easier for regulators to monitor

---

[252] *See* Narechania & Sitaraman, *supra* note 28, at 99–100.

[253] *See, e.g.*, Rebecca Haw Allensworth, *Antitrust's High-Tech Exceptionalism*, 130 YALE L.J. F. 588, 591–92 (2021) (Thus, as American antitrust law entered its second century, it had only low-tech legal tools to confront high-tech market power . . . Scholarly criticisms of antitrust's failure to adapt to the new economy are . . . common. Also easy to find are criticisms of the federal government's underenforcement of the antitrust laws, especially merger enforcement, against the tech sector."); *see also* Christos A. Makridis & Joel Thayer, *The Big Tech Antitrust Paradox: A Reevaluation of the Consumer Welfare Standard for Digital Markets*, 27 STAN. TECH. L. REV. 71, 74–75, 77 (2024) ("Unfortunately . . . [the] lack of legislation, together with lax enforcement, have led to a wide array of abuses within the digital economy, ranging from monopoly power to flagrant exploitation that has ultimately hurt consumers.").

[254] Press Release, Fed. Trade Comm'n, FTC, DOJ, and International Enforcers Issue Joint Statement on AI Competition Issues (July 23, 2024), https://www.ftc.gov/news-events/news/press-releases/2024/07/ftc-doj-international-enforcers-issue-joint-statement-ai-competition-issues.

[255] *See* David McCabe, *U.S. Clears Way for Antitrust Inquiries of Nvidia, Microsoft and OpenAI*, N.Y. TIMES (June 5, 2024).

[256] *See* David McCabe, *What to Know About Trump's Antitrust Efforts Against Tech Giants*, N.Y. TIMES (Apr. 21, 2025).

[257] *See* EXEC. OFF. PRES., *supra* note 1.

and enforce against.[258] However, this baseline of consolidation exacts a heavy price. It strips dominant firms of the competitive pressure to invest in meaningful safety precautions, support open science, or prioritize democratic accountability over market control.[259] It may even create unique national security risks, with the bulk of the nation's AI capabilities resting in the hands of a few large, high-value targets.[260]

2. Reforms

Competition policy offers a powerful toolkit to counteract the monopolistic harms of the current AI ecosystem with calibrated differential openness. As discussed above, merely opening some components like model weights is insufficient when others, like compute and proprietary datasets, remain locked down.[261] True decentralization depends on the ability to inspect, reuse, and build upon core components and on the mobility of the talent that fuels the AI engine. The following interventions, targeting different layers of the AI stack, can be deployed to unlock innovation and strengthen democratic control.

The most direct interventions would regulate incumbent infrastructure using ex ante rules[262] drawn from the broader paradigm of network and platform regulation.[263] For example, a policy of structural separation—prohibiting a single firm from operating in different layers of a market—would break an incumbent's stranglehold on multiple layers of the AI stack by brute force. A firm could not, for instance, both develop foundation models and provide the essential cloud computing services that power the ecosystem. This would prevent vertically integrated firms from disadvantaging osAI competitors by restricting access, raising prices, or prioritizing their own deployments.[264] It also prevents companies, who have visibility into downstream uses of their cloud services,

---

[258] *See, e.g.*, Dakota Foster, *Antitrust Investigations Have Deep Implications for AI and National Security*, BROOKINGS INST. (June 2, 2020), https://www.brookings.edu/articles/antitrust-investigations-have-deep-implications-for-ai-and-national-security/; *see also* SATYA MARAR, MERCATUS CTR., ARTIFICIAL INTELLIGENCE AND ANTITRUST LAW: A PRIMER 13 (2024), https://www.mercatus.org/media/document/4815mararaiantitrustlawssv2pdf.

[259] Narechenia & Sitaraman, *supra* note 28, at 140–43.

[260] *Id.* at 139; U.S. Department of the Treasury, "New Treasury Report Assesses Opportunities, Challenges Facing Financial Sector Cloud-Based Technology Adoption," press release, February 8, 2023, https://home.treasury.gov/news/press-releases/jy1252.

[261] *See supra* Part II.A.2.

[262] Narechenia & Sitaraman, *supra* note 28, at 146–50.

[263] *See generally* MORGAN RICKS ET AL., NETWORKS, PLATFORMS, AND UTILITIES: LAW AND POLICY (2022).

[264] Narechenia & Sitaraman, *supra* note 28, at 159–60.

from providing better versions of model applications built on them once these applications have proven successful.[265]

A less rigid alternative, nondiscrimination, would regulate conduct rather than corporate structure, requiring dominant firms—especially compute providers—to offer services to osAI competitors on fair, reasonable, and non-discriminatory terms.[266] Finally, interoperability mandates, such as requirements to adopt protocols like Anthropic's MCP that standardize the connection between models and data,[267] could address vendor "lock-in." By compelling technical compatibility between components in the AI stack, interoperability empowers users to move between platforms—from proprietary hardware, where interoperability is especially lacking, to an open-source alternative, for instance—without re-engineering their entire system.[268]

While these tools offer a spectrum of options for prying open the market, they come with trade-offs: structural separation could stifle the efficiencies of vertical integration and make it harder to discern safe providers in a newly fragmented market,[269] while nondiscrimination and interoperability rules impose significant ongoing monitoring burdens on regulators.[270]

Alongside these structural rules, renewed ex post antitrust enforcement could meaningfully improve the competitive landscape. Regulators could apply stricter merger controls to block the acquisition of promising research labs or data providers[271]—paying especially close attention to big tech's new strategy of

---

[265] *See* Sharma, *supra* note 208 (discussing this trend in other software environments).

[266] *Id.* at 160–62.

[267] ANTHROPIC, *supra* note 207.

[268] *See* Sharma, *supra* note 208 at 162–64; *see* Sam Adler, *Interoperable Agentic AI: Unlocking the Full Potential of AI Specialization*, TECH POL'Y PRESS (Dec. 3, 2024), https://www.techpolicy.press/interoperable-agentic-ai-unlocking-the-full-potential-of-ai-specialization/.

[269] *See* Richard J. Gilbert, *Separation: A Cure for Abuse of Platform Dominance?*, 54 INFO. ECON. & POL'Y, March 2021, at 11 (2021), https://www.sciencedirect.com/science/article/abs/pii/S0167624520301207 ("Some types of innovations require coordination between complementary businesses, which is impeded if the businesses are confined to separate companies.").

[270] Narechenia & Sitaraman, *supra* note 28, at 161.

[271] *See, e.g.*, Press Release, Fed. Trade Comm'n, FTC Sues to Block $40 Billion Semiconductor Chip Merger, (Dec. 2, 2021), https://www.ftc.gov/news-events/news/press-releases/2021/12/ftc-sues-block-40-billion-semiconductor-chip-merger (detailing FTC to block Nvidia purchase of UK chip designer Arm).

"acquiring" companies not outright but through major investments and licensing deals[272]—as well as renew scrutiny of predatory pricing for cloud services.[273] It can also be used to more aggressively police other exclusionary conduct by dominant firms, such as making APIs to access model weights initially open, or at least permissive, and later exploit third party reliance to extract rent.[274] This approach could also be adapted to combat "open-washing"—the deceptive practice where companies claim the reputational benefits of openness while withholding critical components. Regulators like the FTC could treat misleading openness claims—for example, ones that deviate from the clear benchmark of the OSI's "Open Source AI Definition"—as "unfair or deceptive practices in or affecting commerce."[275] an unfair trade practice, using clear benchmarks to distinguish genuine openness from mere marketing. But enforcement must always be component-specific. While restricting mergers between two major data providers may reverse oligopoly, blocking the acquisition of smaller startups could chill entrepreneurship, as many startup founders are incentivized by the hope of "exiting," or being acquired by a major player.

Finally, competition policy can foster openness in a critical non-technical component: the labor market. Noncompete agreements restrict labor mobility in practice and slow the diffusion of safety knowledge and technical expertise across the industry.[276] As the FTC has noted: "To ensure a competitive and innovative marketplace, it is critical that talented individuals with innovative ideas be permitted to move freely, and, crucially, not be hindered by non-competes."[277] It was this approbation of labor mobility that gave birth to Silicon

---

[272] *See* Mark Isaac, *Cognition AI Buys Windsurf as A.I. Frenzy Escalates*, N.Y. TIMES (July 14, 2025) (describing Google's takeover of AI coding startup Windsurf by poaching its executives and top talent as well as licensing its technology, leaving what was left to either wither or, as they ultimately chose to, agree to an outright acquisition by AI coding competitor Cognition).

[273] *See, e.g.*, *Ofcom Refers UK Cloud Market to CMA for Investigation*, OFCOM (Oct. 5, 2023), https://www.ofcom.org.uk/internet-based-services/cloud-services/ofcom-refers-uk-cloud-market-to-cma-for-investigation ("referr[ing] the UK public cloud infrastructure services market to the [UK government] Competition and Markets Authority (CMA) to carry out a market investigation.").

[274] *See generally* Sharma, *supra* note 208 (discussing this trend in other software environments).

[275] *See* 15 U.S.C. § 45(a).

[276] *See* Ronald J. Gilson, *The Legal Infrastructure of High Technology Industrial Districts: Silicon Valley, Route 128, and Covenants Not to Compete*, 74 N.Y.U. L. REV. 575 (1999).

[277] *Generative AI Raises Competition Concerns*, FTC: OFFICE TECH. BLOG (June 29, 2023), https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/06/generative-ai-raises-competition-concerns.

Valley.[278] Aggressively policing noncompetes and other restrictive employment practices is essential for ensuring that human expertise, a core component of the AI stack, can circulate freely.

C. Intellectual Property

Intellectual property (IP) law, particularly copyright, is a core battleground for AI openness, creating a deep conflict between protecting creators' rights and enabling the data-driven innovation that powers AI. The current legal landscape, governed by the uncertain fair use doctrine, creates a perilous asymmetry: it exposes transparent osAI projects to crippling litigation while shielding proprietary developers who keep their data secret. This legal ambiguity chills the very openness essential for safety research and competition. Moving beyond dependence on the courts' slow adaptation of IP law to the current moment, will require legislative reforms that clarify fair use with component-level precision, explore new licensing frameworks, and potentially establish safe harbors to strike the configuration of differential openness needed to foster a more balanced and innovative AI ecosystem.

1. Baseline

The status quo of copyright enforcement creates a perilous asymmetry for osAI development, placing disproportionate legal pressure on the very forms of openness essential for safety and innovation. While proprietary developers shield their training data and development pipelines as trade secrets, true openness requires documenting methods, data sources, and design choices. This transparency—critical for reproducibility, peer review, and accountability—simultaneously creates a roadmap for infringement litigation, especially since training data is almost inevitably tainted with copyrighted content.[279]

This legal asymmetry between closed and open components is magnified by a stark resource imbalance. Large corporations building closed systems can absorb infringement lawsuits as a cost of doing business and even indemnify

[278] Mike McPhate, *California Today: Silicon Valley's Secret Sauce*, N.Y. TIMES (MAY 19, 2017).

[279] *See* Mark A. Lemley & Bryan Casey, *Fair Learning*, 99 TEX. L. REV. 743, 745 (2021). There have been some recent attempts to train AI systems entirely on public domain content. *See, e.g.*, Nikhil Kandpal et al., *The Common Pile v0.1: An 8TB Dataset of Public Domain and Openly Licensed Text* (June 5, 2025) (unpublished manuscript), https://doi.org/10.48550/arXiv.2506.05209. It is unclear, however, whether such projects will be competitive with the most advanced AI models. *See* Nitasha Tiku, *AI Firms Say They Can't Respect Copyright. These Researchers Tried.*, WASH. POST (June 5, 2025).

their customers,[280] whereas the same litigation can extinguish the smaller, community-driven osAI projects that are the lifeblood of a healthy open ecosystem. This dynamic creates a powerful chilling effect: osAI developers are incentivized to withhold components like training recipes, datasets, or model documentation to avoid IP liability, even when releasing them would unlock significant public value and enable crucial safety research.

The primary legal shield for developers of open components, namely model weights and datasets, is the fair use doctrine, an affirmative defense against copyright infringement. The analysis rests on four statutory factors,[281] but in practice, it often boils down to a tension between the two that most directly address the balance between innovation and protecting the economic rights of copyright holders[282]: (1) "the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes";[283] and (4) "the effect of the use upon the potential market for or value of the copyrighted work."[284] In applying this test, courts interrogate the degree to which an osAI component "transforms" the copyrighted material into something new, rather than merely substituting for the original.[285]

On one hand, osAI developers have a strong argument under the first factor. Courts can be urged to look favorably upon the non-commercial and publicly beneficial nature of open datasets and model weights. The goals of transparency, reproducibility, and academic collaboration directly serve fair use's fundamental purpose of promoting innovation and public knowledge, giving these uses a strong transformative character.

On the other hand, fair use arguments will often founder on the fourth factor: market harm. Courts have repeatedly emphasized that even a clearly transformative, non-commercial project can be infringing if it significantly damages the market for the original work.[286] This is a critical vulnerability for osAI. For instance, an open dataset containing large volumes of copyrighted books could

---

[280] *See, e.g.*, Ron Miller, *Adobe Indemnity Clause Designed to Ease Enterprise Fears About AI-Generated Art,* TECHCRUNCH (June 26, 2023).

[281] 17 U.S.C. § 1071.

[282] *See* Joseph P. Liu, *Two-Factor Fair Use*, 31 COLUM. J.L. & ARTS 571, 572 (2008).

[283] 17 U.S.C. § 1071(1).

[284] *Id.* § 1701(4).

[285] *See* Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 579 (1994).

[286] *See, e.g.*, Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 566 (1985) (describing the market effect factor as "undoubtedly the single most important element of fair use"); *see also* Suneal Bedi & Mike Schuster, *Measuring Fair Use's Market Effect*, 2022 WIS. L. REV. 1467, 1469 (2023) ("Although courts often employ all four factors, one has received substantial attention in fair use determinations—the so called "market effects" . . . factor.").

directly saturate the market those authors rely upon. This creates a "substitution effect" that might outweigh the component's transformative value. Why buy a book when you can access it for free? Similarly, researchers have shown that users can "jailbreak" open model weights to extract near carbon-copy reproductions of training data, directly competing with the original creations.[287] While osAI developers can and should be held responsible for curating the open data they release, they have far less control over how open model weights are misused after they are made public.

This unresolved tension leaves the fair use defense deeply uncertain in the context of osAI. With legal scholarship divided[288] and caselaw still in its infancy—highlighted by a prominent judge recently reversing his own initial ruling on the matter[289]—the doctrine offers more of a litigation gamble than a reliable safe harbor. This legal ambiguity alone is a powerful force, discouraging osAI developers from releasing beneficial components and chilling the very openness that is critical for research and innovation.

2. Reforms

Rather than await decades of conflicting judicial precedent to resolve copyright law's ambiguities, policymakers can intervene to provide the clarity the

---

[287] *See, e.g.*, Nicholas Carlini et al., *Extracting Training Data from Diffusion Models*, PROCS. 32ND USENIX CONF. ON SEC. SYMP. 5253 (2023), https://doi.org/10.5555/3620237.3620531.

[288] *Compare*, *e.g.*, Lemley & Casey, *supra* note 279, at 748–50 (arguing that copying for the purpose of training machine learning models should generally be protected by fair use); James Grimmelmann, *Copyright for Literate Robots*, 101 IOWA L. REV. 657, 664 (2016) ("Verbatim copying of a complete work will be protected as fair use if the copy is used solely as input to a process that does not itself use the works expressively. Or, to put it a little more provocatively, nonexpressive uses do not count as reading."), *with* Robert Brauneis, *Copyright and the Training of Human Authors and Generative Machines*, 48 COLUM. J.L. & ARTS 1, 58 (2024) ("The current case that generative AI training is a fair use is weak."); *see also* U.S. COPYRIGHT OFF., COPYRIGHT AND ARTIFICIAL INTELLIGENCE, PART 3: GENERATIVE AI TRAINING 74 (forthcoming 2025) (2025), https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf ("The copying of expressive works from pirate sources in order to generate unrestricted content that competes in the marketplace, when licensing is reasonably available, is unlikely to qualify as fair use.").

[289] Thomson Reuters Enter. Ctr. GMBH v. Ross Intel. Inc., 765 F. Supp. 3d 382 (D. Del. 2025), *motion to certify appeal granted*, No. 1:20-CV-613-SB, 2025 WL 1488015 (D. Del. May 23, 2025); *see also* Bartz v. Anthropic PBC, No. C 24-05417 WHA, 2025 WL 1741691 (N.D. Cal. June 23, 2025) (finding transformative fair use for AI training and digitizing purchased books, but not for retaining pirated copies); Kadrey v. Meta Platforms, Inc., No. 23-CV-03417-VC, 2025 WL 1752484 (N.D. Cal. June 25, 2025) (holding the use of copyrighted books for AI training to be a transformative fair use).

osAI ecosystem needs. Strong intellectual property policy that curbs the release of datasets improperly compiling copyrighted work can protect content creators, preserving the creativity and generativity that has produced substantial social value as well as bolstering their democratic control over how their work is used.

A foundational step is to provide legislative guidance on the application of the fair use doctrine to osAI, following the lead of jurisdictions that have already enacted clearer rules like the European Union,[290] Japan,[291] and Singapore.[292] Congress can direct courts on how to weigh the four factors for different AI components, reducing ambiguity and litigation risk. For instance, such guidance could instruct courts to give greater weight to the public benefit of transparency in components like system architecture, prompts, and operational records, while still allowing for robust copyright protection against the market harm caused by the open distribution of raw training data or model weights. This guidance must also look past superficial labels like "non-commercial." It should recognize that strategically open releases by large commercial actors, which are often designed to entrench market position,[293] are not equivalent to an osAI project from a university lab and should be treated differently in the analysis.

Beyond clarifying existing law, Congress can enact new statutory frameworks tailored to osAI's unique characteristics. An ambitious intervention would be to establish a compulsory licensing regime, akin to frameworks for digital music streaming,[294] requiring rights-holders to make their works available for AI training under standardized terms.[295] Licensing fees could be tiered, with low or no-cost licenses for academic and non-commercial open projects and higher rates for large, proprietary developers—including those that release some open components—thereby funding creators and leveling the competitive playing field.

Finally, procedural reforms can protect the osAI ecosystem from being weaponized by litigation itself. A tailored safe harbor could shield osAI developers from certain copyright claims when they openly release components that

---

[290] Directive (EU) 2019/790, arts. 3-4, 2019 O.J. (L 130) 92, https://eur-lex.europa.eu/eli/dir/2019/790/oj/eng.

[291] Chosakuken Hō [Copyright Act], Law No. 48 of 1970, art. 30-4 (Japan), https://www.cric.or.jp/english/clj/cl2.html.

[292] Copyright Act 2021 (No. 22 of 2021) § 244 (Sing.), https://sso.agc.gov.sg/Acts-Supp/22-2021/Published/?ProvIds=pr244-.

[293] *See supra* Part I.A.2.

[294] *See* 17 U.S.C. § 115; *see also* Mariana L. Orbay, *Songwriters v. Spotify: Is Spotify the Problem or a Symptom of the Problem?*, 48 PEPP. L. REV. 785, 796–804 (2021).

[295] *See, e.g.*, 88 Fed. Reg. 59942, 59947 (Aug. 30, 2023).

support auditing and reproducibility—such as architectural designs and documentation—provided they adhere to responsible practices.[296] A text-to-data-mining exception to copyright law could permit automated analyses of copyrighted work—but only by certain users for certain purposes, a form of selective openness.[297] Furthermore, fee-shifting provisions, requiring plaintiffs who bring unsuccessful infringement claims against certified non-commercial projects to cover legal costs, would deter frivolous lawsuits designed to drain the resources of smaller developers.[298]

A well-calibrated IP regime does not force a binary choice between protecting creators and enabling innovation; by applying legal standards to osAI with component-level precision, it can cultivate a safer and more competitive AI ecosystem.

D. Trade

Trade policies, particularly export controls, have become a primary lever for managing the national security risks of advanced osAI. The U.S. approach, however, is caught between conflicting goals: restricting adversaries' access to powerful technology while fostering the domestic innovation needed to maintain global leadership. The result is an unstable and often incoherent patchwork of rules targeting different osAI components—from permissive exemptions for open-weight models to shifting restrictions on compute hardware. Navigating a path forward requires analyzing the severe constitutional and strategic consequences of potential reforms, particularly broader restrictions on open model weights, and identifying a more precise, component-focused approach to the complex trade-off's of differential openness.

---

[296] *See* Eleonora Rosati, Eur. Parl., *The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market—Technical Aspects*, PE 604.942 (2018), https://www.europarl.europa.eu/RegData/etudes/BRIE/2018/604942/IPOL_BRI(2018)604942_EN.pdf (describing how text and data mining (TDM) exceptions to copyright laws can promote research and innovation)

[297] *See* Matthew Sag & Peter K. Yu, *The Globalization of Copyright Exceptions for AI Training,* 74 EMORY L. J. 1163 (2025).

[298] *See* 17 U.S.C. § 505 (permitting fee-shifting in copyright cases); *see also* David E. Shipley, *Discouraging Frivolous Copyright Infringement Claims: Fee Shifting under Rule 11 or 28 U.S.C. § 1927 as an Alternative to Awarding Attorney's Fees under Section 505 of the Copyright Act*, 24 J. INTELL. PROP. L. 33 (2016).

1. Baseline

The Commerce Department's Bureau of Industry and Security (BIS), which administers the Export Administration Regulations (EAR) for dual-use technologies—those with both civilian and military applications—first signaled its focus on trade restrictions for AI in January 2020 by controlling geospatial imagery AI software.[299] A more significant policy emerged from the January 2025 "AI diffusion" rule, which imposed export controls on certain advanced closed-weight models but explicitly exempted open-weight osAI models.[300]

This exemption was rooted in a strategic calculation: that the United States' primary competitive advantage lies in its dynamic, open source research community. Although some scholars have forewarned the existential or catastrophic risk of highly capable open-weight models in the hands of adversaries,[301] policymakers determined that the harm to innovation would be more damaging to national interests. For one, reputable researchers argue that there is no real evidence that today's open weight models increase an adversary's capacity to build, for example, a bioweapon.[302] Second, restricting the export of open-weight models would stifle the progress needed to outpace foreign osAI competitors like China's DeepSeek. The decision therefore reflects a deliberate trade-off, prioritizing domestic technological leadership over the risks of proliferation. The rise of powerful open models from China, such as Deepseek-R1 and Qwen3, suggest that the assumption that open models are less powerful than the proprietary counterparts may soon be obsolete.

For now, however, policymakers are turning to a more effective chokepoint to address osAI's national security concerns: the specialized hardware required

---

[299] Addition of Software Specially Designed to Automate the Analysis of Geospatial Imagery to the Export Control Classification Number 0Y521 Series, 85 Fed. Reg. 459 (Jan. 6, 2020); *see also* Kevin J. Wolf, *BIS Publishes a Temporary Unilateral Control on a Type of Machine Learning Software for Automating Analyses of Geospatial Imagery and Point Clouds*, AKIN GUMP (Jan. 6, 2020), https://www.akingump.com/en/insights/alerts/bis-publishes-new-0y521-control-on-certain-machine-learning.

[300] Framework for Artificial Intelligence Diffusion, 90 Fed. Reg. 4544, 4544 (Jan. 15, 2025).

[301] *See, e.g.*, SEGER ET AL., *supra* note 28; BATEMEN ET AL., *supra* note 160.

[302] CHRISTOPHER A. MOUTON ET AL., RAND, THE OPERATIONAL RISKS OF AI IN LARGE-SCALE BIOLOGICAL ATTACKS 1 (Jan. 25, 2024), https://www.rand.org/pubs/research_reports/RRA2977-2.html.

to train and run advanced models. The Biden administration's 2023 chip controls[303] and 2025 AI diffusion rule[304] both aimed to limit China's access to this critical compute infrastructure. This strategy, however, proved politically fragile. Criticism from U.S. allies and industry stakeholders, coupled with a shift in executive priorities, led the subsequent Trump administration not only to rescind the most stringent rules[305] but also to approve substantial compute-related investments in Saudi Arabia and the UAE, effectively undermining the hardware-centric control regime.[306]

Beyond export controls, policymakers have also explored new avenues to restrict the *import* of what they consider dangerous osAI models, citing safety and national security concerns, as illustrated by Senator Josh Hawley's proposed "DeepSeek ban" bill.[307] However, effective implementation of such a ban appears practically infeasible. Once AI model weights are disseminated online, controlling their spread becomes nearly impossible without resorting to highly invasive surveillance or censorship measures, exacerbating First Amendment concerns.[308]

2. Reforms

Current trade policy is in flux, but to the extent that it treats open-weight models more permissively than closed ones, it does so on the assumption that the former are less capable than the latter. But what should be done if and when

---

[303] Implementation of Additional Export Controls: Certain Advanced Computing Items; Supercomputer and Semiconductor End Use; Updates and Corrections, 88 Fed. Reg. 73458 (Oct. 25, 2023).

[304] *See* Framework for Artificial Intelligence Diffusion, 90 Fed. Reg. 4544 (Jan. 15, 2025); Implementation of Additional Due Diligence Measures for Advanced Computing Integrated Circuits, 90 Fed. Reg. 5298 (Jan. 16, 2025).

[305] Press Release, Bureau of Indus. & Sec., U.S. Dep't of Com., Department of Commerce Announces Recission of Biden-Era Artificial Intelligence Diffusion Rule, Strengthens Chip-Related Export Controls (May 13, 2025), https://www.bis.gov/press-release/department-commerce-announces-recission-biden-era-artificial-intelligence-diffusion-rule-strengthens-chip; *see also* David Sacks (@DavidSacks), X (May 9, 2025, 6:38 AM).

[306] *See* Tripp Mickle & Ana Swanson, *Outsourcer in Chief: Is Trump Trading Away America's Tech Future?*, N.Y. TIMES (May 15, 2025).

[307] Decoupling America's Artificial Intelligence Capabilities from China Act of 2025, S. 321, 119th Cong. (2025); *see also* Press Release, Sen. Josh Hawley, Hawley Introduces Legislation to Decouple American AI Development from Communist China (Jan. 29, 2025), https://www.hawley.senate.gov/hawley-introduces-legislation-to-decouple-american-ai-development-from-communist-china.

[308] *See, e.g.,* Reno v. ACLU*,* 521 U.S. 844, 853 (1997) ("The Web is thus comparable, from the readers' viewpoint, to both a vast library including millions of readily available and indexed publications and a sprawling mall offering goods and services.").

this assumption no longer holds? The most direct intervention would be to impose broad restrictions on any open-weight model that becomes as powerful as a closed alternative. However, such a move would not only have severe, counterproductive consequences but also face constitutional challenge.

First, such a policy would destroy the domestic osAI ecosystem. Any osAI project involving international collaborators would trigger "deemed export" rules, which treat the release of controlled technology to a foreign national as an export.[309] The ubiquitous global access granted by platforms like GitHub and Hugging Face would effectively outlaw open publication altogether, even just in the United States. While intended to curb proliferation, these sweeping measures would instead erode safety by suppressing the very transparency that enables independent auditing, red-teaming, and public scrutiny, while also greatly harming domestic research and innovation, much of which relies on the use of open models.[310]

Second, this approach would undermine American power and influence abroad. By making U.S.-led open innovation inaccessible, it would prevent smaller countries from adapting frontier technologies for their own critical needs, from public health in Lagos to crop optimization in Colombia. This could prove devastating to communities who stand to gain from these innovations; but it has geopolitical consequences as well. This would not only cede the technological high ground but also drive these nations toward alternative, and potentially adversarial, AI ecosystems, ultimately weakening the global network of democratic technology partners.

Finally, for the same constitutional reasons that apply to overbroad liability for osAI distribution,[311] any attempt to restrict the export of open-weight models would face formidable First Amendment barriers. It is notable that the *Bernstein* cases that established that source code constitutes protected speech did so specifically by striking down the very kind of export control regime that a ban on model weights would represent.[312] While it remains an open question whether courts would extend the same protection to non-human-readable open model weights, *Bernstein*'s logic poses serious constitutional questions

[309] 15 C.F.R. § 734.13(b) ("Any release in the United States of 'technology' or source code to a foreign person is a deemed export to the foreign person's most recent country of citizenship or permanent residency.").

[310] *See* John Villasenor, *The Tension Between AI Export Control and U.S. AI Innovation*, BROOKINGS INST. (Sep. 24, 2024), https://www.brookings.edu/articles/the-tension-between-ai-export-control-and-u-s-ai-innovation/.

[311] *See supra* notes 248–251 and accompanying text.

[312] Bernstein v. U.S. Dep't of State, 922 F. Supp. 1426, 1435 (N.D. Cal. 1996), *aff'd sub nom.* Bernstein v. U.S. Dep't of Justice, 176 F.3d 1132, 1135 (9th Cir.), reh'g granted, op. withdrawn, 192 F.3d 1308 (9th Cir. 1999).

for any policy that treats their publication as regulatable conduct rather than protected expression.

Ultimately, any trade policy aimed at modulating AI openness must walk a tightrope. A regime that balances these competing interests must approach differential openness with precision. Instead of a blunt ban on information, a more effective long-term strategy would involve tightening access to high-risk, physically controllable components like compute where restriction is enforceable, while preserving the information-sharing and collaborative infrastructure that enables a safe, innovative, and accountable open AI ecosystem.

E. Government Support

Strategic public investment can directly impact osAI's differential openness by supporting specific osAI components—especially compute, data, and labor—in ways that support safety research, innovation, and democratic accountability while strengthening U.S. global leadership.

1. Baseline

As discussed above, the economics of AI development inherently favor large, well-capitalized corporations that can afford the immense costs of compute and data acquisition and often lack the incentives to slow the release of insufficiently tested systems.[313] If policymakers decide to foster a more diverse osAI ecosystem, government support is a primary lever to counteract these market dynamics. Such support can be used to address the significant resource disparities in compute, data, funding, and expertise that place independent and non-commercial osAI development at a disadvantage. While there is ample declared support for direct government support for AI openness—the White House's AI Action Plan, released in July 2025, specifically calls for direct government investment in fostering AI openness[314]—there is little in the way of ongoing support in effect today.

2. Reforms

The most powerful government intervention would be the creation of a public option for osAI resources, establishing critical infrastructure to correct the market asymmetries that block participation from non-commercial actors. So far, government efforts have focused on strengthening the commercial mar-

[313] *See* Sharma, *supra* note 82.

[314] EXEC. OFF. PRES., *supra* note 1, at 4–5.

ketplace for domestic compute. For example, efforts like the CHIPS and Science Act subsidize the construction of semiconductor fabrication plants[315]—and thus democratize access to expensive compute hardware. The AI Action Plan doubles down on this, explicitly endorsing the approach of "improving the financial market for compute" to foster more open spectrum AI.[316] Strengthening the existing compute market, however, risks further solidification of incumbent dominance.

But the AI Action Plan also calls for a more disruptive investment[317]: the National AI Research Resource (NAIRR) program, which would "democratize access to AI resources" by providing researchers with high-performance computing and data that would otherwise be inaccessible.[318] But to be effective, the NAIRR must be designed as true government-run infrastructure—akin to a public supercomputer—rather than a subsidy program that funnels money through incumbent cloud providers. Such an approach would directly enable independent innovation and safety research while counteracting dependence on dominant vendors whose incentives may not align with the public interest.[319]

Just as critical as compute is data, and the disparity between open and proprietary developers is vast. Federal agencies house troves of high-quality datasets—ranging from scientific archives to public health records—that could serve as safer alternatives to the unvetted web-scraped corpora often used to train models.[320] By making these datasets machine-learning-ready, portable, and openly accessible, the government would begin to bridge the divide between proprietary and open developer access to data. To cultivate a sustainable ecosystem of public option data, research supported by government grants must also donate data in formats that meet the aforementioned criteria to this

---

[315] CHIPS and Science Act, Pub. L. No. 117-167, 136 Stat. 1366 (2022).

[316] EXEC. OFF. PRES., *supra* note 1, at 4.

[317] *Id.* at 5.

[318] *See* NAT'L A.I. RESCH. RES. PILOT (last visited Aug. 1, 2025), https://nairrpilot.org/; Madison Adler, *National Science Foundation Rolls Out NAIRR Pilot With Industry, Agency Support*, FEDSCOOP (Jan. 24, 2024), https://fedscoop.com/nsf-launches-nairr-pilot/ ("The pilot for the resource, referred to as the NAIRR, is composed of contributions from 11 federal agencies and 25 private sector partners, including Microsoft, Amazon Web Services, Nvidia, Intel, and IBM. Those contributions range from use of the Department of Energy's Summit supercomputer to datasets from NASA and the National Oceanic and Atmospheric Administration to access for models from OpenAI, Anthropic, and Meta.").

[319] *See* Narechania & Sitaraman, *supra* note 28, at 165–66.

[320] *See* Sean Long & Tom Romanoff, *AI-Ready Open Data*, BIPARTISAN POL'Y CTR. (Feb. 17, 2023), https://bipartisanpolicy.org/explainer/ai-ready-open-data/ ("Government's vast amount of open data can fill this gap: McKinsey estimates that open data can help unlock $3 trillion to $5 trillion in economic value annually").

corpus. This would also enhance auditability and reproducibility—key pillars of safety. To prevent this initiative from simply solidifying incumbent power, access could be selectively granted to noncommercial developers and researchers. While scholars rightly note the risks of public data containing misinformation or being used for abuse,[321] these concerns underscore the need for robust governance and proactive precautions—not inaction.

To serve democratic goals, open infrastructure must be more than just publicly funded; it must be publicly governed. Genuinely civic-oriented osAI infrastructure would institutionalize democratic oversight into how its components are built and used.[322] This requires transparent governance structures with representation from civil society, labor, and affected communities, not just corporate advisors.[323] It means public datasets are vetted for ethical use and compute resources are allocated based on social impact, not just institutional prestige. Governed this way, public infrastructure becomes a platform for societal problem-solving, enabling local governments and independent researchers to build osAI tools responsive to community priorities rather than commercial incentives.

Beyond direct investment, the government can use its power of the purse to create market demand for osAI through procurement policy. By incorporating requirements or preferences for transparency, interoperability, labor diversity, and safety in public-sector AI contracts, the government can provide a sustainable revenue stream for the osAI ecosystem. This approach also helps validate the reliability of osAI solutions for mission-critical applications, addressing a key challenge non-commercial projects face in establishing market credibility. As demonstrated with OSS, procurement preferences do not just fund open components; they set standards for the entire ecosystem.[324]

---

[321] *See, e.g.*, Chinmayi Sharma, Thomas E. Kadri & Sam Adler, *Brokering Safety*, 114 CALIF. L. REV. (forthcoming 2026), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5143114; Janet Freilich, *Government Misinformation Platforms*, 172 U. PA. L. REV. (2024).

[322] *See* Shearer, Davies & Lawrence, *supra* note 97..

[323] *See, e.g.*, Mark Coeckelbergh, *Artificial Intelligence, the Common Good, and the Democratic Deficit in AI Governance*, 5 AI ETHICS 1491 (2025), https://doi.org/10.1007/s43681-024-00492-9.

[324] *See, e.g.*, Iain G Mitchell, *Public Sector and Open Source*, *in* OPEN SOURCE LAW, POLICY AND PRACTICE 429 (Amanda Brock ed., 2022), https://doi.org/10.1093/oso/9780198862345.003.0021; Eunice Mercado-Lara, Shannon Dosemagen & Alison Parker, *Unlocking Innovation: Why Federal Procurement Should Embrace Open Source*, TECH POL'Y PRESS (May 23, 2025), https://www.techpolicy.press/unlocking-innovation-why-federal-procurement-should-embrace-open-source/.

Finally, government support is essential for fostering an open and diverse labor pool, which is itself a critical AI component. First, direct financial support for academic institutions, nonprofits, and smaller developers can drive non-commercial innovation and safety research outside of elite firms. (Notably, AI research is the only area where the Trump administration is not seeking to cut funding in its most recent NSF budget request.)[325] Structuring grants and contracts to prioritize inclusive hiring can create a more representative labor base, expanding the range of problems AI is designed to solve. Second, public investment in inclusive education and training programs—targeting underrepresented groups and supporting mid-career transitions—can correct the talent pipeline imbalances that currently concentrate power and perspective within a homogenous workforce.[326]

This effort to open the labor market must also address institutional and legal barriers. Elite institutions, academic and corporate alike, try to lock in expertise. The government can foster collaboration by funding cross-border initiatives and mandating adherence to open-science principles in the research it supports. Cross-border initiatives like ELIAS in Europe illustrate how formal partnerships, joint training programs, and open-access publishing can foster shared expertise and prevent intellectual concentration.[327]

But perhaps the most overlooked lever is immigration reform.[328] The U.S. currently risks ceding the global AI talent race to more nimble competitors through restrictive immigration policies for professionals and students.[329] Creating stable, inclusive visa pathways, especially for promising students and those working on open spectrum or public-benefit AI projects, would expand the diversity of contributors and directly counter the corporate and nationalist gatekeeping of AI expertise.

---

[325] *See* Nathan Lambert, *The White House's Plan for Open Models & AI Research in the U.S.*, INTERCONNECTS (July 23, 2025), https://www.interconnects.ai/p/the-white-houses-plan-for-open-models.

[326] *See, e.g.*, West, *supra* note 147 ("Only 19.3% of engineering grads [are] female.").

[327] *See About Elias*, ELIAS (last visited Aug. 1, 2025), https://elias-ai.eu/about/ ("ELIAS is a consortium of 34 top European institutions from 17 countries . . . committed to advancing fundamental research in AI.").

[328] *See* ZACHARY ARNOLD ET AL., CTR. SEC. & EMERGING TECH. AT GEO. UNIV., IMMIGRATION POLICY AND THE U.S. AI SECTOR 13 (2019), https://cset.georgetown.edu/publication/immigration-policy-and-the-u-s-ai-sector/.

[329] *See* Judy Wang & Nicol Turner Lee, *Trump's Immigration Policies May Threaten American AI Leadership*, BROOKINGS INST. (July 21, 2025), https://www.brookings.edu/articles/trumps-immigration-policies-may-threaten-american-ai-leadership.

Conclusion

The discourse around AI openness has profound implications for the future of technology, governance, and global power dynamics. Yet, as this Article has demonstrated, the conventional framing of AI openness as a natural extension of open source software invites a misleading binary characterization of AI as "open" or "closed" and an assumption that openness is an inherent good. This misunderstanding leads to regulatory approaches that are ill-equipped to govern AI, potentially stifling innovation, undermining accountability, or creating new security risks.

Because effective governance of osAI demands the utmost precision, this Article has proposed a more sophisticated approach by "unbundling" AI into its constituent components—compute, data, source code, model weights, and operational controls—and mapping each across its own spectrum of openness. In doing so, we introduce the concept of differential openness to capture the matrix of many possible permutations of component openness, each creating distinct risk profiles and governance challenges..

The different policy goals that motivate osAI regulation—such as safety, innovation, democratic access, and national security—often pull in contradictory directions, creating inevitable trade-offs. Openness is often a double-edged sword. For instance, releasing model weights can democratize access and spur innovation, but it also lowers the barrier for malicious actors and makes it impossible to recall a dangerous model once it has proliferated. The unbundling framework enables policymakers to make these trade-offs explicit and to calibrate regulatory approaches—whether through liability, competition policy, intellectual property, trade, and government support—to specific components rather than applying blunt, one-size-fits-all mandates.

AI is here to stay, and with it, openness. Policymakers cannot remain blithely unaware of osAI's complexity if we are to have any hope of shaping a future that is in the public's best interest.

Appendix: Openness of Select Frontier Models

This survey examines existing leading models and the openness of their components. All models with publicly available weights also disclose their architectures. "Disclosed training hardware" means the type of hardware used for training is known, even if access to that hardware is restricted—for example, by Nvidia. "Public inference hardware" indicates that the model can be run on commercially available systems. While "operational metadata" is a broad category, this classification limits it to the availability of ongoing audit logs and performance metrics. The row labeled "OSAID" refers not to a specific model, but to the Open Source Initiative's Open Source AI Definition.

Model	Training Data	Training Code	Training Hardware	Inference Code	Inference Hardware	Model Weights	Operational Metadata
OSAID[330]	Permissive License	Permissive License	Agnostic	Permissive License	Agnostic	Permissive License	Agnostic
Alibaba Qwen3[331]	Private	Private	Private	Public, Permissive License	Public	Public, Permissive License	Private
Anthropic Claude 4[332]	Private	Private	Private	Private	Private	Private	Private
Deepseek R1[333]	Private	Private	Disclosed	Public, Permissive License	Public	Public, Permissive License	Private
Google Gemini 2.5[334]	Private	Private	Private	Private	Private	Private	Private
Google Gemma 3[335]	Private	Private	Private	Public, Permissive License	Public	Public, Restricted License	Private

[330] *See* OPEN SOURCE INITIATIVE, *supra* note 114.

[331] *See* Qwen, *Qwen3*, HUGGING FACE, https://huggingface.co/collections/Qwen/qwen3-67dd247413f0e2e4f653967f (last visited Aug. 1, 2025).

[332] *See Introducing Claude 4*, ANTHROPIC (May 22, 2025), https://www.anthropic.com/news/claude-4.

[333] *See* deepseek-ai, *DeepSeek-R1*, HUGGING FACE, https://huggingface.co/deepseek-ai/DeepSeek-R1/ (last visited Aug. 1, 2025).

[334] *See Gemini 2.5 Pro Release Notes,* GOOGLE CLOUD, (last visited Aug. 1, 2025), https://cloud.google.com/vertex-ai/generative-ai/docs/models/gemini/2-5-pro

[335] *See* google, *Gemma 3 Release*, HUGGING FACE, https://huggingface.co/collections/google/gemma-3-release-67c6c6f89c4f76621268bb6d (last visited Aug. 1, 2025).

Model	Training Data	Training Code	Training Hardware	Inference Code	Inference Hard-ware	Model Weights	Operational Metadata
Meta Llama 4[336]	Private	Private	Disclosed	Public, Permissive License	Public	Public, Restricted License	Private
Nvidia Nemotron 4-340b	Partially Open	Permissive License	Disclosed	Public, Permissive License	Public	Public, Permissive License	Private
OpenAI o4-mini	Private	Private	Private	Private	Private	Private	Private
xAI Grok 1	Private	Private	Private	Public, Permissive License	Public	Public, Permissive License	Private
xAI Grok 3	Private	Private	Private	Private	Private	Private	Private

[336] *See* meta-llama, *Llama 4*, HUGGING FACE, https://huggingface.co/collections/meta-llama/llama-4-67f0c30d9fe03840bc9d0164 (last visited Aug. 1, 2025).