Automated Compliance and the Regulation of AI

LawAI Working Paper Series, No. 1-2026
Cullen O'Keefe & Kevin Frazier

January 2026

law-ai.org

Automated Compliance and the Regulation of AI

Cullen O'Keefe* & Kevin Frazier†

Abstract

Regulation imposes compliance costs on regulated parties. Thus, policy discourse often posits a trade-off between risk reduction and innovation. Without denying this trade-off outright, this paper complicates it by observing that, under plausible forecasts of AI progress, future AI systems will be able to perform many compliance tasks cheaply and autonomously. We call this *automated compliance*. While automated compliance has important implications in many regulatory domains, it is especially important in the ongoing debate about the optimal timing and content of regulations targeting AI itself. Policymakers sometimes face a trade-off in AI policy between potentially regulating too soon or strictly (and thereby stifling innovation and national competitiveness) versus too late or leniently (and thereby risking preventable harms). Under plausible assumptions, automated compliance loosens this trade-off: AI progress itself hedges the costs of AI regulation. Automated compliance implies that, for example, policymakers could reduce the risk of premature regulation by enacting regulations that become effective only when AI is capable of largely automating compliance with such regulations. This regulatory approach would also mitigate concerns that regulations may unduly benefit larger firms, which can bear compliance costs more easily than startups can. While regulations can remain costly even after many compliance tasks have become automated, we hope that the concept of automated compliance can enable a more multidimensional and dynamic discourse around the optimal content and timing of AI risk regulation.

Policy Implications

- AI progress will reduce the cost of certain regulatory compliance tasks.
- Policymakers can reduce risks from premature regulation by using "automatability triggers" to stipulate that regulations become effective only when AI tools are capable of automating compliance with such regulations.
- Proper implementation of compliance-automating AI workflows may provide evidence of regulatory compliance.
- Targeted measures can support the development of compliance-automating AI services.
- Automated compliance is most effective when complemented by the responsible automation of regulator-side regulatory and oversight tasks.

Table of Contents

Introduction

Few contest that rapid advances in artificial intelligence (AI) capabilities and adoption will require regulatory intervention. Instead, some of the deepest disagreements in AI policy concern the general timing, substance, and purpose of those regulations. The stakes, most agree, are high.[1] Those more concerned with risks from AI worry about, for example, risks from misuse of AI systems to make weapons of mass destruction,[2] from strategic destabilization,[3] and from loss of control of advanced AI systems that are not aligned with humanity.[4] In turn, some individuals with this perspective have called for a more aggressive regulatory posture aimed at reducing the odds of worst-case scenarios.[5] Those who are focused on the benefits of AI systems point out that there is a large amount of uncertainty as to the likelihood of these risks and regarding best practices for risk mitigation.[6] They also champion the potential of such systems to drive innovations in medicine and other areas of science,[7] to supercharge economic growth more generally,[8] and to enable strategic applications that could determine the balance of global power.[9] The regulatory posture of these individuals tends to be more hands-off out of a concern for early

[1] *But cf.* Arvind Narayanan & Sayash Kapoor, *AI as Normal Technology*, KNIGHT FIRST AMEND. INST. (Apr. 15, 2025), https://knightcolumbia.org/content/ai-as-normal-technology [https://perma.cc/3RTP-5MTC] (offering a more modest appraisal of AI's likely impacts).

[2] *See, e.g.,* Oliver Meier, *The Fast and the Deadly: When Artificial Intelligence Meets Weapons of Mass Destruction*, EUR. LEADERSHIP NETWORK (June 27, 2024), https://europeanleadershipnetwork.org/commentary/the-fast-and-the-deadly-when-artificial-intelligence-meets-weapons-of-mass-destruction/ [https://perma.cc/DB4F-RHD5].

[3] *See, e.g.*, THE IMPACT OF ARTIFICIAL INTELLIGENCE ON STRATEGIC STABILITY AND NUCLEAR RISK, VOLUME I, EURO-ATLANTIC PERSPECTIVES (Vincent Boulanin ed., 2019), https://www.sipri.org/publications/2019/research-reports/impact-artificial-intelligence-strategic-stability-and-nuclear-risk-volume-i-euro-atlantic [https://perma.cc/8YDP-HWXU].

[4] *See, e.g.,* ELIKA SOMANI ET AL., STRENGTHENING EMERGENCY PREPAREDNESS AND RESPONSE FOR AI LOSS OF CONTROL INCIDENTS (2025), https://www.rand.org/pubs/research_reports/RRA3847-1.html [https://perma.cc/3J4U-YPFP].

[5] *See, e.g.,* Scott Babwah Brennan, *New York's RAISE Act Authors Give Their Pitch as Gov. Hochul Mulls over AI Bill*, TRANSPARENCY COAL. (July 11, 2025), https://www.transparencycoalition.ai/news/new-yorks-raise-act-authors-give-their-pitch-as-gov-hochul-considers-the-bill [https://perma.cc/GWB5-NQDQ].

[6] *See, e.g.*, RISHI BOMMASANI ET AL., THE CALIFORNIA REPORT ON FRONTIER AI POLICY (2025), https://www.gov.ca.gov/wp-content/uploads/2025/06/June-17-2025-%E2%80%93-The-California-Report-on-Frontier-AI-Policy.pdf [https://perma.cc/S36U-J6EV] (identifying areas of disagreement among AI experts as to AI capabilities, risks, and mitigation strategies).

[7] *See, e.g.,* Tal Roded & Peter Slattery, *AI and the Future of Scientific Discovery*, MIT FUTURETECH (Apr. 29, 2025), https://futuretech.mit.edu/news/ai-and-the-future-of-scientific-discovery [https://perma.cc/YJC5-SWNF].

[8] *See, e.g.,* Martin Neil Baily & Aidan T. Kane, *Harnessing AI for Economic Growth*, BROOKINGS (Apr. 8, 2025), https://www.brookings.edu/articles/harnessing-ai-for-economic-growth/ [https://perma.cc/WTS8-SL4V].

[9] *See, e.g.*, BARRY PAVEL ET AL., AI AND GEOPOLITICS: HOW MIGHT AI AFFECT THE RISE AND FALL OF NATIONS? (2023), https://www.rand.org/pubs/perspectives/PEA3034-1.html [https://perma.cc/A7XS-H6LN].

regulations entrenching existing actors and perhaps leading to technological path dependence.[10]

Both perspectives simultaneously find some support from the observed characteristics of currently deployed AI systems[11] but are also necessarily based on a forecast of a large number of uncertain variables, including the trajectories of AI capabilities, societal adaptation, international relations, and public policy. As a result of these disagreements and uncertainties, perspectives on the appropriate course of action range widely, from, at one pole, unilateral or coordinated "pausing" of AI progress,[12] to, at the other pole, deregulation and acceleration of AI progress.[13] We might call this the *proregulatory–deregulatory divide*.[14]

The foregoing is an oversimplified sketch of what is in fact a much more multidimensional debate. Innovation and regulation are not always zero-sum.[15] People—including both authors of this article—tend to hold a combination of proregulatory and deregulatory views depending on the exact AI policy issue. Many policies are preferable

[10] *See* Yiqin Fu & Alasdair Phillips-Robins, *When Should Congress Preempt State AI Law? The Lessons of Past Technologies*, CARNEGIE ENDOWMENT FOR INT'L PEACE (Sept. 9, 2025), https://carnegieendowment.org/research/2025/09/congress-preempt-state-ai-law-the-lessons-of-past-technologies [https://perma.cc/SL93-RXZF] (introducing concerns about "lock in" induced by regulation); *cf.* Stephen Redding, *Path Dependence, Endogenous Innovation, and Growth*, 43 INT'L ECON. REV. 1215 (2002) (providing historical and economic overview of path dependence as it relates to the development and diffusion of new technology).

[11] For empirical support for claims that future AI systems may pose a large risk, see, for example, Ryan Greenblatt et al., Alignment Faking in Large Language Models (Dec. 20, 2024) (unpublished manuscript), http://arxiv.org/abs/2412.14093; Leonard Dung, *Current Cases of AI Misalignment and Their Implications for Future Risks*, 202 SYNTHESE 138 (2023). For research indicating that current AI systems are not yet capable of posing large-scale risks, see, for example, CHRISTOPHER A. MOUTON, CALEB LUCAS & ELLA GUEST, THE OPERATIONAL RISKS OF AI IN LARGE-SCALE BIOLOGICAL ATTACKS: RESULTS OF A RED-TEAM STUDY (2024), https://www.rand.org/pubs/research_reports/RRA2977-2.html; Seth D. Baum, *Assessing the Risk of Takeover Catastrophe from Large Language Models*, 45 RISK ANALYSIS 752 (2025).

[12] *E.g.*, *Pause Giant AI Experiments: An Open Letter*, FUTURE OF LIFE INST. (Mar. 22, 2023), https://futureoflife.org/open-letter/pause-giant-ai-experiments/ [https://perma.cc/TH52-JCJC]; Peter Barnett, Aaron Scher & David Abecassis, Technical Requirements for Halting Dangerous AI Activities (July 13, 2025) (unpublished manuscript), http://arxiv.org/abs/2507.09801.

[13] *E.g.*, Exec. Order No. 14,179, 90 Fed. Reg. 8741 (Jan. 23, 2025); Kevin Frazier & Adam Thierer, *1,000 AI Bills: Time for Congress to Get Serious about Preemption*, LAWFARE (2025), https://www.lawfaremedia.org/article/1-000-ai-bills--time-for-congress-to-get-serious-about-preemption [https://perma.cc/LWK7-UH3R].

[14] In casual discourse, these factions are sometimes called "accelerationists" and "doomers," respectively. *E.g.*, Gabriel Weil, Instrument Choice in AI Governance: Liability as the Indispensable Core 9–10 (June 5, 2025) (unpublished manuscript), https://papers.ssrn.com/abstract=5283275. Other related distinctions are permissive versus precautionary approaches to new technologies, *see* Rebecca Crootof & BJ Ard, *Structuring Techlaw*, 34 HARV. J.L. & TECH. 347, 379–86 (2021), and dynamism versus stasis, *see* Helen Toner, *We're Arguing about AI Safety Wrong*, AI FRONTIERS (May 12, 2025), https://ai-frontiers.org/articles/were-arguing-about-ai-safety-wrong [https://perma.cc/GM5P-QLYT].

[15] *See, e.g.*, Anu Bradford, *The False Choice between Digital Regulation and Innovation*, 119 NW. U. L. REV. 377 (2024); Maarten Herbosch, *Beyond the False Dichotomy: Regulating AI Safety, Ethics and Innovation*, ARIZONA ST. L.J. (forthcoming 2026), https://papers.ssrn.com/abstract=5640031.

under both views; other policies complicate or straddle the divide. Others question whether proregulation versus deregulation is even the right frame for this debate at all.[16] Semantics aside, the fundamental tension between proregulatory and deregulatory approaches to AI policy is, in many cases, real. Indeed, at the risk of oversimplifying our own views, the authors generally locate themselves on opposite sides of the proregulatory–deregulatory divide.[17]

There are many causes of the proregulatory–deregulatory divide, including empirical disagreements about the likely impacts of future AI systems and normative disagreements about how to value different policy risks or outcomes. We do not attempt to resolve these here. We do, however, note one reason to think that the trade-offs between the proregulatory and deregulatory approaches may not remain as harsh as they presently seem: future AI systems will likely be able to automate many compliance tasks, including many required by AI regulations.[18] We can call these *compliance-automating AIs*. Compliance-automating AIs will be able to, for example, perform automated evaluations of AI systems, compile transparency reports of AI systems' performance on those evaluations, monitor for safety and security incidents, and provide incident disclosures to regulators and consumers.[19]

Compliance-automating AIs deserve a larger role in the discussion of regulation across all economic sectors.[20] But their implications for *regulation of AI itself* warrant special attention. This is for two reasons. First, as detailed above, the stakes of AI policy are immense, with weighty considerations on both sides of the proregulatory–deregulatory divide. Compliance-automating AI will affect the costs of AI regulation and therefore be an increasingly important input into the overall cost-benefit analysis of proposed regulations of this important sector. Second, AI policy is unique in that compliance-automating AI can both influence and be influenced by AI policy. As mentioned, the availability of compliance-automating AI will influence the cost-benefit profile of many AI policies, and therefore, one hopes, whether and when such policies are implemented. But AI policies, in turn, will also influence whether and when compliance-automating AI is developed. The interaction between compliance-automating AI and policy is therefore much more complicated in AI than in other policy domains.

---

[16] *See, e.g.*, Helen Toner, *In Search of a Dynamist Vision for Safe Superhuman AI*, RISING TIDE (May 12, 2025), https://helentoner.substack.com/p/dynamism-vs-stasis [https://perma.cc/U8BG-E6XK].

[17] *Compare, e.g.*, Kevin Frazier, *True Confessions of an AI Flip Flopper*, APPLESEED AI (July 14, 2025), https://appleseedai.substack.com/p/true-confessions-of-an-ai-flip-flopper [https://perma.cc/GQR4-B2YA], *with* Markus Anderljung et al., Frontier AI Regulation: Managing Emerging Risks to Public Safety (Nov. 7, 2023) (unpublished manuscript), http://arxiv.org/abs/2307.03718.

[18] *Cf.* BOMMASANI ET AL., *supra* note 6, at 34 (raising the possibility of automated data processing to lower administrative and operational burdens associated with regulation).

[19] *See infra* Part III.

[20] *See* Paul Ohm, *Toward Compliance Zero: AI and the Vanishing Costs of Regulatory Compliance*, NETWORK L. REV. (Sept. 2, 2025), https://www.networklawreview.org/ohm-ai-regulation/ [https://perma.cc/R8YX-VKNF].

This paper proceeds as follows. In Part I, we briefly note the potentially high costs associated with regulatory compliance. In Part II, we survey how current AI policy proposals attempt to manage compliance costs. In Part III, we introduce the concept of *automated compliance*: a prediction that, as AI capabilities advance, AI systems will themselves be capable of automating some regulatory compliance tasks, leading to reduced compliance costs.[21] In Part IV, we note several implications of automated compliance for the design of AI regulations. First, we propose the novel concept of *automatability triggers*: regulatory mechanisms that specify that AI regulations become effective only when automation has reduced the costs to comply with regulations below a predetermined level. Second, we note how AI policies could use automated compliance as evidence of compliance. Third, we identify several tasks policymakers, entrepreneurs, and civic technologists could take to accelerate automated compliance. Finally, we note the possible synergies between automated compliance and automated *governance*.

I. Regulatory Costs

Regulation can be very costly to the regulated party, the regulator, and (therefore) the economy as a whole. While a holistic survey of the potential costs associated with regulation is well beyond the scope of this paper,[22] we briefly note some data points indicating the potential costs of regulation.

Compliance costs are perhaps the most easily observable costs. By "compliance costs," we mean "the costs that are incurred by businesses . . . at whom regulation may be targeted in undertaking actions necessary to comply with the regulatory requirements, as

[21] During the writing of this paper, Professor Paul Ohm released an excellent short essay making similar points to many of those in this paper, especially those in Part III. *See* Ohm, *supra* note 20. Professor Ohm argues that because "AI automation will drive the cost of regulatory compliance down so close to zero" we should consider "that, going forward, . . . most regulations [will] impose no compliance burden whatsoever." *Id.* While we agree with many of Professor Ohm's claims, our thesis diverges from his in several respects. First, ours is attentive to a sequencing problem that Professor Ohm largely does not address. While we share Professor Ohm's optimism about the potential for AI systems to reduce compliance costs, some of the trickiest questions in AI policy concern the optimal timing of regulation: specifically, whether to regulate AI *before* AI advances to the point where compliance work is largely automated. *See infra* Section IV.A (discussing the respective risks of regulating too early versus too late). Since compliance costs have not yet reached zero, the projection that they will soon do so bears less on the optimal regulatory approach in the interim, even if that projection is correct. Second, while Professor Ohm notes that there will be "compliance work that will indeed be difficult to automate," *see* Ohm, *supra* note 20, it seems that we expect that such "irreducible costs," *id.*, will be larger than Professor Ohm expects. *Compare infra* Part III (detailing compliance costs that automation will not largely eliminate), *with* Ohm, *supra* note 20 ("The burden of regulatory compliance will soon be nothing but a bad memory of days past, a story the old-timers tell the young ones about how the law used to operate").

[22] For one overview, see OECD, OECD REGULATORY COMPLIANCE COST ASSESSMENT GUIDANCE (2014), https://www.oecd.org/en/publications/oecd-regulatory-compliance-cost-assessment-guidance_9789264209657-en.html [https://perma.cc/CJH3-VDCE].

well as the costs to government of regulatory administration and enforcement."[23] This includes both administrative costs (e.g., the costs associated with producing and processing paperwork required by regulation) and substantive costs (e.g., the costs associated with reengineering a product to comply with substantive standards imposed by regulation).[24]

To take just a few examples:

- The state of California estimated that initial compliance costs for the California Consumer Privacy Act totaled approximately $55 billion.[25]
- A large share of the increased costs of generating nuclear power is due to increased (and changing) safety regulations.[26] For example, quality control and quality assurance requirements cause there to be a "nuclear premium" for commodity components of nuclear power plants.[27] The nuclear premium has been estimated at 23% of the cost of concrete and 41% of the cost of steel in nuclear plants.[28]
- Financial crime compliance costs in the US and Canada are estimated at $61 billion,[29] becoming major cost centers for financial institutions.[30]

[23] *Id.* at 12.

[24] *See id.* at 12–13.

[25] DAVID ROLAND-HOLST ET AL., STANDARDIZED REGULATORY IMPACT ASSESSMENT: CALIFORNIA CONSUMER PRIVACY ACT OF 2018 REGULATIONS 11 (2019), https://dof.ca.gov/media/docs/forecasting/economics/major-regulations/major-regulations-table/CCPA_Regulations-SRIA-DOF.pdf [https://perma.cc/N64S-DQEQ].

[26] *See* Brian Potter, *Why Does Nuclear Power Plant Construction Cost so Much?*, INST. FOR PROGRESS (May 1, 2023), https://ifp.org/nuclear-power-plant-construction-costs/ [https://perma.cc/JJR2-WKXW] (citing CHARLES KOMANOFF, POWER PLANT COST ESCALATION: NUCLEAR AND COAL CAPITAL COSTS, REGULATION, AND ECONOMICS (1981); Soon Paik & William R. Schriver, *The Effect of Increased Regulation on Capital Costs and Manual Labor Requirements of Nuclear Power Plants*, 26 ENG. ECON. 223 (1980); Robert A. Szalay, *Stability in Licensing Requirements: A Technical Perspective*, INT'L CONF. ON REGULATING NUCLEAR ENERGY (1978), https://inis.iaea.org/records/bzpza-3pn83/preview/10421637.pdf [https://perma.cc/KA8N-N7QE]; Nathan Hultman & Jonathan Koomey, *Three Mile Island: The Driver of US Nuclear Power's Decline?*, 69 BULL. ATOMIC SCIENTISTS 63 (2013)).

[27] *See id.*

[28] *See* KAREN DAWSON & PIYUSH SABHARWALL, A REVIEW OF LIGHT WATER REACTOR COSTS AND COST DRIVERS iii–iv (2017), http://www.osti.gov/servlets/purl/1466793/ [https://perma.cc/PSM7-WGGQ].

[29] *See* Ade O'Connor, *Study Reveals Annual Cost of Financial Crime Compliance Totals $61 Billion in the United States and Canada*, LEXISNEXIS RISK SOLUTIONS (Feb. 21, 2024), https://risk.lexisnexis.com/about-us/press-room/press-release/20240221-true-cost-of-compliance-us-ca [https://perma.cc/P6WK-RJTP].

[30] *See* Lanier Saperstein, Geoffrey Sant & Michelle Ng, *The Failure of Anti-Money Laundering Regulation: Where Is the Cost-Benefit Analysis?*, 91 NOTRE DAME L. REV. 1, 4–5 (2015) ("HSBC recently estimated it now devotes $750 million to $800 million per year on compliance [with financial crime regulation]—an amount equivalent to one quarter of the operating budget of its entire U.S. operations.").

- One paper estimates that "[a]n average firm spends 1.34 percent of its total labor costs on performing regulation-related tasks."[31]

Regulations also impose costs on the government, such as the costs associated with promulgating regulation, bringing enforcement actions, and adjudicating cases.[32] California Governor Gavin Newsom recently vetoed[33] AB 1064, which would have regulated companion chatbots made available to minors.[34] A previous version of the bill would have established a Kids Standards Board,[35] which would have cost the state "between $7.5 million and $15 million annually."[36] To take another infamous example, consider the National Environmental Policy Act (NEPA),[37] which, inter alia, requires the federal government to prepare an environmental impact statement (EIS) prior to "major Federal actions significantly affecting the quality of the human environment."[38] These EISs have grown to become incredibly burdensome, with a 2014 report from the Government Accountability Office estimating that the average EIS took 1,675 days to complete[39] and cost between $250,000 and $2 million.[40]

It is also worth considering the *opportunity costs* associated with regulation. Regulation requires firms to divert resources away from their most productive use.[41] The opportunity cost of a regulation to a firm is therefore the difference between the return the firm would have earned from its most productive use of resources dedicated to compliance and the return those resources in fact earned.[42] While more difficult to observe than compliance costs, opportunity costs may be significantly larger over time due to compounding growth.[43] Opportunity costs also include the cost to consumers when

[31] Francesco Trebbi & Miao Ben Zhang, *The Cost of Regulatory Compliance in the United States* 2 (Nat'l Bureau of Econ. Rsch., Working Paper No. 30691, 2022), https://www.nber.org/papers/w30691 [https://perma.cc/T8B7-LJQY].

[32] *See* OECD, *supra* note 22, at 12 (including "the costs to government of regulatory administration and enforcement" as part of compliance costs).

[33] Letter from Governor Newsom to the California State Assembly Vetoing AB1064 (Oct. 13, 2025), https://www.gov.ca.gov/wp-content/uploads/2025/10/AB-1064-Veto.pdf [https://perma.cc/EWP5-TQRX].

[34] A.B. 1064, 2025-2026 Reg. Sess. (Cal. 2025), https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202520260AB1064 [https://perma.cc/LHR4-CTB5].

[35] *See id.* § 22757.22 (May 1, 2025 Amended Assemb. version).

[36] ANNIKA CARLSON, ANALYSIS OF AB-1064, 2025-2026 Reg. Sess., at 2 (Cal. 2025), https://perma.cc/R32J-BBDF.

[37] National Environmental Policy Act of 1969, 42 U.S.C. § 4321 *et seq*.

[38] 42. U.S.C. § 4332(2)(C).

[39] *See* U.S. GOV'T ACCOUNTABILITY OFF., GAO-14-370, NATIONAL ENVIRONMENTAL POLICY ACT: LITTLE INFORMATION EXISTS ON NEPA ANALYSES 14 (2014), https://www.gao.gov/assets/gao-14-370.pdf [https://perma.cc/GND6-G9NF].

[40] *See id.* at 13.

[41] *See* OECD, *supra* note 22, at 15.

[42] *See id.* at 15.

[43] John W. Dawson & John J. Seater, *Federal Regulation and Aggregate Economic Growth*, 18 J. ECON. GROWTH 137, 161 (2013).

regulations prevent a product from reaching the market. For example, a working paper found that the European Union's (EU's) General Data Protection Regulation (GDPR) "induced the exit of about a third of available apps" on the Google Play Store, ultimately reducing consumer surplus in the app market by about a third.[44]

Finally, regulation can have strategic costs not easily captured in economic terms. A number of commentators worry that overregulation of AI in the US could cause the US to lose its lead in AI research and development to foreign competitors, especially China.[45] Of course, China for its part has hardly taken a laissez-faire attitude toward its domestic AI industry.[46] Nevertheless, it is reasonable to carefully consider international competitiveness when evaluating domestic regulatory proposals.

To be clear, this paper does not argue that the potentially high costs of regulation are a conclusive argument against any particular regulatory proposal. Regulations often have pro tanto benefits, and such benefits must be weighed against costs to decide whether the regulation is socially beneficial on net.[47] Nor is this intended to be a comprehensive survey of regulatory burdens. We are merely restating the banal observation that regulations can come at significant cost to societal goods. Any serious discussion of AI policy must be willing to concede that point and entertain approaches to capturing the benefits of well-designed regulation at lower cost to producers, consumers, and society as a whole.

---

[44] *See* Rebecca Janßen et al., *GDPR and the Lost Generation of Innovative Apps* (Nat'l Bureau of Econ. Rsch., Working Paper No. 30028, 2022), https://www.nber.org/papers/w30028 [https://perma.cc/ZC2F-G4P5]. For more estimates of costs from GDPR, see Mert Demirer et al., *Data, Privacy Laws and Firm Production: Evidence from the GDPR* (Nat'l Bureau of Econ. Rsch., Working Paper No. 32146, 2024), https://www.nber.org/papers/w32146 [https://perma.cc/TC6F-J8CC]; *The Internet and Digital Communications: Examining the Impact of Global Internet Governance: Hearing Before the Subcomm. on Commc'ns, Tech., Innovation, & the Internet of the S. Comm. on Com., Sci., & Transp.*, 115th Cong. 35 (2018) (statement of Denise E. Zheng, Vice President, Policy, Business Roundtable), https://www.commerce.senate.gov/services/files/A741272F-5789-48DC-9C06-DDB5F66F4321 [https://perma.cc/Q8ZW-YSYP].

[45] *See, e.g.*, *AI Regulation and the Future of US Leadership: Hearing Before the Subcomm. on Com., Mfg., & Trade of the Comm. On Energy & Com.*, 119th Cong. (2025) (statement of Adam Thierer, Resident Senior Fellow, Tech. & Innovation, R Street Inst.), https://democrats-energycommerce.house.gov/sites/evo-subsites/democrats-energycommerce.house.gov/files/evo-media-document/thierer_testimony.pdf [https://perma.cc/7YL7-RDFV]; Keegan McBride, *The Threat of "AI Safety" to American AI Leadership*, NAT'L INT. (Apr. 28, 2024), https://nationalinterest.org/blog/techland/threat-ai-safety-american-ai-leadership-210780/ [https://perma.cc/59Q4-QG3X].

[46] *See, e.g.*, Meaghan Tobin, *China Announces Rules to Keep AI Bound by 'Core Socialist Values,'* WASH. POST, July 14, 2023, https://www.washingtonpost.com/world/2023/07/14/china-ai-regulations-chatgpt-socialist/ [https://perma.cc/EA33-VUYM].

[47] *See generally, e.g.*, Cass R. Sunstein, *The Cost-Benefit State* (Coase-Sandor Inst. for L. & Econ. Working Paper No. 39, 1996), https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1497&context=law_and_economics [https://perma.cc/V22H-4WNA].

II. Current Approaches to Managing Compliance Costs in AI Policy

Regulators and policy entrepreneurs often make some efforts to reduce the costs associated with regulation. Traditional regulatory literature often proposes using *performance-based* regulation, which requires regulated parties to achieve certain results rather than use certain methods or technologies, on the logic that performance-based standards allow the regulatees to find and adopt more efficient methods of achieving compliance.[48] Tort-based approaches to AI policy[49] similarly incentivize firms to identify and implement the most effective means for reducing actionable harms from their systems.[50]

Nevertheless, there remains substantial interest in prescriptive regulation for AI.[51] To date, the most carefully designed proregulatory proposals have tended to use some method for *regulatory targeting* to attempt to limit regulatory costs to only those firms that are both (a) engaging in the riskiest behaviors and (b) best able to bear compliance costs. Early proposals tended to rely on *compute thresholds*, wherein only AI models that were trained using more than a certain number of computational operations would be regulated.[52] A number of proposed and enacted laws and regulations have used compute thresholds for this reason:

[48] *See, e.g.*, Cary Coglianese, *The Limits of Performance-Based Regulation*, 50 U. MICH. J.L. REFORM 525, 526–28 (2017) (collecting sources in favor of performance-based standards, then later arguing that performance-based regulation has underappreciated limitations).

[49] *E.g.*, Weil, *supra* note 14; Gabriel Weil, Tort Law as a Tool for Mitigating Catastrophic Risk from Artificial Intelligence (Jan. 13, 2024) (unpublished manuscript), https://papers.ssrn.com/abstract=4694006.

[50] *See* Weil, *supra* note 14, at 12–14.

[51] For one argument why liability-based approaches will be inadequate, see Daniel Schwarcz & Josephine Wolff, *The Limits of Regulating AI Safety through Liability and Insurance: Lessons from Cybersecurity* (Minnesota Legal Studies Research Paper 2025-46, 2025), https://papers.ssrn.com/abstract=5411062. On the difficulties in implementing performance-based regulation for AI, see, for example, Rachel L. Thomas & David Uminsky, *Reliance on Metrics Is a Fundamental Challenge for AI*, 3 PATTERNS 100476 (2022).

[52] *E.g.*, Anderljung et al., *supra* note 17, at 35–36; Matteo Pistillo et al., *The Role of Compute Thresholds for AI Governance*, 1 GEO. WASH. J.L. & TECH. 26 (2025); Lennart Heim & Leonie Koessler, Training Compute Thresholds: Features and Functions in AI Regulation (Aug. 6, 2024) (unpublished manuscript), http://arxiv.org/abs/2405.10799; Girish Sastry et al., Computing Power and the Governance of Artificial Intelligence 19–33 (Feb. 13, 2024) (unpublished manuscript), http://arxiv.org/abs/2402.08797; *White House Policy on AI: Hearing Before the Subcomm. on Cybersecurity, Info. Tech., & Gov't Innovation of the H. Comm. on Oversight & Gov't Reform*, 118th Cong. 2 (2023) (statement of Samuel Hammond, Senior Economist, Found. for Am. Innovation), https://oversight.house.gov/wp-content/uploads/2023/12/Testimony-Hammond.pdf [https://perma.cc/P8UA-MQYJ]. *But see infra* notes 65–66 and accompanying text (discussing how reasoning models and other AI developments have complicated the relationship between compute usage and AI capabilities).

- The EU AI Act states that "A general-purpose AI model shall be presumed to have high impact capabilities . . . when the cumulative amount of computation used for its training measured in floating point operations is greater than 10^{25}."[53]
- Regulations proposed under the Biden administration would have required AI companies to report the development of "dual-use foundation models," which were in turn defined as models that "utilize[] more than 10^{26} computational operations (e.g., integer or floating-point operations)."[54]
- The much-debated—and ultimately vetoed[55]—California SB 1047[56] would have initially[57] applied to models that were "trained using a quantity of computing power greater than 10^{26} integer or floating-point operations, the cost of which exceeds one hundred million dollars . . . ."[58]
- California's recently enacted SB 53[59]—a successor to SB 1047 focused on transparency—defines "frontier model" as "a foundation model that was trained using a quantity of computing power greater than 10^{26} integer or floating-point operations."[60]

Compute-based thresholds are a reasonable proxy for the financial and operational resources needed to comply with regulation because compute (in the amounts typically proposed) is expensive: when proposed in 2023, the 10^{26} operations threshold

[53] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), art. 51(2), 2024 O.J. (L. 1689) 83 [hereinafter EU AI Act].

[54] Establishment of Reporting Requirements for the Development of Advanced Artificial Intelligence Models and Computing Clusters, 89 Fed. Reg. 73612, 73615 (Sept. 11, 2024).

[55] Letter from Cal. Governor Gavin Newsom to the Cal. State Senate (Sept. 29, 2024), https://www.gov.ca.gov/wp-content/uploads/2024/09/SB-1047-Veto-Message.pdf [https://perma.cc/JPN7-FRHV].

[56] Safe and Secure Innovation for Frontier Artificial Intelligence Models Act, S.B. 1047, 2023-2024 Reg. Sess. (Cal. 2024) (as enrolled, Sept. 3, 2024), https://calmatters.digitaldemocracy.org/bills/ca_202320240sb1047 [https://perma.cc/H8XW-KGFW].

[57] From January 1, 2027, onwards, the Government Operations Agency could have changed the operations-based compute threshold, but the $100 million dollar threshold would have remained, *see id.* sec. 3, § 22602(e)(1)(B)(i)(I), after being adjusted for inflation, *see id.* sec. 3 § 22602(e)(2).

[58] *Id.* sec. 3, § 22602(e)(1)(A)(i). The law would have also initially applied to models that were fine-tuned from an otherwise-covered model when such fine-tuning used $3 * 10^{25}$ operations and cost more than $10 million. *Id.* sec. 3, § 22602(e)(1)(A)(ii). As with the primary definition, the operations-based threshold could have been changed by the Government Operations Agency from January 1, 2027, onwards, but the dollar-based threshold would have remained (after being adjusted for inflation). *Id.* sec. 3, § 22602(e)(1)(B)(i)(II), (e)(2).

[59] *Governor Newsom Signs SB 53, Advancing California's World-Leading Artificial Intelligence Industry*, GOVERNOR CAL. (Sept. 29, 2025), https://www.gov.ca.gov/2025/09/29/governor-newsom-signs-sb-53-advancing-californias-world-leading-artificial-intelligence-industry/ [https://perma.cc/SV6T-T3GM].

[60] Transparency in Frontier Artificial Intelligence Act, sec. 2 § 22757.11(i)(1), 2025 Cal. Legis. Serv. ch. 138 (West) (to be codified at CAL. BUS. & PROF. CODE § 22757.11(i)(1)).

corresponded to roughly $100 million in model training costs.[61] Any firm using such amounts of compute will necessarily be well-capitalized. Training compute is also reasonably predictive of model performance,[62] so compute thresholds target only the most capable models reasonably well.[63] Of course, improvements in compute price-performance[64] will steadily erode the cost needed to reach any given compute threshold. This is why some proposals, like SB 1047's, use a conjunctive test, wherein regulation is only triggered if the cost to develop a covered model surpasses both an operations-based *and* a dollar-based threshold.

However, newer AI paradigms, such as reasoning models, have complicated the relationship between training compute and AI capabilities.[65] Thus, more recent proposals have also argued for regulatory targeting based not on training compute, but rather on some *entity-based* threshold, as measured by the total amount an AI company spends on AI research or compute (including both training and inference compute).[66]

III. Automated Compliance

This paper does not suggest abandoning existing approaches to managing regulatory costs in AI policy. It does, however, suggest that such approaches are insufficiently ambitious in the era of AI. In particular, we think that existing approaches

[61] *See* Anderljung et al., *supra* note 17, at 36 n.85.

[62] *See id.* appx. B; Sastry et al., *supra* note 52, at 21.

[63] It remains debated whether a focus on compute as a proxy for risk objectionably overlooks smaller models with narrower purposes that may nevertheless have similar risk profiles. *See* Letter from Rep. Zoe Lofgren et al., to Cal. Governor Gavin Newsom 2 (Aug. 15, 2024), https://democrats-science.house.gov/imo/media/doc/2024-08-15%20to%20Gov%20Newsom_SB1047.pdf [https://perma.cc/QE7R-8KV8]; Andrew W. Reddie & Kevin Frazier, *Why Context, Not Compute, Is the Key to AI Governance*, TECH POL'Y PRESS (Aug. 5, 2025), https://techpolicy.press/why-context-not-compute-is-the-key-to-ai-governance [https://perma.cc/BC4T-H4ES].

[64] *See generally* Marius Hobbhahn & Tamay Besiroglu, *Trends in GPU Price-Performance*, EPOCH AI (June 27, 2022), https://epoch.ai/blog/trends-in-gpu-price-performance [https://perma.cc/CP6V-F8YW].

[65] *See, e.g.*, Toby Ord, *Inference Scaling Reshapes AI Governance*, FORETHOUGHT (Feb. 13, 2025), https://www.forethought.org/research/inference-scaling-reshapes-ai-governance [https://perma.cc/N5E9-EXZQ]; Dean W. Ball, *An Extraordinary Alien*, HYPERDIMENSIONAL (Apr. 10, 2025), https://www.hyperdimensional.co/p/an-extraordinary-alien [https://perma.cc/S9YH-F4FN]; Venkat Somala et al., *Three Challenges Facing Compute-Based AI Policies*, EPOCH AI (Sept. 11, 2025), https://epochai.substack.com/p/three-issues-undermining-compute [https://perma.cc/44RZ-U4PR]. *But cf.* Lennart Heim, *Inference Compute: GPT-O1 and AI Governance*, BLOG - LENNART HEIM (Sept. 22, 2024), https://blog.heim.xyz/inference-compute/ [https://perma.cc/RTY9-PEHV] (arguing that the rise of reasoning models "doesn't fundamentally challenge how training compute thresholds are being used right now.").

[66] *See* Dean W. Ball & Ketan Ramakrishnan, *Entity-Based Regulation in Frontier AI Governance*, CARNEGIE ENDOWMENT FOR INT'L PEACE (July 7, 2025), https://carnegieendowment.org/research/2025/06/artificial-intelligence-regulation-united-states [https://perma.cc/2XAF-UT5W].

often ignore the extent to which AI technologies *themselves* could reduce regulatory costs by largely automating compliance tasks.[67]

Compliance professionals already report significant benefits from the use of AI tools in their work,[68] and there is no shortage of companies claiming to be able to automate core compliance tasks.[69] Such claims must of course be treated with appropriate skepticism, coming, as they do, from companies trying to attract customers. Nevertheless, the large amount of interest in existing compliance-automating AI suggests some reason for optimism.

The most significant promise for compliance automation, however, comes from future AI systems. AI companies are developing agentic AI systems: AI systems that can competently perform an increasingly broad range of computer-based tasks.[70] If these companies succeed, then AI systems will be able to autonomously perform an increasingly broad range of computer-based compliance tasks[71]—possibly more quickly, reliably, and cheaply than human compliance professionals.[72] We can call this general hypothesis—that future AI systems will be able to automate many core compliance tasks—*automated compliance*.

Automated compliance has significant implications for the proregulatory–deregulatory debate within AI policy.[73] Before exploring the implications of automated compliance, however, we should be clear about its content. Not all compliance tasks are equally automatable. Unfortunately, we cannot here provide a comprehensive account of which compliance tasks are most automatable.[74] However, we can provide some initial, tentative thoughts on which compliance tasks might be more or less automatable.

---

[67] *See* Ohm, *supra* note 20.

[68] *See, e.g.*, *Perceptions of Risk and Compliance Professionals on AI*, THOMSON REUTERS L. BLOG (Aug. 22, 2024), https://legal.thomsonreuters.com/blog/many-risk-compliance-professionals-see-ai-as-a-force-for-good-in-their-industry/ [https://perma.cc/6X97-33RF].

[69] *See* Ohm, *supra* note 20, nn.19–21 and accompanying text; *see also AI Regulatory Compliance*, TRUSTIBLE, https://trustible.ai/ai-regulatory-compliance/ [https://perma.cc/UX3Z-J3TS].

[70] *See generally* Cullen O'Keefe et al., *Law-Following AI: Designing AI Agents to Obey Human Laws*, 94 FORDHAM L. REV. 57, 66–69 (2025).

[71] *See* Ohm, *supra* note 20.

[72] Thomas Kwa et al., *Measuring AI Ability to Complete Long Tasks*, METR BLOG (Mar. 19, 2025), https://metr.org/blog/2025-03-19-measuring-ai-ability-to-complete-long-tasks/ [https://perma.cc/QT2F-MQCQ].

[73] Of course, this observation has implications for *all* areas of policy with nontrivial compliance costs, not just AI policy. However, our focus here is on how the rise of automated compliance will affect trade-offs in AI policy specifically.

[74] For prior work estimating the automatability of existing economic tasks, see generally Jonathan I. Dingel & Brent Neiman, *How Many Jobs Can Be Done at Home?*, 189 J. PUB. ECON. 104235 at 1 (2020); Tyna Eloundou et al., *GPTs Are GPTs: Labor Market Impact Potential of LLMs*, 384 SCIENCE 1306 (2024). For a proposed evaluation suite to measure model performance on economically valuable, real-world tasks, see *Measuring the Performance of Our Models on Real-World Tasks*, OPENAI (Sept. 25, 2025), https://openai.com/index/gdpval/ [https://perma.cc/3AFT-TPJ8].

To start, recall that we limited our definition of "agentic AI" to "AI systems that can competently perform an increasingly broad range of *computer-based* tasks."[75] Thus, by definition, agentic AI would only be able to automate computer-based compliance tasks; those requiring physical interactions would remain non-automatable.[76] An AI agent, under our definition, would not be able to, for example, provide physical security to a sensitive data center.[77] Fortunately, many compliance tasks required by AI policy proposals would be computer-based; hence, this definitional constraint does not itself significantly limit the implications of automated compliance for AI policy. Consider the following processes that plausible AI safety and security policies might require:

- *Automated red-teaming*, in which AI models themselves attempt to "find flaws and vulnerabilities" in another AI system.[78] Some frontier AI developers already use automated red-teaming as part of their safety workflows.[79]
- *Cybersecurity measures* to prevent unauthorized access to frontier model weights.[80] While cybersecurity has some inherently physical components (e.g., limitations on the types of hardware used, key personnel protection),[81] many components of strong cybersecurity should be highly automatable in principle.[82]
- *Implementation of some AI alignment techniques.* Some AI alignment techniques, such as reinforcement learning from human feedback,[83] require significant human input. However, more recent techniques, such as

[75] *Supra* text accompanying note 70 (emphasis added).

[76] *Cf.* O'Keefe et al., *supra* note 70, at 67 n.56 (explaining why tasks requiring physical interaction might be harder to automate than computer-based tasks).

[77] *See, e.g.*, SELLA NEVO ET AL., SECURING AI MODEL WEIGHTS: PREVENTING THEFT AND MISUSE OF FRONTIER MODELS 25 (2024), https://www.rand.org/pubs/research_reports/RRA2849-1.html [https://perma.cc/SJX9-LYL6].

[78] Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, 88 Fed. Reg. 75191 § 3(d) (Oct. 30, 2023).

[79] Lama Ahmad et al., OpenAI's Approach to External Red Teaming for AI Models and Systems (Jan. 24, 2025) (unpublished manuscript), http://arxiv.org/abs/2503.16431; Mantas Mazeika et al., HarmBench: A Standardized Evaluation Framework for Automated Red Teaming and Robust Refusal (Feb. 27, 2024) (unpublished manuscript), http://arxiv.org/abs/2402.04249.

[80] *See generally* NEVO ET AL., *supra* note 77.

[81] *See, e.g., id.* at 92–93 (describing the most rigorous security protocol).

[82] *See, e.g., id.* at 87 ("After inference, outputs are randomly modified while minimizing effects on legitimate use—for example, by adding small noise to the logprobs or running outputs through an independent autoencoder."); *id.* ("A separate component (possibly an independent AI model) is used to classify outputs that seem to be the result of malicious activity and block them from being further processed or returned to the user.").

[83] *See generally* Paul Christiano et al., Deep Reinforcement Learning from Human Preferences (Feb. 17, 2023) (unpublished manuscript), http://arxiv.org/abs/1706.03741.

constitutional AI,[84] leverage AI feedback. Research into scalable forms of AI alignment, in which AI systems themselves are charged with aligning next-generation AI systems, is ongoing.[85]

- *Automated evaluations* of AI systems on safety-relevant benchmarks.[86]
- *Automated interpretability*,[87] in which AI systems help us understand how other AI models make decisions, in human-comprehensible terms.[88]

Importantly, automated compliance extends beyond computer science tasks. Advanced AI agents could also help reduce compliance costs by, for example:

- Keeping abreast of updates in the regulatory and legal landscape.[89]
- Translating regulatory requirements into concrete changes to internal procedure.[90]
- Filing mandated transparency reports.[91]
- Designing and delivering employee training.[92]
- Corresponding with regulators.[93]

---

[84] *See generally* Yuntao Bai et al., Constitutional AI: Harmlessness from AI Feedback (Dec. 15, 2022) (unpublished manuscript), http://arxiv.org/abs/2212.08073.

[85] *See generally* Jan Leike, *A Minimal Viable Product for Alignment*, MUSINGS ON THE ALIGNMENT PROBLEM (Mar. 29, 2022), https://aligned.substack.com/p/alignment-mvp [https://perma.cc/8VS3-XUAR].

[86] *See, e.g.*, Laura Weidinger et al., Holistic Safety and Responsibility Evaluations of Advanced AI Models 7 (Apr. 22, 2024), http://arxiv.org/abs/2404.14068.

[87] On interpretability, see generally TIM RUDNER & HELEN TONER, KEY CONCEPTS IN AI SAFETY: INTERPRETABILITY IN MACHINE LEARNING (2021), https://cset.georgetown.edu/publication/key-concepts-in-ai-safety-interpretability-in-machine-learning/ [https://perma.cc/PS3V-88X2].

[88] *See, e.g.*, Gonçalo Paulo et al., Automatically Interpreting Millions of Features in Large Language Models (Dec. 4, 2024) (unpublished manuscript), http://arxiv.org/abs/2410.13928; Jan Leike et al., *Language Models Can Explain Neurons in Language Models*, OPENAI (Feb. 14, 2024), https://openai.com/index/language-models-can-explain-neurons-in-language-models/ [https://perma.cc/P3FD-7G9C]; Yoav Gur-Arieh et al., *Enhancing Automated Interpretability with Output-Centric Feature Descriptions*, 1 PROCS. 63D ANN. MEETING ASS'N FOR COMPUTATIONAL LINGUISTICS 5757 (2025), https://aclanthology.org/2025.acl-long.288/.

[89] For one company using AI systems to do this, see *Why Quorum?*, QUORUM, https://www.quorum.us/about/why-quorum/ [https://perma.cc/FNJ3-DGT8] (archived Dec. 19, 2025).

[90] For one company using AI systems to do this, see NORM AI, https://www.norm.ai/ [https://perma.cc/X384-2A57] (archived Dec. 19, 2025).

[91] On transparency requirements in AI policy, see generally Julius Hattingh, *New AI Transparency Rules Have a Trade Secrets Problem*, LAWFARE (Sept. 15, 2025), https://www.lawfaremedia.org/article/new-ai-transparency-rules-have-a-trade-secrets-problem [https://perma.cc/4LEM-2A62]. On AI-assisted regulatory filings, see, for example, Anton Mihic et al., *Faster Regulatory Submissions for Pharma with AI*, MCKINSEY & CO. (Aug. 1, 2025), https://www.mckinsey.com/industries/life-sciences/our-insights/rewiring-pharmas-regulatory-submissions-with-ai-and-zero-based-design [https://perma.cc/Q3YY-58MC].

[92] *See, e.g.,* Ian Bird, *AI for Employee Training*, IBM, https://www.ibm.com/think/insights/generative-ai-for-employee-training [https://perma.cc/WW44-TW3P] (archived Dec. 19, 2025).

[93] For one company using AI systems to do this, see *Trajan*, VULCAN TECH., https://vulcan-tech.com/products/trajan [https://perma.cc/7XLX-4C9H] (archived Dec. 23, 2025).

However, not all computer-based compliance tasks would be made much cheaper by AI agents.[94] Some compliance tasks might require *human input* of some sort, which by its nature would not be automatable.[95] For example, some forms of red-teaming "involve[] humans actively crafting prompts and interacting with AI models or systems to simulate adversarial scenarios, identify new risk areas, and assess outputs."[96] Automation may also be unable to reduce compliance costs associated with tasks that have an explicit *time requirement*. For example, suppose that a new regulation requires frontier AI developers to implement a six-month "adaptation buffer" during which they are not permitted to distribute the weights of their most advanced models.[97] The costs associated with this calendar-time requirement could not be automated away. Nevertheless, there will be some compliance tasks that will be, to some significant degree, automatable by advanced AI agents.[98]

To summarize, as AI capabilities progress, AI systems will themselves be able to perform an increasing fraction of compliance-related tasks.[99] The simplest implication of automated compliance is that, holding regulation levels constant, compliance costs should decline relative to the pre-AI era.[100] This is hardly a novel observation given the large number of both legacy firms and startups that are already integrating frontier AI technologies into their legal and compliance workflows.[101] However, we think that automated compliance has even more significant implications for the design of optimal AI policy. We turn to those implications in the next section.

[94] *See* Ohm, *supra* note 20.
[95] *See id.*
[96] Ahmad et al., *supra* note 79, at 3.
[97] The idea of adaptation buffers is first articulated in Helen Toner, *Nonproliferation Is the Wrong Approach to AI Misuse*, RISING TIDE (Apr. 5, 2025), https://helentoner.substack.com/p/nonproliferation-is-the-wrong-approach [https://perma.cc/3FL2-BQJ4]. *See also* Anderljung et al., *supra* note 17, at 14.
[98] *See* Ohm, *supra* note 20.
[99] *See id.*
[100] *See id.*
[101] *See id.*

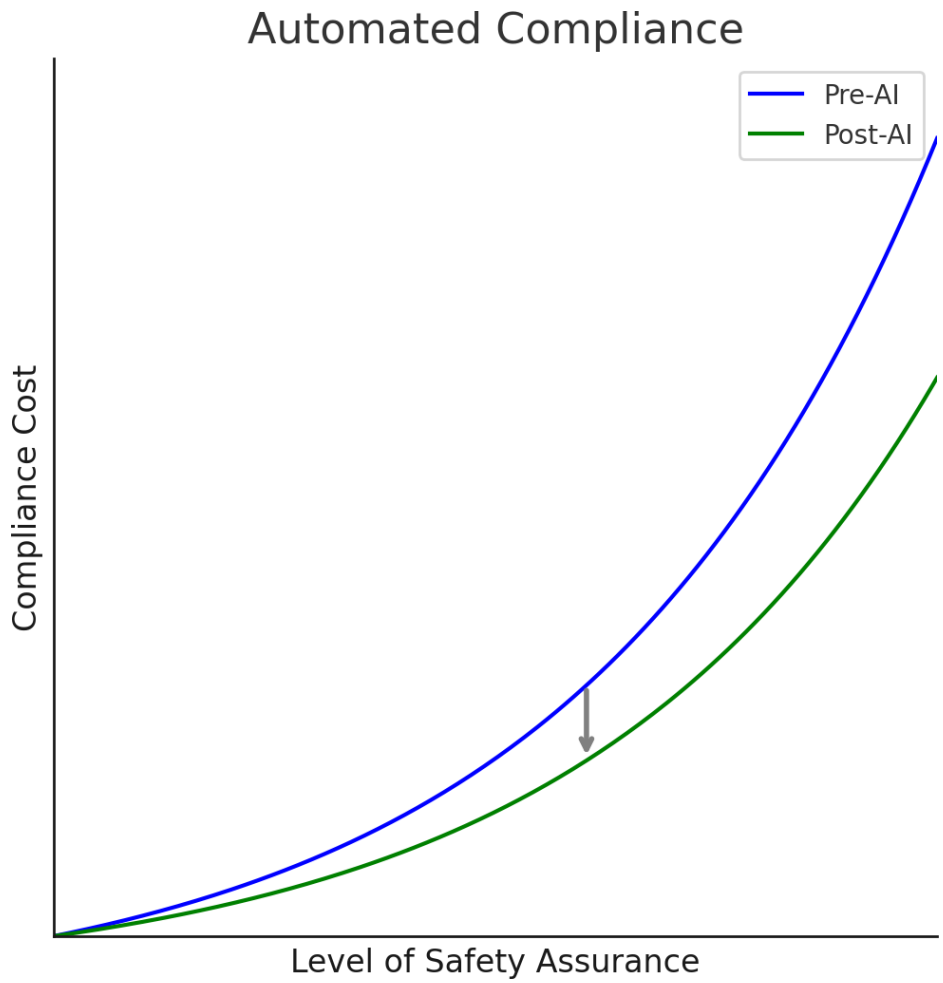

Figure 1: Automated Compliance illustrated. AI causes the cost to achieve any given level of safety assurance to decline.

IV. Implications of Automated Compliance for Policy Design

A. Automated Compliance and the Optimal Timing of Regulation

Automated compliance can reduce compliance costs associated with some forms of regulation. However, this is only true insofar as the AI technology necessary to automate compliance arrives *before* (or at least, simultaneously with) the regulatory requirements to be automated.

Thus, even those who agree with our prediction might still worry that automated compliance might become an excuse to implement costly regulation prematurely, in the expectation that technological progress will eventually reduce compliance costs. To be sure, such projections are often reasonable. For example, compliance costs for the Obama

Administration's Clean Power Plan[102] were significantly lower than initially estimated due in large part to "[o]ngoing declines in the costs of renewable energy."[103] Nevertheless, when there are genuine concerns about the downsides of premature AI regulation, or simply outstanding uncertainties over the pace or cadence of further progress in compliance-automating AI applications, merely hoping that compliance automation technologies will eventually reduce excessive compliance costs imposed today may seem like a risky proposition.

This *sequencing problem* suggests a natural solution: certain AI safety policies could only be triggered when AI technology has progressed to the point where compliance with such policies is largely automatable. We could call such legal mechanisms *automatability triggers*.

Developing statutory language for automatability triggers must be left for future work, partly because such triggers need tailoring to their broader regulatory context. However, an illustrative example may help build intuition and catalyze further refinements. Consider a bill that would impose a fine on any person who, without authorization, exports[104] the weights of a neural network if such neural network:

> (a) was trained with an amount of compute exceeding [$10 million] at fair market rates,[105] and
> (b) can, if used by a person without advanced training in synthetic biology, either:
> (i) reliably increase the probability of such person successfully engineering a pathogen by [50%], or
> (ii) reduce the cost thereof by [50%],
> in each case as evaluated against the baseline of such a person without access to such models but with access to the internet.[106]

[102] Carbon Pollution Emission Guidelines for Existing Stationary Sources: Electric Utility Generating Units, 80 Fed. Reg. 64662 (Oct. 23, 2015).

[103] DENISE A. GRAB & JACK LIENKE, THE FALLING COST OF CLEAN POWER PLAN COMPLIANCE 1 (2017), https://policyintegrity.org/files/publications/Falling_Cost_of_CPP_Compliance.pdf [https://perma.cc/ZP5P-XZK7].

[104] *See generally* Alan 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 [https://perma.cc/4973-D3QU].

[105] For similar proposals, see *supra* Part II.

[106] On this threat model, see generally MOUTON, LUCAS & GUEST, *supra note* 11. Consistent with legal drafting conventions, language in brackets is illustrative and variable depending on the policy preferences of drafters.

Present methods for assessing whether frontier AI models are capable of such "uplift" rely heavily on manual evaluations by human experts.[107] This is exactly the type of evaluation method that could be manageable for a large firm but prohibitive for a smaller firm. And although $10 million is a lot of money for individuals, it seems plausible that there will be many firms that would spend that much on compute but for whom this type of regulatory requirement could be quite costly. Under our proposed approach, the legislature might consider an automatability trigger like the following:

> The requirements of this Act will only come into effect [one month] after the date when the [Secretary of Commerce], in their reasonable discretion, determines that there exists an automated system that:
>
> (a) can determine whether a neural network is covered by this Act;
>
> (b) when determining whether a neural network is covered by this Act, has a false positive rate not exceeding [1%] and false negative rate not exceeding [1%];
>
> (c) is generally available to all firms subject to this Act on fair, reasonable, and nondiscriminatory terms, with a price per model evaluation not exceeding [$10,000]; and,
>
> (d) produces an easily interpretable summary of its analysis for additional human review.

This sample language is intended to introduce one way to incorporate an automatability trigger into law, and various considerations may justify alternative implementations or additional provisions. For example, concerns about disproportionate compliance costs being borne by smaller labs might justify a subsidy for the use of the tool. Similarly, lawmakers would need to consider whether to make the use of such a tool mandatory. Though it seems likely that most firms would prefer to adopt state-approved automated compliance tools, some may insist on doing things the "old-fashioned way." Whether that option should be available to firms will likely depend on the extent to which alternative systems would frustrate the ability of the regulator to easily assess compliance. While surely imperfect in many ways, an automatability trigger like this could probably allay many concerns about the regulatory burdens associated with our hypothetical bill.[108]

Automatability triggers could improve AI policy through two related mechanisms. First, of course, they reduce compliance costs throughout the entire time that a regulation is in force. But more importantly, they also aim to reduce the possibility of premature

[107] *See, e.g.*, *id.*; Tejal Patwardhan et al., *Building an Early Warning System for LLM-Aided Biological Threat Creation*, OPENAI (Feb. 14, 2024), https://openai.com/index/building-an-early-warning-system-for-llm-aided-biological-threat-creation/ [https://perma.cc/67M2-PZZV].

[108] We also note that this regime might entail risks of its own. For example, there is a risk that the automated evaluation system described could be used as a *reward signal* for a malicious actor who wished to intentionally train an AI system capable of engineering pathogens.

regulation: they allow AI progress to happen, unimpeded by regulation, until such time as compliance with such regulation would be much less burdensome than at present. Of course, the reverse is also true: automatability triggers might also increase the probability that AI regulations are implemented too late if the risk-producing AI capabilities arrive earlier than compliance-automating capabilities. Thus, the desirability of automatability triggers depends sensitively on policymakers' preferences over regulating too soon or too late.

Automatability triggers also have key benefits over the primary approaches to controlling compliance costs within AI policy proposals: compute thresholds and monetary thresholds.[109] Existing approaches tend not to (directly)[110] control *absolute* costs of compliance, but rather tend to ensure that compliance costs are only borne by well-capitalized firms. Automatability triggers, by contrast, aim to cap compliance costs for all regulated firms.

The prospective enactment of automatability triggers in regulations also sends a useful signal to AI developers: it makes clear that there will be a market for compliance-automating AI and therefore incentivizes development towards that end. This, in turn, implies that trade-offs between safety regulation and compliance costs would loosen much more quickly than by default.

Finally, regulations that incorporate automatability triggers may prove far more adaptable and flexible than traditional, static regulatory approaches—a quality that most experts regard as essential for effective AI governance.[111] Traditional rules often struggle to keep pace with rapid technological change, requiring lengthy amendment processes whenever new risks or practices emerge. By contrast, automatability triggers allow regulations to evolve in step with the development of compliance technologies. As automated compliance tools become more sophisticated, legislators and regulators could use them to focus on the precise types of information from labs that are most relevant to the regulatory issue at hand, rather than demanding broad, costly disclosures. This targeted approach not only reduces unnecessary burdens on regulated entities but also increases the likelihood that regulators receive timely, actionable data. Importantly, amendments to laws designed with automatability triggers would not require firms to reinvent their compliance systems from the ground up. Instead, updates would simply involve ensuring that the relevant information is transmitted through the approved compliance tools—making the regulatory framework more resilient, responsive, and sustainable over time.

Although the idea of automatability triggers is straightforward, their design might not be. Policymakers would need to be able to define and measure the cost-reducing potential of AI technologies. This seems difficult to do even in isolation; ensuring that such

[109] *See supra* Part II.
[110] Of course, firms always have an incentive to reduce compliance costs.
[111] BOMMASANI ET AL., *supra* note 6, at 3–4, 14.

measures accurately predict the cost-savings realizable by regulated firms seems more difficult still.

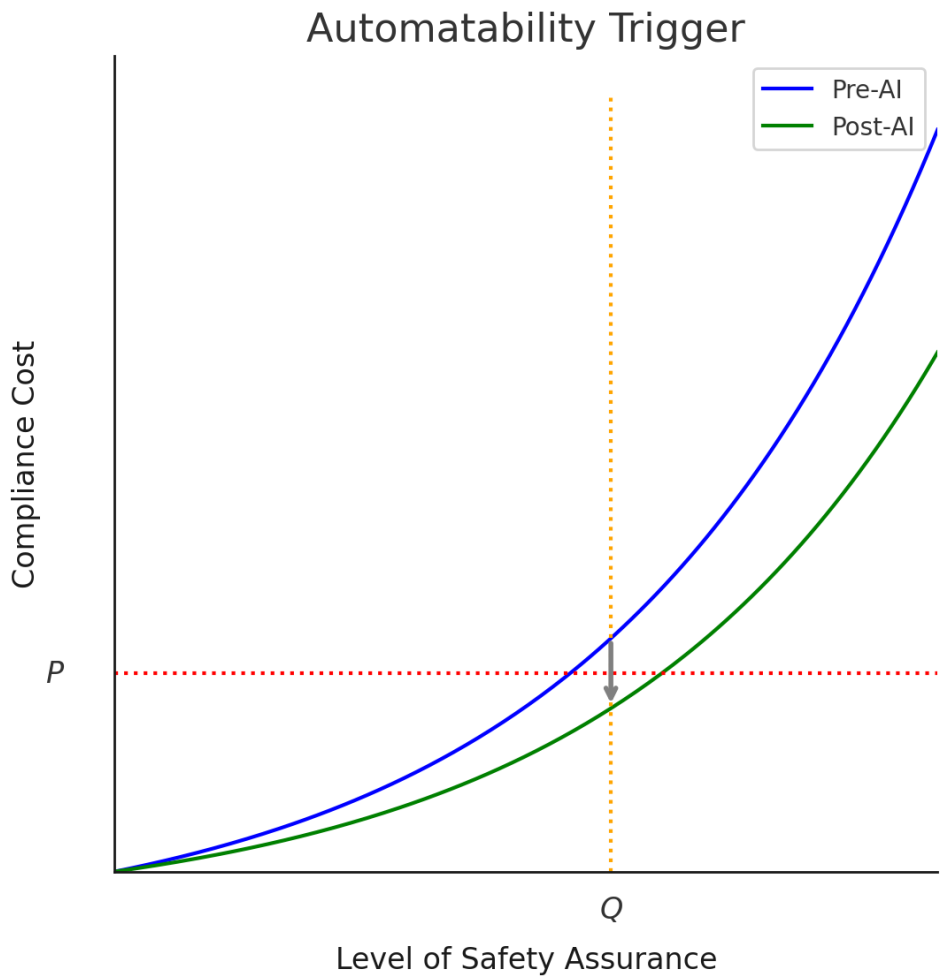

Figure 2: Automatability Triggers illustrated. A regulation that provides safety assurance level *Q* is implemented with an automatability trigger at *P*: the regulation is only effective when the cost to implement the regulation falls below *P*. Regulated parties are thus guaranteed to always have compliance costs below *P*.

B. Adoption of Compliance-Automating AI as Evidence of Compliance

Automatability triggers assume that regulators can competently identify AI services that, if properly adopted, enable regulated firms to comply with regulations at an acceptable cost. If so, then we might also consider a legal rule that says that regulated firms that properly implement such "approved" compliance-automating AI systems are presumptively (but rebuttably) entitled to some sort of preferential treatment in regulatory enforcement actions. For example, such firms might be inspected less frequently or less

invasively than firms that have not implemented approved compliance-automating AI systems. Or, in enforcement actions, they might be entitled to a rebuttable presumption that they were in compliance with regulations while using such systems.[112] Or a statute might provide that proper adoption of such a system is conclusive evidence that the firm was exercising reasonable care so as to preclude negligence suits.

Of course, it is important that such safe harbors be carefully designed. They should only be available to firms that were *properly* implementing compliance-automating AI. This would also mean denying protection to firms who, for example, provided incomplete or misleading information to the compliance-automating AI or knowingly manipulated it, causing it to falsely deem the firm to be compliant. Compliance-automating AIs, in turn, should ideally be robust to such attempts at manipulation. Regulators would also need to be confident that, if properly implemented, compliance-automating AI systems really would achieve the desired safety results in deployment settings. Finally, in the ideal case, regulators would ensure that the market for compliance-automating AI services is competitive; reliance on a small number of vendors who can set supracompetitive prices would reduce the cost-saving potential of compliance-automating AI and concentrate its benefits among the firms with deeper pockets. For example, perhaps regulators could accomplish this by only implementing such an approval regime if there were multiple compliant vendors.[113]

C. Differential Acceleration of Automated Compliance

Automated compliance can be analyzed through the lens of "risk-sensitive innovation": a strategy for deliberately structuring the timing and order of technological advances to "reduce specific risks across a technology portfolio."[114] Targeted acceleration of the development of compliance-automating AI systems could reduce painful trade-offs between safety and innovation in a very fraught and uncertain policy environment. It is therefore worth considering what, if anything, AI policy actors can do to differentially accelerate automated compliance.

---

[112] By analogy, parties subject to specific obligations under the EU AI Act may rely on "codes of practice" authorized by the Act to "demonstrate compliance with" those obligations. *E.g.*, EU AI Act, *supra* note 53, art. 55(2).

[113] This has similarities to the regulatory market proposals in Gillian K. Hadfield & Jack Clark, Regulatory Markets: The Future of AI Governance (Apr. 25, 2023) (unpublished manuscript), http://arxiv.org/abs/2304.04914; Dean W. Ball, *Putting Private AI Governance into Action*, HYPERDIMENSIONAL (Apr. 10, 2025), https://www.hyperdimensional.co/p/putting-private-governance-into-action [https://perma.cc/EY6Y-WR5D].

[114] Jonas B. Sandbrink et al., *Risk-Sensitive Innovation: Leveraging Interactions between Technologies to Navigate Technology Risks*, 51 SCI. PUB. POL'Y 1028, 1028 (2024); *see also* LIZKA VAINTROB & OWEN COTTON-BARRATT, AI TOOLS FOR EXISTENTIAL SECURITY (2025), https://www.forethought.org/downloads/ai-tools-for-existential-security.pdf [https://perma.cc/JH8E-6HMU].

Of course, there will be a natural market incentive to develop such technologies: regulated parties will need to comply with regulations and will be willing to pay anyone that can reduce the costs of compliance. Indeed, we have mentioned some examples of how firms are already using AI to reduce compliance costs.[115] But prosocial actors may be able to further accelerate automated compliance by, for example:[116]

- Building curated data sets that would be useful for creating compliance-automating AI systems.[117]
- Subsidizing compute access for developers attempting to build compliance-automating AI systems.[118]
- Building proof-of-concept compliance-automating AI systems for existing regulatory regimes.
- Preferentially developing and advocating for AI policy proposals that are likely to be more automatable.
- Educating policymakers on which types of regulations are more likely to be compatible with regulatory compliance.
- Instituting monetary incentives, such as advance market commitments, for compliance-automating AI applications.[119]
- Ensuring that firms working on automated compliance have early access to restricted AI technologies.[120]
- Building AI systems that automate key technical AI safety tasks that might be required by AI safety regulation, such as those listed above.[121]
- Building AI systems that automate key non-technical legal and compliance tasks, such as those listed above.[122]

D. Automated Compliance Meets Automated Governance

So far, we have been focusing on the costs and benefits to regulated parties. However, automated compliance might be especially synergistic with automation of core

[115] *See supra* notes 68–69, 79.
[116] This list is heavily inspired by VAINTROB & COTTON-BARRATT, *supra* note 114, at 4.
[117] *See id.*
[118] *See id.*; *see also* Sastry et al., *supra* note 52, at 43–44.
[119] *See generally, e.g.*, Michael Kremer, Jonathan Levin & Christopher M. Snyder, *Advance Market Commitments: Insights from Theory and Experience*, 110 AEA PAPERS & PROCS. 269 (2020). For an initiative to establish similar advance market commitments, see FORMAL MARKETS, https://www.marketshapingai.org/ [https://perma.cc/5WLQ-TAU6] (archived Dec. 19, 2025).
[120] *See* Toner, *supra* note 97.
[121] *See supra* Part III.
[122] *See supra* Part III.

regulatory and administrative processes.[123] For example, regulatory AI systems would be well-positioned to know how proposed regulations would affect regulated companies and could therefore be used to write responses to proposed rules.[124] Regulatory AI systems, in turn, could compile and analyze these comments.[125] Indeed, AI systems might be able to draft and analyze many more variations of rules than human-staffed bureaucracies could,[126] thus enabling regulators to receive and review in-depth, tailored responses to many possible policies and select among them more easily.

Compliance-automating AI systems could also request guidance from regulatory AI systems, who could review and respond to the request nearly instantaneously.[127] Such guidance-providing regulatory AI systems could be engineered to ensure that business information disclosed by the requesting party was stored securely and never read by human regulators (unless, perhaps, such materials became relevant to a subsequent dispute), thus reducing the risk that the disclosed information is subsequently used to the detriment of the regulated party.

Of course, there are many governance tasks that should remain exclusively human, and automating core governance tasks carries its own risks.[128] But future AI systems could offer significant benefits to both regulators and regulatees alike, and may be even more beneficial still when allowed to interact with each other according to predetermined rules designed to mitigate the potential for abuse by either party.

---

[123] On the latter, see generally Cary Coglianese, *Administrative Law in the Automated State*, 150 DAEDALUS 104 (2021); BRYAN L. FRYE, ARTIFICIAL REGULATION (2025), https://cdn.sanity.io/files/d8lrla4f/staging/2dd465fb68e3e10c094a567ba92630194359d090.pdf [https://perma.cc/9D5G-6GSN]; Nicholas Caputo, (When) Should We Delegate AI Governance to AIs? Some Lessons from Administrative Law (Sept. 24, 2025) (unpublished manuscript), http://arxiv.org/abs/2509.22717.

[124] For discussion of use of generative AI to draft comments in notice-and-comment rulemaking, see, for example, Stephen M Johnson, *Rulemaking 3.0: Incorporating AI and ChatGPT into Notice and Comment Rulemaking*, 88 MISS. L. REV. 1021, 1053–69 (2023).

[125] *See id.* pt. VI.C.

[126] *See* John J. Nay & Troy A. Paredes, *How Regulators Can Use AI*, 71 VAND. L. REV. EN BANC 63, 67–69 (2025) (discussing using AI systems to generate and analyze variations of proposed regulations).

[127] *Cf. id.* at 69–71 ("To promote innovation and commerce, regulators could develop and deploy AI agents that businesses could use as a form of automated self-service preclearance. . . . An application of this capability could enable regulated entities to request preclearance review of regulated (or potentially regulated) activity from government-hosted regulatory-AI agents built to assess regulatory compliance—in other words, AI systems that would indicate whether the proposed activity being reviewed complies with the relevant regulation. Businesses or individuals, if they chose to, could engage with these AI agents to receive interactive and explanatory responses based on complete and accurate information inputted into the model. Within minutes, AI agents could provide feedback indicating if the activity is approved, if more information is needed, if the outcome is uncertain and requires additional human perspective and judgment, or if the activity raises concerns and might need to be adjusted to comply." (footnote omitted)).

[128] *See generally* Caputo, *supra* note 123.

Conclusion

AI systems are capable of automating an increasingly broad range of tasks. Many regulatory compliance tasks will be similarly automatable. This insight has important implications for the ongoing debate about whether and how to regulate AI. On the one hand, forecasts of regulatory compliance costs will be overstated if they fail to account for this fact; AI progress *itself* hedges the costs of many forms of AI regulation. Regulatory design should account for this dynamic. However, to maximize the benefits of automated compliance, regulators must successfully navigate a tricky sequencing problem. If regulations are triggered too soon—that is, before compliance costs have fallen sufficiently—they will hinder desirable forms of AI progress. On the other hand, if they are triggered too late—that is, after the risks from AI would justify the regulations—then the public may be exposed to excessive risks from AI. Smart AI policy must be attentive to these dynamics.

To be clear, our claim is fairly modest: AI progress will reduce compliance costs in some cases. Automated compliance is only relevant when compliance tasks are in fact automatable, and not all compliance tasks will be. Accordingly, the costs of some forms of AI regulation might remain high, even if many compliance tasks are automated. And of course, regulations are only justified when their expected benefits outweigh their expected costs. Furthermore, regulations have costs in excess of their directly measurable compliance costs;[129] these costs are no less real than are compliance costs. The availability of compliance-automating AI should not be used as an excuse to jettison careful analysis of the costs and benefits of regulation. Nevertheless, AI policy discourse should internalize the fact that AI progress implies reduced compliance costs, all else equal, due to automated compliance.

[129] *See supra* Part I.